BMJ Open

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: an application comparison between logistic regression model and generalized partially linear tree-based regression model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014246
Article Type:	Research
Date Submitted by the Author:	20-Sep-2016
Complete List of Authors:	Wang, Yuyan; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Zhang, Biao; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hou, Lei; Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, , Department of Epidemiology and Biostatistics Han, Wei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Xue, Fang; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Yanhong; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Tang, Yong; The People's Hospital in Altay Region Liang, Shaohua; The People's Hospital in Altay Region Wang, Weizhi; The People's Hospital in Altay Region Asaiti, Kuliqian; Hong Dun Town hospital, Altay Region Wang, Zixing; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hu, Yaoda; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Lei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Qiu, Changchun; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences Zhang, Mingtao; The People's Hospital in Altay Region Jiang, Jingmei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods, Genetics and genomics
Keywords:	Hypertension < CARDIOLOGY, Gene environment interaction, ACE genotype, Salt intake, Logistic model, GPLTR model

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: an application comparison between logistic regression model and generalized partially linear tree-based regression model

Yuyan Wang[1†], Biao Zhang[1†], Lei Hou[1], Wei Han[1], Fang Xue[1], Yanhong Wang[1], Yong Tang[2], Shaohua Liang[2], Weizhi Wang[2], Kuliqian.Asaiti[3], Zixing Wang[1], Yaoda Hu[1], Lei Wang[1], Changchun Qiu[1], Mingtao Zhang[2*], Jingmei Jiang[1*]

[1] *Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, China*

[2] *The People's Hospital in Altay Region, Xinjiang, China*

[3] *Hong Dun Town hospital, Altay Region, Xinjiang, China*

[†] These authors contributed equally to this article.

[*] Corresponding Authors:

Prof. Jingmei Jiang, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, 5 Dong Dan San Tiao, Beijing, 100005, China. E-mail: jingmeijiang@ibms.pumc.edu.cn. Tel: +86 10 69155935, Fax: +86 10 65123696.

Mr. Mingtao Zhang. The People's Hospital in Altay Region, Xinjiang, China. E-mail: altrmyy@163.com. Tel:+86018896519163.

ABSTRACT

Objective: To explore the interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs, and meanwhile to compare the interaction application between logistic model and GPLTR model.

Design: Population-based cross-sectional study.

Setting: Hong Dun town in the North Xinjiang, China.

Participants: 916 Chinese Kazakh participants without consanguinity (342 men and 574 women) aged 30 years or over.

Main outcome measures: Association between ACE gene and hypertension, association between salt intake and hypertension, and interaction of ACE genotype and salt intake on hypertension in two models.

Results: Associations between salt intake and hypertension were diverse stratified by ACE genotype as II and ID+DD. For logistic models, main and interaction effects were not observed for men, but in opposite directions for women (main effect of ACE: OR=0.20, p=0.003; interaction effect: OR=1.07, p=0.027). For GPLTR models, BIC trees included both salt intake and ACE genotype as split variables, Salt intake≥19.5g/day and ID+DD genotypes had 3.99 times (p=0.004) higher risk for hypertension compared with II genotype for men, and salt intake<20.1g/day and ID+DD genotypes had an OR=0.55 (p=0.014) compared with II genotype for women.

Conclusions: Interaction of ACE genotype and salt intake on hypertension was observed among Chinese Kazakhs but in different ways by gender, and GPLTR model might be more suitable interaction exploration of complex diseases.

Strengths and limitations of this study

1. Provide evidences that association of ACE gene polymorphism with hypertension was sex-specific.

2. Interaction of ACE genotype and salt intake on hypertension was observed among Chinese Kazakhs but in different ways by gender.

3. Provide an important clues for the candidate gene–environment interaction exploration in epidemiology, which might overthrow the existing logistic regression methods.

4. Salt intake was assessed by urinary sodium excretion and might be underestimated because salt can be removed from the body through other means, for example, sweat.

INTRODUCTION

Hypertension is a complex disorder resulting from an interaction of genetic and environmental factors,[1] and blood pressure (BP) is known to be regulated by the renin–angiotensin–aldosterone system (RAAS).[2] Salt intake is the most influential dietary factor for the RAAS, and has been recognized as the most important environmental contributor to hypertension.[3] For genetic influence, the insertion/deletion polymorphism of the angiotensin I-converting enzyme (ACE) gene is the most common genetic variation in RAAS,[4] and its association with blood pressure has been reported to be dependent of gender and race/ethnicity.[4-7] In the meantime, the presence of interaction of ACE genotype and dietary salt intake have been supported by some researches,[8,9] however, there are few reports on the interaction effect of them on hypertension in different populations.[10]

The measurement and explanation of candidate gene–environment interaction (cG×E) in populations have been explored for a long time,[11] and it is generally considered that statistical concerns that arise with modeling cG×E are frequently overlooked.[12] Popular logistic regression models with interaction terms can be problematic in many cases.[13,14] The newly introduced generalized partially linear tree-based regression (GPLTR) model, which combines tree structure and linear adjustment, may have more flexibility and power in the process of interaction exploration for complex diseases.[15]

The aims of this study was to explore the interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs, whose high prevalence of hypertension and dietary intake of salt are quite impressive, and meanwhile to compare the interaction application between logistic model and GPLTR model.

METHODS

Study subjects

The designing scheme, sampling process and data collection of the Chinese Xinjiang Altay Kazakh Heart Study (XAKHS) have been presented elsewhere.[16,17] Hong Dun town is the base of this study locating in the urban-rural fringe of Altay county-level city in the North Xinjiang, China. Subjects were required to have at least three paternal/maternal biological generations living in the same region without history of intermarriage. From October 2012 to February 2013, a total of 1805 people aged 30 years or over were recruited, and 1668 participants from 601 families completed the baseline survey and examination under standard procedures. In this study, spouses without consanguinity and one member of biological relatives from each family were extracted to ensure the independence (figure 1). Approval for the study was given by the Ethics Committee of Institute of Basic Medical Sciences Chinese Academy of Medical Sciences and written informed consent was obtained from all participants.

Study variables

Blood pressure was measured with an appropriate arm cuff after a resting period of at least 10 minutes in the supine position, which used the same standard procedure and calibrated mercury sphygmomanometers by trained and certified staff, and was defined as the mean of two readings. A third measurement was taken if a difference of more than 5 mmHg was observed between the first and second measurement, and then the final reading was the mean of all three measurements. Subjects were classified as hypertensive if they had a systolic BP ≥ 140 mmHg, and/or a diastolic BP ≥ 90 mmHg, or both, or the use of antihypertensive

medications within the last 2 weeks.

Venous blood samples were drawn after an overnight fast of at least 10 hours and centrifuged within 20 minutes, then divided into two, one for biochemistry assays performed within 4 hours, and the another stored in a portable refrigerator at -20℃and transferred as soon as possible to Beijing, where they were stored at -80℃ for genetic analysis. The D and I alleles were identified on the basis of polymerase-chain-reaction (PCR) amplification of the respective fragments from intron 16 of the ACE gene and size fractionation and visualization by electrophoresis. The ACE genotype was classified as II, ID or DD depending on each allele with 99.76% genotyping success rate.

The second urine sample after waking was collected, and urinary sodium was measured with a Caretium XI-921CT automatic electrolyte analyzer (Shenzhen, China). The salt intake was assessed by urinary sodium excretion per day from the amounts of urinary sodium, urine creatinine (Cr), and the 24-h urinary Cr excreted as estimated from height, body weight and age,[18] for which each 43 mmol of sodium is approximately equivalent to 1 g of sodium or 2.5 g of salt (sodium chloride).

Other covariates were also collected in the standard procedure. Demographic information (including age, gender and occupation), diet, alcohol consumption, cigarette smoking and medical history were obtained from a face-to-face questionnaire survey. Body weight and height were measured in light indoor clothes and without shoes. Body mass index (BMI) was calculated as body weight in kilograms divided by height in meters squared (kg/m^2). Fasting blood glucose (FBG) was collected according to standardized protocols by a BECKMAN COULTER AU2700 clinical chemistry analyzer (CA, USA).

Statistical analysis

All analyses were conducted separately by gender. Basic characteristics were compared with regard to the status of hypertension. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the associations of ACE polymorphism with hypertension. To examine the presence of interaction, associations of salt intake with hypertension were illustrated stratified by the ACE genotype as II and ID+DD. To compare the effect of interaction, two models were applied. On the one hand, logistic regression model was built, on the other hand, GPLTR models adjusting for the same covariates were conducted to extract the high-risk group of hypertension using Akaike information criterion (AIC) and Bayesian information criterion (BIC). Description and preliminary analyses were carried out by SAS Version 9.3 (SAS Institute, Cary, NC, USA), and GPLTR models were conducted by R Version 3.1.3 with 'GPLTR' package. All *p*-values were two sided except *p* trend tests based on logistic model, in which one-sided *p* values were used.

RESULTS

This study enrolled 916 independent participants (342 men and 574 women) who had complete data for all study variables. The prevalence of hypertension was 49.4% for men and 35.9% for women. Participant characteristics are shown in table 1. Briefly, participants with hypertension had increased age, body mass index (BMI) and fasting blood glucose (FBG) compared with normotensive participants for both sexes.

Table 1 Basic characteristics of enrolled 916 subjects by gender and hypertension status

	Men (N=342)			Women (N=574)		
	Hypertension (n=169)	Normotension (n=173)	*p*	Hypertension (n=206)	Normotension (n=368)	*p*
Age, years	50.1 ± 12.1	42.8 ± 10.5	<0.001	51.0 ± 11.2	39.6 ± 7.8	<0.001
Occupation, n (%)						
Nomad	68 (40.2)	67 (38.7)	0.900	87 (42.2)	114 (31.0)	0.001
Farmer	75 (44.4)	81 (46.8)		94 (45.6)	169 (45.9)	
City worker	26 (15.4)	25 (14.5)		25 (12.2)	85 (23.1)	
BMI, kg/m^2	26.9 ± 4.8	24.6 ± 3.4	<0.001	28.7 ± 5.4	25.8 ± 4.5	<0.001
Smoking, n (%)	116 (69.8)	133 (76.9)	0.140	12 (5.8)	11 (3.0)	0.097
Drinking, n (%)	53 (31.4)	38 (22.0)	0.049	0 (0.0)	0 (0.0)	—
SBP, mm Hg	152.3 ± 19.2	122.2 ± 9.5	<0.001	154.1 ± 21.7	117.8 ± 9.8	<0.001
DBP, mm Hg	95.7 ± 11.2	77.9 ± 7.2	<0.001	94.1 ± 11.4	75.5 ± 7.7	<0.001
Family history of hypertension, n (%)	92 (57.1)	94 (56.3)	0.876	116 (60.1)	213 (58.2)	0.663
FBG, mmol/L	5.4 ± 0.8	5.2 ± 0.4	<0.001	5.3 ± 0.8	5.0 ± 0.5	<0.001
Salt intake, g/day	19.5 ± 26.7	18.3 ± 15.9	0.599	17.1 ± 7.4	15.9 ± 9.7	0.129

BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; FBG: fasting blood glucose.
Values are means ± SD and frequency (percent).

Independent associations

The independent associations of the ACE genotype and salt intake with hypertension were shown in table 2. D allele had a protective effect on hypertension in women, with DD genotype having an OR=0.47 (95% CI: 0.25-0.88) compared with the II genotype. For men, there was not the same protective effect, with the DD genotype having an OR=1.27 (95% CI: 0.66-2.46) compared with the II genotype. Salt intake by quartiles did not correlate with hypertension for either sex (*p* for trend=0.095 for men and 0.152 for women).

Table 2 Association of hypertension with ACE Genotype and salt intake

	ACE Genotype				Quartiles of salt intake with median (interquartile range), g/day				
	II (n=356)	ID (n=399)	DD (n=161)	*p*	Q1	Q2	Q3	Q4	*p* for trend
Men	128	148	66		11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
HTN, %	46.9	49.3	54.6		48.8	47.1	44.2	57.7	
Crude OR (95% CI)	1.00	1.10 (0.69, 1.77)	1.36 (0.75, 2.47)	0.327	1.00	0.93 (0.51, 1.70)	0.83 (0.46, 1.51)	1.43 (0.78, 2.61)	0.334
Age-adjusted OR (95% CI)	1.00	1.08 (0.66, 1.78)	1.31 (0.70, 2.46)	0.416	1.00	0.95 (0.50, 1.79)	0.88 (0.47, 1.66)	1.79 (0.95, 3.39)	0.105
Multi-adjusted OR*(95% CI)	1.00	1.02 (0.60, 1.72)	1.27 (0.66, 2.46)	0.522	1.00	0.94 (0.48, 1.83)	0.90 (0.46, 1.74)	1.88 (0.96, 3.71)	0.095
Women	228	251	95		10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
HTN, %	39.9	35.1	28.4		28.5	36.4	35.0	43.8	
Crude OR (95% CI)	1.00	0.81 (0.56, 1.18)	0.60 (0.36, 1.00)	0.048	1.00	1.44 (0.87, 2.36)	1.35 (0.82, 2.23)	1.95 (1.20, 3.19)	0.013
Age-adjusted OR (95% CI)	1.00	0.71 (0.46, 1.09)	0.50 (0.27, 0.92)	0.017	1.00	1.59 (0.89, 2.85)	1.25 (0.69, 2.26)	1.87 (1.06, 3.30)	0.073
Multi-adjusted OR* (95% CI)	1.00	0.69 (0.44, 1.09)	0.47 (0.25, 0.88)	0.013	1.00	1.39 (0.75, 2.55)	1.13 (0.60, 2.13)	1.68 (0.92, 3.06)	0.152

HTN: hypertension prevalence;

*Men adjusted for age, body mass index, drinking, and fasting blood glucose; Women adjusted for age, body mass index, occupation and fasting blood glucose.

Interaction effect

Stratified analysis

Associations between salt intake and hypertension stratified by ACE II genotype and ID+DD genotypes were evaluated (table 3). No correlation was observed in II genotype group for either men (p for trend=0.967) or women (p for trend=0.904). An increase in hypertension with higher salt intake was observed in the ID+DD genotypes group for both men (p for trend=0.028) and women (p for trend=0.045).

Logistic model

Results from the logistic regression model adjusted for multiple confounders were shown in table 4. For men, main effects of the ACE genotype (OR=1.33, p=0.426) and salt intake (OR=1.00, p=0.415) were not observed, and there was no interaction effect (OR=0.99, p=0.464). For women, main effect of salt intake (OR=0.99, p=0.523) was not observed, but main effect of the ACE genotype (OR=0.20, p=0.003) and interaction effect (OR=1.07, p=0.027) were observed to be present in opposite directions.

GPLTR model

All iterative procedures converged after 15 iterations, and trees selected by AIC and BIC with different degrees of pruning resulted in 10 and seven leaves for men, and nine and three leaves for women, respectively (figure 2). Considering for the stability, BIC trees were chosen for the final interpretation. Risk assessments of trees were shown in table 4. For men, leaves were identified from left as:

(1) Node 1: 57 participants characterized by salt intake ≥ 19.5 grams per day (g/day) and II genotype for ACE;

(2) Node 2: 40 participants characterized by salt intake ≥ 19.5 g/day and ID+DD genotype for ACE;

(3) Node 3: 127 participants characterized by 15.3 g/day ≤ salt intake < 19.5 g/day;

(4) Node 4: 13 participants characterized by 14.6 g/day ≤ salt intake < 15.3 g/day;

(5) Node 5: 57 participants characterized by 12.2 g/day ≤ salt intake < 14.6 g/day;

(6) Node 6: 17 participants characterized by salt intake < 9.6 g/day;

(7) Node 7: 31 participants characterized by 9.6 g/day ≤ salt intake < 12.2 g/day.

Split variables for the selected tree included both salt intake and ACE genotype as II and ID+DD. After adjusting for the same confounders in the multivariable logistic model, the final tree showed that Node 2 expressed 3.99 times (p=0.004) higher risk for hypertension compared with Node 1, and node 3 had an OR=0.93 (p=0.860). For women, leaves were identified from left as:

(1) Node 1: 185 participants characterized by salt intake < 20.1 g/day and II genotype for ACE;

(2) Node 2: 288 participants characterized by salt intake < 20.1 g/day and ID+DD genotype for ACE;

(3) Node 3: 101 participants characterized by salt intake ≥ 20.1 g/day.

Split variables also included both salt intake and ACE genotype as II and ID+DD. Results showed that Node 2 had an OR=0.55 (p=0.014) and Node 3 had OR=1.26 (p=0.444) compared with Node 1.

Table 3 Multivariable-adjusted odds ratios (95% confidence interval) in the ACE genotypes of II and ID+DD by different salt intake level in men and women

Multi-adjusted OR* (95% CI)	Quartiles of salt intake with median (interquartile range), g/day				
	Q1	Q2	Q3	Q4	*p* for trend
Men	11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
II (n=128)	1.00	2.04 (0.66, 6.30)	1.59 (0.51, 5.01)	1.09 (0.37, 3.19)	0.967
ID+DD (n=214)	1.00	0.59 (0.25, 1.41)	0.68 (0.30, 1.58)	3.15 (1.25, 7.93)	0.028
Women	10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
II (n=228)	1.00	1.19 (0.47, 2.99)	0.75 (0.28, 2.04)	1.15 (0.44, 3.00)	0.904
ID+DD (n=346)	1.00	1.55 (0.67, 3.59)	1.53 (0.66, 3.52)	2.35 (1.06, 5.24)	0.045

*Men adjusted for age, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, body mass index, occupation and fasting blood glucose.

Table 4 ACE gene and salt intake interaction explorations by logistic model and partial linear tree-based regression model

Logistic model			PLTR model		
Terms in the model	Multi-adjusted OR* (95% CI)	*p*	Nodes in the model	Multi-adjusted OR* (95% CI)	*p*
Men			Men		
Salt intake, g/day	1.00 (0.99, 1.03)	0.415	Node 1: Salt intake ≥ 19.5 and ACE(II)	1.00	
ACE(ID+DD) vs. ACE(II)	1.33 (0.66, 2.65)	0.426	Node 2: Salt intake ≥ 19.5 and ACE(ID+DD)	3.99 (1.55, 10.26)	0.004
ACE(ID+DD)×Salt intake vs. ACE(II)×Salt intake	0.99 (0.96, 1.02)	0.464	Node 3: 15.3 ≤ Salt intake < 19.5	0.93 (0.41, 2.09)	0.860
Women			Women		
Salt intake, g/day	0.99 (0.94, 1.03)	0.523	Node 1: Salt intake < 20.1 and ACE(II)	1.00	
ACE(ID+DD) vs. ACE(II)	0.20 (0.07, 0.59)	0.003	Node 2: Salt intake < 20.1 and ACE(ID+DD)	0.55 (0.34, 0.89)	0.014
ACE(ID+DD)×Salt intake vs. ACE(II)×Salt intake	1.07 (1.01, 1.13)	0.027	Node 3: 20.1 ≤ Salt intake	1.26 (0.70, 2.25)	0.444

*Men adjusted for age, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, body mass index, occupation and fasting blood glucose.

DISCUSSION

There are three main findings from the study. First, we observed that association between ACE gene polymorphism and hypertension was sex-specific. Second, stratified analysis indicated that the interaction of ACE genotype and salt intake existed for both sexes. Third, application comparison of the two models suggested that logistic model were unsound, and that the GPLTR model could effectively determine the high-risk characteristics of hypertension.

Comparison with previous findings

The distribution of the ACE genotype observed in this study is similar to that reported for an isolated group of Kazakh in 2003,[19] with the frequency of DD being less than 20%, which is a large difference compared with other studies on the Kazakhs [20] and with studies of other ethnicities.[4-7] It is suggested that the difference may arise from diverse structure of the population and racial heterogeneity.

The results of the current study are consistent with previous studies that have found that the ACE genotype is a sex-specific candidate gene for hypertension.[5,6,21,22] While there is not a positive association between the D allele and hypertension in males, there is a protective effect of the D allele for women, presumably through different ethnicities and other factors including gene-gene or gene-environment interactions.[6]

With regard to the interaction of ACE genotype and salt intake on hypertension, a study that examined 284 Japanese men using logistic regression analysis [10] reported a negative association for ACE (II+ID vs. DD: OR=0.1, p=0.024) and a positive association for interaction (ACE/II+ID × salt intake level vs. DD × salt intake level: OR=3.6, p=0.047) in

opposite directions, which was similar to the results in the current study for when the II genotype was set as the reference.

Possible mechanisms

There is strong evidence from association and linkage studies to suggest that the ACE gene had a profound effect on the variance of plasma levels of ACE,[23] the function of which is to maintain blood pressure homeostasis. Among the major factors that activate RAAS, dietary sodium/potassium balance is the most important, and is mainly influenced by dietary salt intake.[3] It is reported that sodium status affects the phenotype of ACE I/D polymorphism, and enhances blood pressure and renal function responses, and that conversion of aldosterone to angiotensin I with the DD genotype are blunted by low sodium intake.[8] It is collectively suggested from above reports that the ACE polymorphism interacts with dietary salt intake, with subsequent effects on the regulation of blood pressure.

Gene-targeting experiments with inactivation of the ACE gene in mice have provided evidence for the sex-specific effect of the ACE gene on blood pressure.[24] Population-based research also supports the proposition that the ACE gene influences blood pressure variability in a sex-specific manner.[4,25] Moreover, the diet habits of subjects were also assessed in current study (data not shown). Average dietary salt intake was determined as 17.3 g/day, which is far greater than other populations and World Health Organization recommendations.[26] It is suggested that, as the daily salt intake for this population is of such a high level, the potential harmful effects of salt intake for Chinese Kazakhs during hypertension development should be emphasized and care should be taken for the explanation. For men, from previous studies on the relationship between the ACE gene and salt sensitivity,

it is thought that D allele may lead to an increased response to a high salt intake,[4,27] thus carriers of the D allele may trigger hypertension with a high salt diet. For women, the mechanism behind the interaction is uncertain. It is, however, thought that the combination of the D allele and a lower salt intake diet may have a synergistic effect on hypertension.

Exploration of the interaction

Discussion on the gene–environment interactions have progressed through three main stages, including eugenics, IQ, and serotonin transporter gene controversy,[28] and has resulted in an apparent divide between variation-partitioning approaches and mechanism-elucidation approaches. With the former interpreting interaction to be a strictly statistical phenomenon, as suggested by Fisher, and the latter emphasizing the need to probe and understand the causal mechanisms behind interaction responsible for the developmental process, as suggested by Hogben.[11] For complex diseases like hypertension, failure to pay proper attention to the methodological checks that must be undertaken for cG×E has resulted in some studies making dubious claims.[29,30]

Logistic regression requiring main effects and linear conditions is usually restricted in low-order interactions, and exploration relying on cross-product terms to measure the cause of the variation responsible for the interaction can be particularly problematic. Besides, explanatory and confounding variables are entered equally into logistic regression model, which may have the effect of dispersing or concealing the actual main and interactive effects.[31] For example, the main protective effects from the ACE genotype and the risk interaction effects of ACE genotype and salt intake for women in current study are not able to be readily explained, since the ACE genotype has no reported independent effect on

hypertension at most, but it barely shows the opposite effect after interacting with other factors.[8] Therefore, new models suit for complicated high-order interaction and appropriate confounder adjustment, designed to be distinguished from generalized linear regression models (GLMs), are warranted to reveal gene-environment interaction for complex diseases.[32]

The generalized partially linear tree-based regression model is semi-parametric and integrates the advantages of the tree structure and GLMs. Its purpose is to decompose a data space recursively into smaller areas defined by a set of explanatory variables with adjustment for confounding variables,[15,33,34] where explanatory and confounding variables are not treated in the same way. The modified procedure [15] can explicitly identify sub-groups of individuals at high risk for hypertension with less computational demands, where highlighted features of subgroup provide clues for the mechanisms of disease development. Taking the application of the GPLTR model in the current study as an example, results indicated that the D allele might be the trigger for hypertension among men with a high salt intake, and carriers of the D allele combined with a lower salt intake could help protect women from the development of hypertension.

Strengths and limitations

The high-quality study design, a response rate of 92.4% and high concentrations of Chinese Kazakhs in the current study provide good internal validity of the results. Sex-stratified analyses were performed to reveal potential differences in mechanisms behind the interaction of ACE genotype and salt intake on hypertension. The high prevalence of hypertension (49.4% for men and 35.9% for women) allowed for robust comparisons between people with

hypertension and non-hypertensive control participants. Application comparison between the logistic model and the GPLTR model was based on the same confounding variables by gender, which were cautiously chosen according to baseline results.

Study also need to be considered in light of limitations. The relatively small sample size may allow statistical instability. However, all enrolled participants did not have consanguinity, which is optimal for gene-environment interaction research. Salt intake was assessed by urinary sodium excretion and might be underestimated because salt can be removed from the body through other means, for example, sweat. Regardless, this sodium excretion estimation method has been proved to be closely correlated with the value determined from 24-h pooled urine.[18] Additionally, because of field conditions, this method was determined to be superior to an estimation from a dietary questionnaire.

CONCLUSION

The interaction of ACE genotype and salt intake on hypertension was observed for both men and women among Chinese Kazakhs, but in different ways. Application comparison of the two models should be of interest to epidemiologists and geneticists for cG×E exploration, and the GPLTR model might be more suitable for complex diseases to provide valuable clues as to the etiology of the disease. Further research is needed to determine the causes for the observed differences behind the interaction mechanisms for men and women.

Acknowledgements We thank the government of Altay City and Hong Dun Town and the Health Bureau of Altay City who supported this project, the village head and staff of each sample village, clinical doctors and nurses of town hospital and people hospital in Altay

region, Kazakh college students from Xinjiang Medical University who participated in the surveys, and all participants of this study.

Funding This work was supported by National Natural Science Foundation of China (81273181).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Committee of Institute of Basic Medical Sciences Chinese Academy of Medical Sciences.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Contributorship statement YW and BZ abstracted data, performed data analysis and interpreted results with substantial contributions from the other co-authors, and drafted the manuscript, which was critically revised by all co-authors. JJ and LH conceived the study and assisted in data analysis, design of figures and tables. WH, FX and YW assisted in data analysis and interpretation, added important background knowledge and improved the manuscript by repeated readings and rephrasing as well as critical discussions of the intellectual content. MZ, YT, SL, WW, KA and CQ contributed to the data collection, quality control, general editing purposes and comments regarding medical terminology and critical review of the manuscript. ZW, YH and LW mainly performed statistical analyses, designed tables and figures and reviewed the manuscript. All authors have read and approved the final

version of this manuscript for submission.

REFERENCES

1. Padmanabhan S, Caulfield M, Dominiczak AF. Genetic and Molecular Aspects of Hypertension. *Circ Res* 2015;116(6):937-59.
2. Te RL, van Esch JH, Roks AJ, et al. Hypertension: renin-angiotensin-aldosterone system alterations. *Circ Res* 2015;116(6):960-75.
3. Adrogué HJ, Madias NE. Sodium and potassium in the pathogenesis of hypertension. *N Engl J Med* 2007;356(19):1966-78.
4. Sayed-Tabatabaei FA, Oostra BA, Isaacs A, et al. ACE polymorphisms. *Circ Res* 2006;98(9):1123-33.
5. Ned RM, Yesupriya A, Imperatore G, et al. The ACE I/D polymorphism in US adults: limited evidence of association with hypertension-related traits and sex-specific effects by race/ethnicity. *Am J Hypertens* 2012;25(2):209-15.
6. O'Donnell CJ, Lindpaintner K, Larson MG, et al. Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. *Circulation* 1998;97(18):1766-72.
7. Ji LD, Zhang LN, Shen P, et al. Association of angiotensinogen gene M235T and angiotensin-converting enzyme gene I/D polymorphisms with essential hypertension in Han Chinese population: a meta-analysis. *J Hypertens* 2010;28(3):419-28.
8. Calculator B. Hypertension and the ACE gene. http://www.gbhealthwatch.com/GND-Hypertension-ACE.php (accessed 5 Jul 2016).
9. Fg VDK, de Jong PE, Henning RH, et al. Enhanced responses of blood pressure, renal function, and aldosterone to angiotensin I in the DD genotype are blunted by low sodium intake. *J Am Soc Nephrol* 2002;13(4):1025-33.
10. Zhang L, Miyaki K, Araki J, et al. Interaction of angiotensin I-converting enzyme insertion-deletion polymorphism and daily salt intake influences hypertension in Japanese men. *Hypertens Res* 2006;29(10):751-8.
11. Tabery J. Debating interaction: the history, and an explanation. *Int J Epidemiol* 2015;44(4):1117-23.
12. Dick DM, Agrawal A, Keller MC, et al. Candidate gene-environment interaction research: reflections and recommendations: Brookings Institution 37-59 p.
13. Aliev F, Latendresse SJ, Bacanu SA, et al. Testing for Measured Gene-Environment Interaction:

Problems with the use of Cross-Product Terms and a Regression Model Reparameterization Solution. *Behav Genet* 2014;44(2):165-81.

14. Ganzach Y. Misleading interaction and curvilinear terms. *Psychol Methods* 1997;2(3):235-47.
15. Mbogning C, Perdry H, Toussile W, et al. A novel tree-based procedure for deciphering the genomic spectrum of clinical disease entities. *J Clin Bioinforma* 2014;4(1):1-11.
16. Jiang J, Zhang B, Zhang M, et al. Prevalence of conventional cardiovascular disease risk factors among Chinese Kazakh individuals of diverse occupational backgrounds in Xinjiang China. *Int J Cardiol* 2015;179:558-60.
17. Hou L, Zhang M, Han W, et al. Influence of Salt Intake on Association of Blood Uric Acid with Hypertension and Related Cardiovascular Risk. *PLoS One* 2016;11(4).
18. Kawano Y, Tsuchihashi T, Matsuura H, et al. Report of the Working Group for Dietary Salt Reduction of the Japanese Society of Hypertension: (2) Assessment of salt intake in the management of hypertension. *Hypertens Res* 2007;30(10):887-93.
19. Wang X, Wang S, Lin R, et al. Relationship of I/D polymorphism of angiotensin converting enzyme gene with hypertension in Xinjiang Kazakh isolated group. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2003;20(3):253-5.
20. Zhang B, Simayi A, Sun Y, et al. Interaction Between Genetic and Environmental Risk Factors on Essential Hypertension in Kazakh Population. *J Med Res* 2009.
21. Stankovic A, Zivkovic MD. Angiotensin I-converting enzyme gene polymorphism in a Serbian population: a gender-specific association with hypertension. *Scandinavian Journal of Clinical & Laboratory Investigation* 2009;62(6):469-75.
22. Chiang FT, Chern TH, Lai ZP, et al. Age- and gender-dependent association of the angiotensin-converting enzyme gene with essential hypertension in a Chinese population. *J Hum Hypertens* 1996;10(12):823-6.
23. Busjahn A, Knoblauch H, Knoblauch M, et al. Angiotensin-Converting Enzyme and Angiotensinogen Gene Polymorphisms, Plasma Levels, Cardiac Dimensions A Twin Study A Twin Study. *Hypertension* 1997;29(2):165-70.
24. Krege JH, John SWM, Langenbach LL, et al. Male–female differences in fertility and blood pressure in ACE-deficient mice. *Nature* 1995;375(6527):146-8.
25. Fornage M, Amos CI, Kardia S, et al. Variation in the region of the angiotensin-converting enzyme

gene influences interindividual differences in blood pressure levels in young white males. *Circulation* 1998;97(97):1773-9.

26. O'Donnell M, Mente A, Yusuf S. Sodium intake and cardiovascular health. *Circ Res* 2015;116(6):1046-57.

27. Giner V, Poch E, Bragulat E, et al. Renin-Angiotensin System Genetic Polymorphisms and Salt Sensitivity in Essential Hypertension. *Hypertension* 2000;35(1 Pt 2):512-7.

28. Caspi A, Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2010;301(5631):386-9.

29. Rutter M. Some of the complexities involved in gene-environment interplay. *Int J Epidemiol* 2015;44(4):1128-9.

30. Chair MR. Chapter 12. Gene–Environment Interaction in Complex Diseases: Asthma as an Illustrative Case: John Wiley & Sons, Ltd; 2008. 184-97.

31. Keller MC. Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biol Psychiatry* 2014;75(1):18-24.

32. Bastone L, Reilly M, Rader DJ, et al. MDR and PRP: a comparison of methods for high-order genotype-phenotype associations. *Hum Hered* 2004;58(2):82-92.

33. Chen J, Yu K, Hsing A, et al. A partially linear tree ‐ based regression model for assessing complex joint gene–gene and gene– environment effects. *Genet Epidemiol* 2007;31(3):238–51.

34. Mbogning C, Perdry H, Broët P. A Bagged, Partially Linear, Tree-Based Regression Procedure for Prediction and Variable Selection. *Hum Hered* 2015;79(3-4):182-93.

Figure legends

Figure 1 Flow chart for enrolled 916 independent subjects.

Figure 2 Optimal trees obtained from 342 independent men and 574 independent women using AIC and BIC selection criterion. (Leaves are denoted by rectangles and number in each node represents the number of subjects falling in this leave)

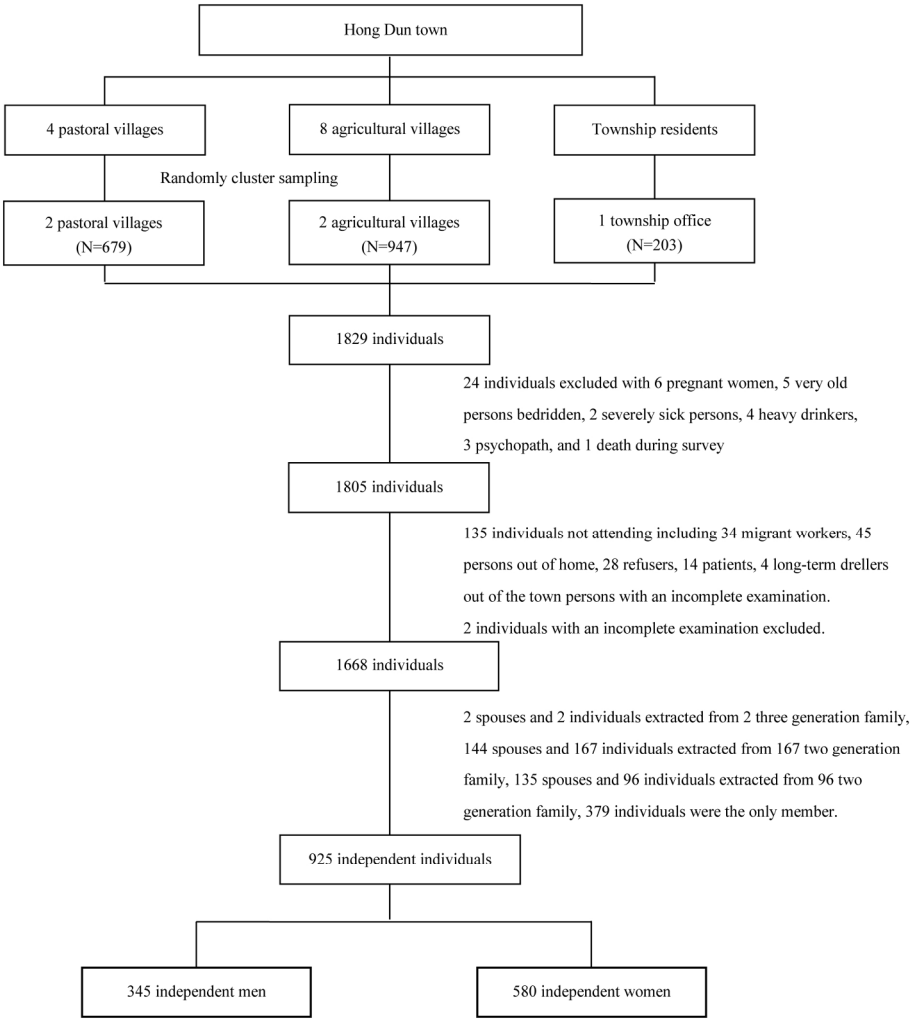

Fig 1. Flow chart for recruitment of participants and enrolled 925 subjects.

Figure 1 Flow chart for enrolled 916 independent subjects.

170x209mm (300 x 300 DPI)

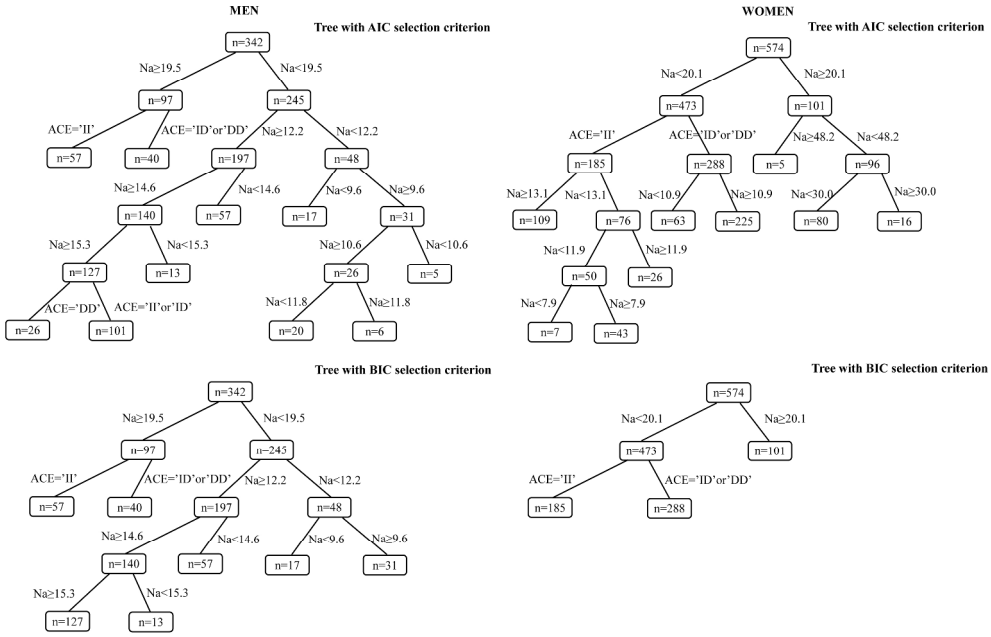

Figure 2 Optimal trees obtained from 342 men and 574 women using AIC and BIC selection criteria. Leaves are denoted by rectangles and the number in each node represents the number of participants falling in this leaf.

292x188mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-6
Bias	9	Describe any efforts to address potential sources of bias	5-6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-6
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7
		(*b*) Describe any methods used to examine subgroups and interactions	7
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Results			Page No
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7
		(b) Give reasons for non-participation at each stage	7
		(c) Consider use of a flow diagram	7
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8
		(b) Indicate number of participants with missing data for each variable of interest	8
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	7-8
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-12
		(*b*) Report category boundaries when continuous variables were categorized	8-12
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8-12
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	16-17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	13-16
Generalisability	21	Discuss the generalisability (external validity) of the study results	16-17
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: results from a population-based cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014246.R1
Article Type:	Research
Date Submitted by the Author:	03-Feb-2017
Complete List of Authors:	Wang, Yuyan; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Zhang, Biao; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hou, Lei; Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, , Department of Epidemiology and Biostatistics Han, Wei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Xue, Fang; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Yanhong; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Tang, Yong; The People's Hospital in Altay Region Liang, Shaohua; The People's Hospital in Altay Region Wang, Weizhi; The People's Hospital in Altay Region Asaiti, Kuliqian; Hong Dun Town hospital, Altay Region Wang, Zixing; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hu, Yaoda; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Lei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Qiu, Changchun; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences Zhang, Mingtao; The People's Hospital in Altay Region Jiang, Jingmei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods, Genetics and genomics
Keywords:	Hypertension < CARDIOLOGY, Interaction, ACE genotype, Salt intake, Logistic model, GPLTR model

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: results from a population-based cross-sectional study

Yuyan Wang[1†], Biao Zhang[1†], Lei Hou[1], Wei Han[1], Fang Xue[1], Yanhong Wang[1], Yong Tang[2], Shaohua Liang[2], Weizhi Wang[2], Kuliqian.Asaiti[3], Zixing Wang[1], Yaoda Hu[1], Lei Wang[1], Changchun Qiu[1], Mingtao Zhang[2*], Jingmei Jiang[1*]

[1] *Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, China*

[2] *The People's Hospital in Altay Region, Xinjiang, China*

[3] *Hong Dun Town hospital, Altay Region, Xinjiang, China*

[†] These authors contributed equally to this article.

[*] Corresponding Authors:

Prof. Jingmei Jiang, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, 5 Dong Dan San Tiao, Beijing, 100005, China. E-mail: jingmeijiang@ibms.pumc.edu.cn. Tel: +86 10 69155935, Fax: +86 10 65123696.

Mr. Mingtao Zhang, The People's Hospital in Altay Region, Xinjiang, China. E-mail: altrmyy@163.com. Tel:+86018896519163.

ABSTRACT

Objectives: To explore the effect of interaction between angiotensin I-converting enzyme (ACE) genotype and salt intake on hypertension among Chinese Kazakhs, and to compare applications of interactions between logistic model and generalized partially linear tree-based regression (GPLTR) model.

Design: Population-based cross-sectional study.

Setting: Hong Dun, North Xinjiang, China.

Participants: Non-consanguineous Chinese Kazakh participants (n=916, 342 men and 574 women) aged 30 years or over.

Main outcome measures: Association between ACE genotype and hypertension, association between salt intake and hypertension, and interaction of ACE genotype and salt intake on hypertension in two models.

Results: Associations between salt intake and hypertension were different in ACE genotype of II and ID+DD. Under the logistic models, main and interaction effects were not observed for men, but effects were present in opposite directions for women (main effect of ACE: OR=0.20, p=0.003; interaction effect: OR=1.07, p=0.027). Under the GPLTR model, Bayesian information criterion (BIC) trees included both salt intake and ACE genotype as split variables. Individuals with a salt intake≥19.5g/day and ID+DD genotypes had a 3.99-fold (p=0.004) higher risk for hypertension compared with the II genotype for men, while salt intake<20.1g/day and ID+DD genotypes had an OR=0.55 (p=0.014) compared with the II genotype for women.

Conclusions: An interaction of ACE genotype and salt intake on hypertension was

observed among Chinese Kazakhs but in different ways according to sex. The GPLTR model appears to be more suitable for an exploration of interactions in complex diseases.

Strengths and limitations of this study

1. This is one of the few studies to explore the interaction of ACE genotype and salt intake on hypertension, and is also the first such study in Chinese Kazakhs.

2. The design strengths of the study include: the extraction of participants without consanguinity, the use of stratified analyses to test the existence of an interaction effect, and a comparison of the interaction exploration results from two models.

3. The cross-sectional design cannot realize an investigation of the causal pathway of the association between the ACE genotype and salt intake interaction with hypertension, although possible mechanisms were suggested for the interaction effect on hypertension.

4. Salt intake was assessed by urinary sodium excretion and might be underestimated because salt can be removed from the body through other means.

INTRODUCTION

Hypertension is a complex disorder resulting from an interaction of genetic and environmental factors.[1] Blood pressure (BP) is regulated by the renin–angiotensin–aldosterone system (RAAS),[2] which is heavily influenced by dietary salt intake. Indeed, dietary salt is recognized as the most important environmental contributor to hypertension.[3] The most common genetic variation in RAAS is the insertion/deletion polymorphism of the angiotensin I-converting enzyme gene (ACE),[4] and its association with BP has been reported to be dependent on sex and race/ethnicity.[4-7] Although the presence of an interaction between ACE genotype and dietary salt intake has been supported by research findinges,[8 9] few reports have investigated the effect of this interaction on hypertension in different populations.[10]

Measurements and explanations of candidate gene–environment interaction (cG×E) in populations have long been explored,[11] resulting in a general consensus that statistical concerns about their modeling are often overlooked.[12] Popular logistic regression models including interaction terms can be problematic,[13 14] but the newly introduced generalized partially linear tree-based regression (GPLTR) model, which combines tree structure and linear adjustment, may offer more flexibility and power to the exploration of interactions in complex diseases.[15]

The aims of this study were to explore the interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs, who have a high prevalence of hypertension and a high dietary salt intake, and to compare the logistic model and the GPLTR model in interaction applications.

METHODS

Study participants

The design scheme, sampling process and data collection of the Chinese Xinjiang Altay Kazakh Heart Study have been presented elsewhere.[16 17] Hong Dun town is located in the urban rural fringe of Altay prefecture in North Xinjiang, China. From October 2012 to February 2013, a total of 1805 individuals aged 30 years or over were recruited, and 1668 participants from 601 families completed the baseline survey and examination under standard procedures. Participants were required to have at least three paternal/maternal biological generations living in the same region without a history of intermarriage. Non-consanguineous spouses without consanguinity and only one member of biological relatives from each family were extracted to ensure independence (figure 1). Approval for the study was given by the Ethics Committee of the Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, and written informed consent was obtained from all participants prior to enrollment.

Study variables

We complied with current Chinese Guidelines for Blood Pressure Measurement.[18] A quiet room in a village clinic was selected for BP measurements and all people in the room were required to keep quiet. Participant were comfortably seated, with legs uncrossed and the back and arm supported, such that the middle of the cuff on the upper arm was at the level of the right atrium (the mid-point of the sternum). BP was measured with an appropriate arm cuff and a mercury column sphygmomanometer on the left arm after a resting period of at least 10 min in the supine position. All measurements used the same standard procedure and calibrated mercury sphygmomanometers, and were conducted by trained and qualified staff. Currently, a large-size arm cuff for a mercury column sphygmomanometer is not available in the market of

China, so radial artery auscultation on the upper forearm was used for obese individuals with large upper arms, as recommended by BP measurement guidelines.[18] BP was defined as the mean of two readings, or three readings if there was a difference of more than 5 mmHg between the initial readings. Participants were classified as hypertensive if they met at least one of the following three criteria: a) a systolic BP ≥ 140 mmHg; b) a diastolic BP ≥ 90 mmHg; c) the use of antihypertensive medications within the last 2 weeks.

Venous blood samples were drawn after an overnight fast of at least 10 hours and centrifuged within 20 minutes, then divided into two. One sample was used for biochemistry assays performed within 4 hours, and the other was stored in a portable refrigerator at -20°C then transferred as soon as possible to Beijing, where it was kept at -80°C for genetic analysis. D and I ACE alleles were identified by polymerase-chain-reaction (PCR) amplification of the respective fragments from ACE intron 16 and size fractionation and visualization by electrophoresis. The ACE genotype was classified as II, ID or DD depending on the presence of each allele, with a 99.76% genotyping success rate.

The second urine sample after waking was collected, and urinary sodium was measured with a Caretium XI-921CT automatic electrolyte analyzer (Shenzhen, China). Salt intake was assessed by urinary sodium excretion per day from the amounts of urinary sodium, urine creatinine (Cr), and the 24-h urinary Cr excreted as estimated from height, body weight and age,[19] for which each 43 mmol of sodium is approximately equivalent to 1 g of sodium or 2.5 g of salt (sodium chloride).

Other covariates were also collected using standard procedures. Demographic information (including age, sex, occupation and education), and information about diet, alcohol

consumption, cigarette smoking and medical history were obtained from a face-to-face questionnaire survey. Body weight and height were measured in light indoor clothes and without shoes. Body mass index (BMI) was calculated as body weight in kilograms divided by height in meters squared (kg/m^2). Fasting blood glucose (FBG) was collected according to standardized protocols by a BECKMAN COULTER AU2700 clinical chemistry analyzer (CA, USA).

The GPLTR model and selection criteria

Y was denoted as the outcome of interest, X were the confounding variables to be modeled linearly, and G were the group of risk factors whose interaction effect was to be approximated by the tree-based structure. The main purpose of this GPLTR model was to test the combined effect of G on Y while adjusting for the confounders X; the model was expressed as

$$g(E(Y|X,Z)) = X'\theta + \beta_T F(T(Z))$$

where $g(.)$ is a specific link function (generalized linear model), such as the logit link function $g(\mu) = \log(\frac{\mu}{1-\mu})$ for a binary outcome, and $F(T(Z))$ is the vector of indicator variables signifying the leaves of the tree $T(Z)$.

In this study, we used Y to represent the presence of hypertension, G to be the ACE genotype and the estimated salt intake, and X to be the environmental risk factors having an apparent impact on the hypertension rate. First, we constructed a maximal tree with successive steps of initialization, iteration and ending conditions. The goodness of a candidate split was assessed for each node by the deviance of a generalized linear model, which equals the test deviance of the parent node minus the sum of the deviance of the two child nodes.

Second, the sequence of nested candidate subtrees was created, and Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to determine the best tree in the sequence.

Statistical analysis

All analyses were conducted separately by sex. Basic characteristics were compared with regard to the status of hypertension. Variables significantly different between hypertension and normotension were chosen as confounders. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the associations of the ACE polymorphism with hypertension. To examine the presence of an interaction, associations of salt intake with hypertension were illustrated as stratified by ACE genotype as II and ID+DD. To compare the effect of the interaction, two models were applied, a logistic regression model, and a GPLTR model adjusting for the same covariates to extract the high-risk group of hypertension using AIC and BIC. Description and preliminary analyses were carried out by SAS Version 9.3 software (SAS Institute, Cary, NC, USA), and GPLTR models were conducted by R Version 3.1.3 with the 'GPLTR' package. One-sided hypotheses were established for trend tests based on the logistic model, and all p-values were two-tailed and considered significant at <0.05.

RESULTS

This study enrolled 916 independent participants (342 men and 574 women) who had complete data for all study variables. The prevalence of hypertension was 49.4% for men and 35.9% for women. Participant characteristics are shown in table 1. Briefly, participants with hypertension had increased age, BMI, and FBG compared with normotensive participants for

both sexes.

Table 1 Basic characteristics of enrolled 916 participants by sex and hypertension status

	Men (N=342)			Women (N=574)		
	Hypertension (n=169)	Normotension (n=173)	p	Hypertension (n=206)	Normotension (n=368)	p
Age, years	50.1 ± 12.1	42.8 ± 10.5	<0.001	51.0 ± 11.2	39.6 ± 7.8	<0.001
Occupation, n (%)						
Nomad	68 (40.2)	67 (38.7)	0.900	87 (42.2)	114 (31.0)	0.001
Farmer	75 (44.4)	81 (46.8)		94 (45.6)	169 (45.9)	
City worker	26 (15.4)	25 (14.5)		25 (12.2)	85 (23.1)	
Education, n (%)			0.028			<0.001
Primary school and below	61 (36.1)	40 (23.1)		75 (36.4)	63 (17.1)	
Junior middle school	69 (40.8)	89 (51.5)		86 (41.8)	161 (43.8)	
High school and above	39 (23.1)	44 (25.4)		45 (21.8)	144 (39.1)	
BMI, kg/m^2	26.9 ± 4.8	24.6 ± 3.4	<0.001	28.7 ± 5.4	25.8 ± 4.5	<0.001
Smoking, n (%)	116 (69.8)	133 (76.9)	0.140	12 (5.8)	11 (3.0)	0.097
Drinking, n (%)	53 (31.4)	38 (22.0)	0.049	0 (0.0)	0 (0.0)	—
SBP, mm Hg	152.3 ± 19.2	122.2 ± 9.5	<0.001	154.1 ± 21.7	117.8 ± 9.8	<0.001
DBP, mm Hg	95.7 ± 11.2	77.9 ± 7.2	<0.001	94.1 ± 11.4	75.5 ± 7.7	<0.001
Family history of hypertension, n (%)	92 (57.1)	94 (56.3)	0.876	116 (60.1)	213 (58.2)	0.663
FBG, mmol/L	5.4 ± 0.8	5.2 ± 0.4	<0.001	5.3 ± 0.8	5.0 ± 0.5	<0.001
Salt intake, g/day	19.5 ± 26.7	18.3 ± 15.9	0.599	17.1 ± 7.4	15.9 ± 9.7	0.129

BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; FBG: fasting blood glucose.
Values are means ± SD and frequency (percent).

Independent associations

The independent associations of the ACE genotype and salt intake with hypertension are shown in table 2. The D allele had a protective effect on hypertension in women, with the DD genotype having an OR=0.47 (95% CI: 0.25-0.88) compared with the II genotype. This protective effect was not seen in men, with the DD genotype having an OR=1.27 (95% CI: 0.66-2.46) compared with the II genotype. Salt intake by quartiles did not correlate with hypertension for either sex (p for trend=0.095 for men and 0.152 for women).

Table 2 Association of hypertension with ACE Genotype and salt intake

	ACE Genotype				Quartiles of salt intake with median (interquartile range), g/day				
	II (n=356)	ID (n=399)	DD (n=161)	p	Q1	Q2	Q3	Q4	p for trend
Men	128	148	66		11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
HTN, %	46.9	49.3	54.6		48.8	47.1	44.2	57.7	
Crude OR (95% CI)	1.00	1.10 (0.69, 1.77)	1.36 (0.75, 2.47)	0.327	1.00	0.93 (0.51, 1.70)	0.83 (0.46, 1.51)	1.43 (0.78, 2.61)	0.334
Age-adjusted OR (95% CI)	1.00	1.08 (0.66, 1.78)	1.31 (0.70, 2.46)	0.416	1.00	0.95 (0.50, 1.79)	0.88 (0.47, 1.66)	1.79 (0.95, 3.39)	0.105
Multi-adjusted OR*(95% CI)	1.00	1.01 (0.60, 1.72)	1.24 (0.64, 2.40)	0.567	1.00	0.98 (0.50, 1.95)	0.90 (0.46, 1.76)	1.94 (0.97, 3.90)	0.099
Women	228	251	95		10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
HTN, %	39.9	35.1	28.4		28.5	36.4	35.0	43.8	
Crude OR (95% CI)	1.00	0.81 (0.56, 1.18)	0.60 (0.36, 1.00)	0.048	1.00	1.44 (0.87, 2.36)	1.35 (0.82, 2.23)	1.95 (1.20, 3.19)	0.013
Age-adjusted OR (95% CI)	1.00	0.71 (0.46, 1.09)	0.50 (0.27, 0.92)	0.017	1.00	1.59 (0.89, 2.85)	1.25 (0.69, 2.26)	1.87 (1.06, 3.30)	0.073
Multi-adjusted OR* (95% CI)	1.00	0.71 (0.45, 1.13)	0.47 (0.25, 0.87)	0.014	1.00	1.36 (0.74, 2.51)	1.11 (0.59, 2.09)	1.64 (0.89, 2.99)	0.178

HTN: hypertension prevalence;

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose; Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

Interaction effect

Stratified analysis

Associations between salt intake and hypertension stratified by ACE II genotype and ID+DD genotypes were evaluated (table 3). No correlation was observed in the II genotype group for either men (p for trend=0.967) or women (p for trend=0.904). An increase in hypertension with a higher salt intake was observed in the ID+DD genotypes group for both men (p for trend=0.028) and women (p for trend=0.045).

Logistic model

Results from the logistic regression model adjusted for multiple confounders are shown in table 4. For men, no main effects of the ACE genotype (OR=1.33, p=0.426) or salt intake (OR=1.00, p=0.415) were observed, and there was no interaction effect (OR=0.99, p=0.464). For women, no main effect of salt intake (OR=0.99, p=0.523) was observed, but a main effect of the ACE genotype (OR=0.20, p=0.003) and an interaction effect (OR=1.07, p=0.027) were observed to be present in opposite directions. Recessive model was also built as sensitivity analysis, and ORs were turned as reciprocals as expected, which confirms the robustness of the findings.

GPLTR model

All iterative procedures converged after 15 iterations, and trees selected by AIC and BIC with different degrees of pruning resulted in 10 and seven leaves for men, and nine and three leaves for women, respectively (figure 2). BIC trees were chosen for the final interpretation based on their stability. Risk assessments of trees are shown in table 4. For men, leaves are identified from the left as:

(1) Node 1: 57 participants characterized by salt intake ≥ 19.5 g/day and ACE II genotype;

(2) Node 2: 40 participants characterized by salt intake ≥ 19.5 g/day and ACE ID+DD genotype;

(3) Node 3: 127 participants characterized by 15.3 g/day ≤ salt intake < 19.5 g/day;

(4) Node 4: 13 participants characterized by 14.6 g/day ≤ salt intake < 15.3 g/day;

(5) Node 5: 57 participants characterized by 12.2 g/day ≤ salt intake < 14.6 g/day;

(6) Node 6: 17 participants characterized by salt intake < 9.6 g/day;

(7) Node 7: 31 participants characterized by 9.6 g/day ≤ salt intake < 12.2 g/day.

Split variables for the selected tree included both salt intake and ACE genotype as II and ID+DD. After adjusting for the same confounders in the multivariable logistic model, the final tree showed that Node 2 had 3.99-fold (p=0.004) higher risk for hypertension compared with Node 1, and the Node 3 had an OR=0.93 (p=0.860). For women, leaves were identified from the left as:

(1) Node 1: 185 participants characterized by salt intake < 20.1 g/day and ACE II genotype;

(2) Node 2: 288 participants characterized by salt intake < 20.1 g/day and ACE ID+DD genotype;

(3) Node 3: 101 participants characterized by salt intake ≥ 20.1 g/day.

Split variables also included both salt intake and ACE genotype as II and ID+DD. Node 2 had an OR=0.55 (*p*=0.014) and Node 3 had an OR=1.26 (*p*=0.444) compared with Node 1.

Table 3 Multivariable-adjusted odds ratios (95% confidence interval) in the ACE genotypes of II and ID+DD by different salt intake level in men and women

Multi-adjusted OR* (95% CI)	Quartiles of salt intake with median (interquartile range), g/day				
	Q1	Q2	Q3	Q4	p for trend
Men	11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
II (n=128)	1.00	2.04 (0.66, 6.30)	1.59 (0.51, 5.01)	1.09 (0.37, 3.19)	0.967
ID+DD (n=214)	1.00	0.59 (0.25, 1.41)	0.68 (0.30, 1.58)	3.15 (1.25, 7.93)	0.028
Women	10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
II (n=228)	1.00	1.19 (0.47, 2.99)	0.75 (0.28, 2.04)	1.15 (0.44, 3.00)	0.904
ID+DD (n=346)	1.00	1.55 (0.67, 3.59)	1.53 (0.66, 3.52)	2.35 (1.06, 5.24)	0.045

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

Table 4 ACE genotype and salt intake interaction explorations by logistic model and generalized partially linear tree-based regression model

Logistic model			GPLTR model		
Terms in the model	Multi-adjusted OR* (95% CI)	p	Nodes in the model	Multi-adjusted OR* (95% CI)	p
Men (ROC=0.747)			Men (ROC=0.809)		
Salt intake, g/day	1.00 (0.99, 1.03)	0.415	Node 1: Salt intake ≥ 19.5 and ACE(II)	1.00	
ACE(ID+DD vs. II)	1.33 (0.66, 2.65)	0.426	Node 2: Salt intake ≥ 19.5 and ACE(ID+DD)	3.99 (1.55, 10.26)	0.004
ACE(ID+DD vs. II)×Salt intake	0.99 (0.96, 1.02)	0.464	Node 3: 15.3 ≤ Salt intake < 19.5	0.93 (0.41, 2.09)	0.860
Women (ROC=0.839)			Women (ROC=0.839)		
Salt intake, g/day	0.99 (0.94, 1.03)	0.523	Node 1: Salt intake < 20.1 and ACE(II)	1.00	
ACE(ID+DD vs. II)	0.20 (0.07, 0.59)	0.003	Node 2: Salt intake < 20.1 and ACE(ID+DD)	0.55 (0.34, 0.89)	0.014
ACE(ID+DD vs. II)×Salt intake	1.07 (1.01, 1.13)	0.027	Node 3: 20.1 ≤ Salt intake	1.26 (0.70, 2.25)	0.444

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

DISCUSSION

There are three main findings from the present study. First, we observed that the association between ACE gene polymorphism and hypertension was sex-specific. Second, stratified analysis indicated that an interaction between ACE genotype and salt intake existed for both sexes. Third, an application comparison of the two models suggested that the traditional logistic model was not appropriate in this situation, and that the GPLTR model could effectively determine the high-risk characteristics of hypertension.

The distribution of the ACE genotype observed in this study is similar to that reported for an isolated group of Kazakhs in 2003,[20] with a DD genotype frequency of less than 20%, which differs from that observed in other studies on Kazakhs[21] and in studies of other ethnicities.[4-7] This difference may arise from the diverse structure of the population and racial heterogeneity. Our current results are consistent with previous studies that reported ACE genotype to be a sex-specific candidate gene for hypertension.[5 6 22 23] No positive association was detected between the D allele and hypertension in men, but a protective effect of the D allele was observed for women, presumably through sexes, ethnicities and other factors including gene-gene or gene-environment interactions.[6] Nonlinear association between salt intake and risk of hypertension was confirmed by restricted cubic spline with automatically computed 'knots' (data not shown). With regard to the interaction of ACE genotype and salt intake on hypertension, a previous study that examined 284 Japanese men using logistic regression analysis[10] reported a negative association for ACE (II+ID vs. DD: OR=0.1, p=0.024) and a positive association for interaction (ACE(II+ID vs. DD)× salt intake level: OR=3.6, p=0.047) in opposite directions, which was similar to our current results when the II

genotype was set as the reference.

Strong evidence from association and linkage studies suggests that ACE has a profound effect on the variance of plasma ACE levels,[24] which maintain BP homeostasis. Among the major factors that activate RAAS, dietary sodium/potassium balance is the most important, and is mainly influenced by dietary salt intake.[3] The sodium status has been reported to affect the ACE I/D polymorphism phenotype, and to enhance BP and renal function responses, while the conversion of aldosterone to angiotensin I with the DD genotype is blunted by low a sodium intake.[8] These reports suggest that the ACE polymorphism interacts with dietary salt intake, with subsequent effects on the regulation of BP. Gene-targeting experiments that inactivate ACE in mice have provided evidence for the sex-specific effect of ACE gene on blood pressure.[25] Population-based research also supports the sex-specific influence of ACE on BP variability.[4 26] Assessment of the dietary habits of individuals in the current study (data not shown) revealed average salt intake of 17.3 g/day, which is far greater than that of other populations and World Health Organization recommendations.[27] Given the high daily salt intake for this population, we suggest that the potential harmful effects of high salt intake on the development of hypertension should be emphasized to Chinese Kazakhs. Previous studies into the relationship between ACE and salt sensitivity indicated possession of the ACE D allele by men may lead to an increased response to a high salt intake;[4 28] thus carriers of the D allele may trigger hypertension if they consume a high salt diet. For women, the mechanism behind the interaction is uncertain; however, the combination of the ACE D allele and a lower salt intake may have a synergistic effect on hypertension.[28]

Previous discussions into gene–environment interactions have historically progressed

through three main stages, including eugenics, IQ, and the serotonin transporter gene controversy.[29] This has resulted in an apparent divide between variation-partitioning and mechanism-elucidation approaches. The former approach interpreted interactions as strictly statistical phenomena, as suggested by Fisher,[11] while the latter emphasized the need to probe and understand the causal mechanisms behind interaction responsible for the developmental process, as suggested by Hogben.[11] For complex diseases like hypertension, failure to pay proper attention to the methodological checks that must be undertaken for cG×E has resulted in some studies making dubious claims.[30 31] Logistic regression requiring main effects and linear conditions is usually restricted to low-order interactions, and exploration relying on cross-product terms to measure the cause of the variation responsible for the interaction can be particularly problematic. We have built a full model using three main factors (i.e., gender, ACE genotype, and salt intake), in which the three main effects and the second-order interaction term are positive, but the three first-order interaction terms are all negative. Moreover, explanatory and confounding variables are entered equally into logistic regression models, which may disperse or conceal the actual main and interactive effects.[32] For example, the main protective effects from the ACE genotype and the risk interaction effects of the ACE genotype and salt intake in women of the current study cannot be readily explained because the ACE genotype has no reported independent effect on hypertension, and rarely shows the opposite effect after interacting with other factors.[8] Therefore, new models suitable for complicated high-order interactions and appropriate confounder adjustments, designed to be distinguished from generalized linear regression models (GLMs), are warranted to reveal gene-environment interactions for complex diseases.[33] The GPLTR model is semi-parametric

and integrates the advantages of the tree structure and GLMs. Its purpose is to decompose a data space recursively into smaller areas defined by a set of explanatory variables with adjustment for confounding variables,[15 34 35] where explanatory and confounding variables are not treated in the same way. The modified procedure [15] can explicitly identify sub-groups of individuals at high risk for hypertension with fewer computational demands, where highlighted features of subgroup provide clues for the mechanisms of disease development. Taking the results of the GPLTR model in the current study as an example, the ACE D allele was suggested to be the trigger for hypertension among men with a high salt intake, and female carriers of the D allele combined with a lower salt intake may offer help protection from the development of hypertension. Our calculation of the 10-fold cross-validation prediction error (data not shown) of the logistic regression model and GPLTR model for both men and women revealed the overall improved accuracy of GPLTR model.

The high-quality study design, a response rate of 92.4% and high aggregation of Chinese Kazakhs in the current study provide good internal validity of the results. Sex-stratified analyses were performed to reveal potential differences in mechanisms behind the effect of the ACE genotype and salt intake interaction on hypertension. The high prevalence of hypertension (49.4% for men and 35.9% for women) allowed for robust comparisons between individuals with hypertension and non-hypertensive controls. Application comparisons between the logistic model and the GPLTR model was based on the same confounding variables by sex, which were cautiously chosen according to baseline results. The study should also be considered in light of its limitations. The cross-sectional design cannot realize investigation of the causal pathway of the association between the ACE genotype and salt

intake interaction with hypertension, so confirmatory studies in other large population-based samples are warranted. The relatively small sample size may cause statistical instability. However, all enrolled participants did not have consanguinity, which is necessary for gene-environment interaction research. Salt intake was assessed by urinary sodium excretion so might be underestimated because salt can be removed from the body through other means, for example, sweat. Nevertheless, this sodium excretion estimation method was shown to be closely correlated with the value determined from 24-h pooled urine.[19] Additionally, because of field conditions, this method was deemed superior to an estimation from a dietary questionnaire.

CONCLUSION

The interaction of ACE genotype and salt intake on hypertension was observed for both men and women among Chinese Kazakhs, albeit in different ways. Application comparison of the two models should be of interest to epidemiologists and geneticists for cG×E exploration, and revealed that the GPLTR model might be more suitable to help elucidate the etiology of complex diseases. Further research is needed to determine the causes for the observed differences behind the interaction mechanisms for men and women.

Acknowledgements We thank the government of Altay City and Hong Dun Town and the Health Bureau of Altay City who supported this project, the village head and staff of each sample village, clinical doctors and nurses of town hospital and people hospital in Altay region, Kazakh college students from Xinjiang Medical University who participated in the surveys, and all participants of this study.

Funding This work was supported by National Natural Science Foundation of China (81273181).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Committee of Institute of Basic Medical Sciences Chinese Academy of Medical Sciences.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Contributors YW and BZ abstracted data, performed data analysis and interpreted results with substantial contributions from the other co-authors, and drafted the manuscript, which was critically revised by all co-authors. JJ and LH conceived the study and assisted in data analysis, design of figures and tables. WH, FX and YW assisted in data analysis and interpretation, added important background knowledge and improved the manuscript by repeated readings and rephrasing as well as critical discussions of the intellectual content. MZ, YT, SL, WW, KA and CQ contributed to the data collection, quality control, general editing purposes and comments regarding medical terminology and critical review of the manuscript. ZW, YH and LW mainly performed statistical analyses, designed tables and figures and reviewed the manuscript. All authors have read and approved the final version of this manuscript for submission.

REFERENCES

1. Padmanabhan S, Caulfield M, Dominiczak AF. Genetic and Molecular Aspects of Hypertension. *Circ Res* 2015;116(6):937-59.
2. Te RL, van Esch JH, Roks AJ, et al. Hypertension: renin-angiotensin-aldosterone system alterations. *Circ Res* 2015;116(6):960-75.
3. Adrogué HJ, Madias NE. Sodium and potassium in the pathogenesis of hypertension. *N Engl J Med* 2007;356(19):1966-78.
4. Sayed-Tabatabaei FA, Oostra BA, Isaacs A, et al. ACE polymorphisms. *Circ Res* 2006;98(9):1123-33.
5. Ned RM, Yesupriya A, Imperatore G, et al. The ACE I/D polymorphism in US adults: limited evidence of association with hypertension-related traits and sex-specific effects by race/ethnicity. *Am J Hypertens* 2012;25(2):209-15.
6. O'Donnell CJ, Lindpaintner K, Larson MG, et al. Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. *Circulation* 1998;97(18):1766-72.
7. Ji LD, Zhang LN, Shen P, et al. Association of angiotensinogen gene M235T and angiotensin-converting enzyme gene I/D polymorphisms with essential hypertension in Han Chinese population: a meta-analysis. *J Hypertens* 2010;28(3):419-28.
8. Calculator B. Hypertension and the ACE gene. http://www.gbhealthwatch.com/GND-Hypertension-ACE.php (accessed 5 Jul 2016).
9. Fg VDK, de Jong PE, Henning RH, et al. Enhanced responses of blood pressure, renal function, and aldosterone to angiotensin I in the DD genotype are blunted by low sodium intake. *J Am Soc Nephrol* 2002;13(4):1025-33.
10. Zhang L, Miyaki K, Araki J, et al. Interaction of angiotensin I-converting enzyme insertion-deletion polymorphism and daily salt intake influences hypertension in Japanese men. *Hypertens Res* 2006;29(10):751-8.
11. Tabery J. Debating interaction: the history, and an explanation. *Int J Epidemiol* 2015;44(4):1117-23.
12. Dick DM, Agrawal A, Keller MC, et al. Candidate gene-environment interaction research: reflections and recommendations: Brookings Institution 37-59 p.
13. Aliev F, Latendresse SJ, Bacanu SA, et al. Testing for Measured Gene-Environment Interaction:

Problems with the use of Cross-Product Terms and a Regression Model Reparameterization Solution. *Behav Genet* 2014;44(2):165-81.

14. Ganzach Y. Misleading interaction and curvilinear terms. *Psychol Methods* 1997;2(3):235-47.
15. Mbogning C, Perdry H, Toussile W, et al. A novel tree-based procedure for deciphering the genomic spectrum of clinical disease entities. *J Clin Bioinforma* 2014;4(1):1-11.
16. Jiang J, Zhang B, Zhang M, et al. Prevalence of conventional cardiovascular disease risk factors among Chinese Kazakh individuals of diverse occupational backgrounds in Xinjiang China. *Int J Cardiol* 2015;179:558-60.
17. Hou L, Zhang M, Han W, et al. Influence of Salt Intake on Association of Blood Uric Acid with Hypertension and Related Cardiovascular Risk. *PLoS One* 2016;11(4).
18. Wang W, Weizhong Z, Ninglin S, et al. Chinese Guidelines for Blood Pressure Measurement. *Chinese Journal of Hypertension* 2011, 6(12):1101-15.
19. Kawano Y, Tsuchihashi T, Matsuura H, et al. Report of the Working Group for Dietary Salt Reduction of the Japanese Society of Hypertension: (2) Assessment of salt intake in the management of hypertension. *Hypertens Res* 2007;30(10):887-93.
20. Wang X, Wang S, Lin R, et al. Relationship of I/D polymorphism of angiotensin converting enzyme gene with hypertension in Xinjiang Kazakh isolated group. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2003;20(3):253-5.
21. Zhang B, Simayi A, Sun Y, et al. Interaction Between Genetic and Environmental Risk Factors on Essential Hypertension in Kazakh Population. *J Med Res* 2009.
22. Stankovic A, Zivkovic MD. Angiotensin I-converting enzyme gene polymorphism in a Serbian population: a gender-specific association with hypertension. *Scandinavian Journal of Clinical & Laboratory Investigation* 2009;62(6):469-75.
23. Chiang FT, Chern TH, Lai ZP, et al. Age- and gender-dependent association of the angiotensin-converting enzyme gene with essential hypertension in a Chinese population. *J Hum Hypertens* 1996;10(12):823-6.
24. Busjahn A, Knoblauch H, Knoblauch M, et al. Angiotensin-Converting Enzyme and Angiotensinogen Gene Polymorphisms, Plasma Levels, Cardiac Dimensions A Twin Study A Twin Study. *Hypertension* 1997;29(2):165-70.
25. Krege JH, John SWM, Langenbach LL, et al. Male–female differences in fertility and blood pressure in

ACE-deficient mice. *Nature* 1995;375(6527):146-8.

26. Fornage M, Amos CI, Kardia S, et al. Variation in the region of the angiotensin-converting enzyme gene influences interindividual differences in blood pressure levels in young white males. *Circulation* 1998;97(97):1773-9.

27. O'Donnell M, Mente A, Yusuf S. Sodium intake and cardiovascular health. *Circ Res* 2015;116(6):1046-57.

28. Giner V, Poch E, Bragulat E, et al. Renin-Angiotensin System Genetic Polymorphisms and Salt Sensitivity in Essential Hypertension. *Hypertension* 2000;35(1 Pt 2):512-7.

29. Caspi A, Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2010;301(5631):386-9.

30. Rutter M. Some of the complexities involved in gene-environment interplay. *Int J Epidemiol* 2015;44(4):1128-9.

31. Chair MR. Chapter 12. Gene–Environment Interaction in Complex Diseases: Asthma as an Illustrative Case: John Wiley & Sons, Ltd; 2008. 184-97.

32. Keller MC. Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biol Psychiatry* 2014;75(1):18-24.

33. Bastone L, Reilly M, Rader DJ, et al. MDR and PRP: a comparison of methods for high-order genotype-phenotype associations. *Hum Hered* 2004;58(2):82-92.

34. Chen J, Yu K, Hsing A, et al. A partially linear tree ‐ based regression model for assessing complex joint gene–gene and gene– environment effects. *Genet Epidemiol* 2007;31(3):238–51.

35. Mbogning C, Perdry H, Broët P. A Bagged, Partially Linear, Tree-Based Regression Procedure for Prediction and Variable Selection. *Hum Hered* 2015;79(3-4):182-93.

Figure legends

Figure 1 Flow chart for the 916 enrolled independent participants.

Figure 2 Optimal trees obtained from 342 men and 574 women using AIC and BIC selection criteria. Leaves are denoted by rectangles and the number in each node represents the number of participants falling in this leaf.

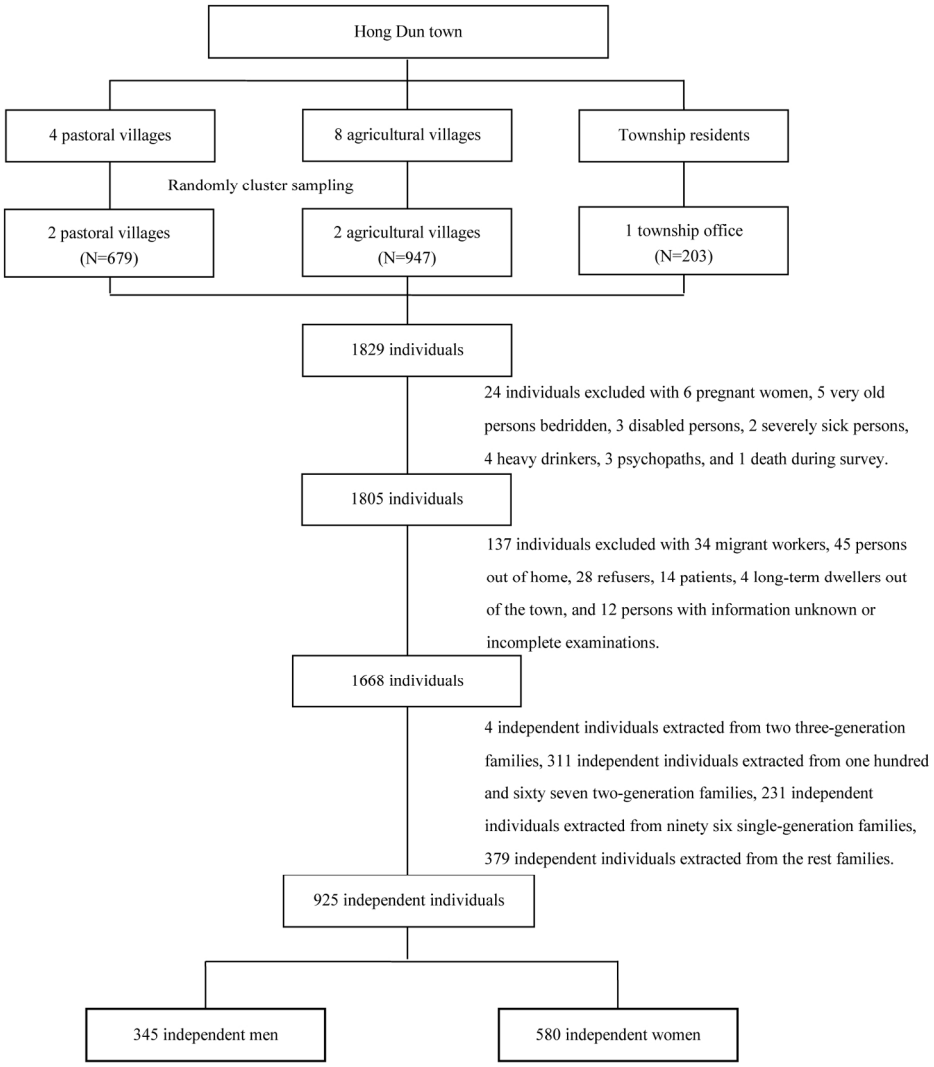

Figure 1 Flow chart for the 916 enrolled independent participants.

167x181mm (300 x 300 DPI)

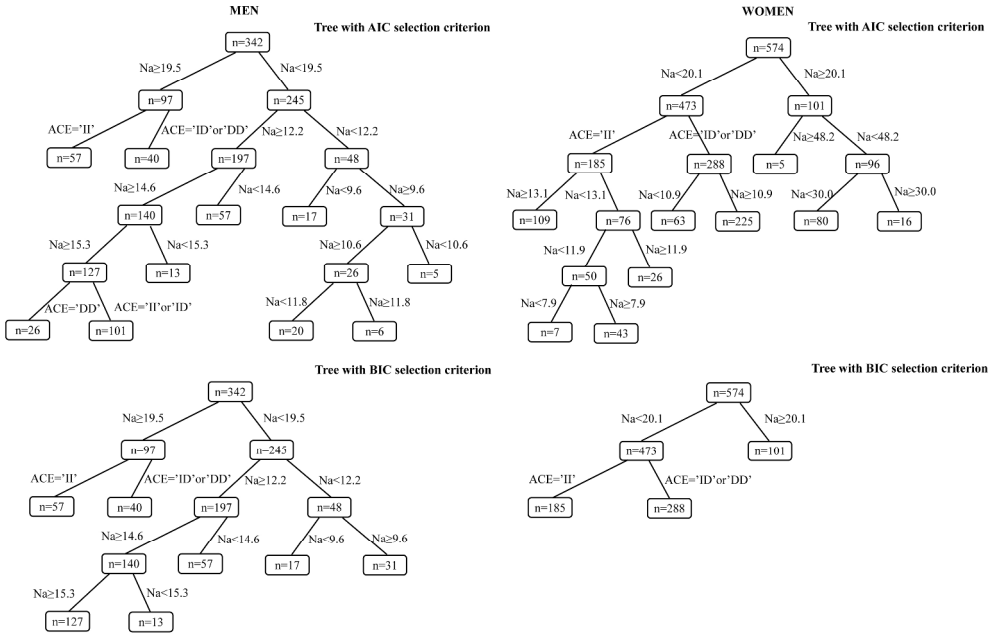

Figure 2 Optimal trees obtained from 342 men and 574 women using AIC and BIC selection criteria. Leaves are denoted by rectangles and the number in each node represents the number of participants falling in this leaf.

292x188mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-6
Bias	9	Describe any efforts to address potential sources of bias	5-6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8
		(*b*) Describe any methods used to examine subgroups and interactions	8
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Results			**Page No**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	5
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8
		(b) Indicate number of participants with missing data for each variable of interest	8
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	8-9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-13
		(*b*) Report category boundaries when continuous variables were categorized	8-13
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-13
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14-17
Generalisability	21	Discuss the generalisability (external validity) of the study results	17-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: results from a population-based cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-014246.R2
Article Type:	Research
Date Submitted by the Author:	16-Mar-2017
Complete List of Authors:	Wang, Yuyan; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Zhang, Biao; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hou, Lei; Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, , Department of Epidemiology and Biostatistics Han, Wei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Xue, Fang; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Yanhong; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Tang, Yong; The People's Hospital in Altay Region Liang, Shaohua; The People's Hospital in Altay Region Wang, Weizhi; The People's Hospital in Altay Region Asaiti, Kuliqian; Hong Dun Town hospital, Altay Region Wang, Zixing; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Hu, Yaoda; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Wang, Lei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics Qiu, Changchun; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences Zhang, Mingtao; The People's Hospital in Altay Region Jiang, Jingmei; Chinese Academy of Medical Sciences Institute of Basic Medical Sciences, Department of Epidemiology and Biostatistics
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Research methods, Genetics and genomics
Keywords:	Hypertension < CARDIOLOGY, Interaction, ACE genotype, Salt intake, Logistic model, GPLTR model

Interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs: results from a population-based cross-sectional study

Yuyan Wang[1†], Biao Zhang[1†], Lei Hou[1], Wei Han[1], Fang Xue[1], Yanhong Wang[1], Yong Tang[2], Shaohua Liang[2], Weizhi Wang[2], Kuliqian.Asaiti[3], Zixing Wang[1], Yaoda Hu[1], Lei Wang[1], Changchun Qiu[1], Mingtao Zhang[2*], Jingmei Jiang[1*]

[1] *Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, China*

[2] *The People's Hospital in Altay Region, Xinjiang, China*

[3] *Hong Dun Town hospital, Altay Region, Xinjiang, China*

[†] These authors contributed equally to this article.

[*] Corresponding Authors:

Prof. Jingmei Jiang, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, 5 Dong Dan San Tiao, Beijing, 100005, China. E-mail: jingmeijiang@ibms.pumc.edu.cn. Tel: +86 10 69155935, Fax: +86 10 65123696.

Mr. Mingtao Zhang, The People's Hospital in Altay Region, Xinjiang, China. E-mail: altrmyy@163.com. Tel:+86018896519163.

ABSTRACT

Objectives: To explore the effect of interaction between angiotensin I-converting enzyme (ACE) genotype and salt intake on hypertension among Chinese Kazakhs, and to compare applications of interactions between logistic model and generalized partially linear tree-based regression (GPLTR) model.

Design: Population-based cross-sectional study.

Setting: Hong Dun, North Xinjiang, China.

Participants: Non-consanguineous Chinese Kazakh participants (n=916, 342 men and 574 women) aged 30 years or over.

Main outcome measures: Association between ACE genotype and hypertension, association between salt intake and hypertension, and interaction of ACE genotype and salt intake on hypertension in two models.

Results: Associations between salt intake and hypertension were different in ACE genotype of II and ID+DD. Under the logistic models, main and interaction effects were not observed for men, but effects were present in opposite directions for women (main effect of ACE: OR=0.20, p=0.003; interaction effect: OR=1.07, p=0.027). Under the GPLTR model, Bayesian information criterion (BIC) trees included both salt intake and ACE genotype as split variables. Individuals with a salt intake≥19.5g/day and ID+DD genotypes had a 3.99-fold (p=0.004) higher risk for hypertension compared with the II genotype for men, while salt intake<20.1g/day and ID+DD genotypes had an OR=0.55 (p=0.014) compared with the II genotype for women.

Conclusions: An interaction of ACE genotype and salt intake on hypertension was

observed among Chinese Kazakhs but in different ways according to sex. The GPLTR model appears to be more suitable for an exploration of interactions in complex diseases.

Strengths and limitations of this study

1. This is one of the few studies to explore the interaction of ACE genotype and salt intake on hypertension, and is also the first such study in Chinese Kazakhs.

2. The design strengths of the study include: the extraction of participants without consanguinity, the use of stratified analyses to test the existence of an interaction effect, and a comparison of the interaction exploration results from two models.

3. The cross-sectional design cannot realize an investigation of the causal pathway of the association between the ACE genotype and salt intake interaction with hypertension, although possible mechanisms were suggested for the interaction effect on hypertension.

4. Salt intake was assessed by urinary sodium excretion and might be underestimated because salt can be removed from the body through other means.

INTRODUCTION

Hypertension is a complex disorder resulting from an interaction of genetic and environmental factors.[1] Blood pressure (BP) is regulated by the renin–angiotensin–aldosterone system (RAAS),[2] which is heavily influenced by dietary salt intake. Indeed, dietary salt is recognized as the most important environmental contributor to hypertension.[3] The most common genetic variation in RAAS is the insertion/deletion polymorphism of the angiotensin I-converting enzyme gene (ACE),[4] and its association with BP has been reported to be dependent on sex and race/ethnicity.[4-7] Although the presence of an interaction between ACE genotype and dietary salt intake has been supported by research findinges,[8 9] few reports have investigated the effect of this interaction on hypertension in different populations.[10]

Measurements and explanations of candidate gene–environment interaction (cG×E) in populations have long been explored,[11] resulting in a general consensus that statistical concerns about their modeling are often overlooked.[12] Popular logistic regression models including interaction terms can be problematic,[13 14] but the newly introduced generalized partially linear tree-based regression (GPLTR) model, which combines tree structure and linear adjustment, may offer more flexibility and power to the exploration of interactions in complex diseases.[15]

The aims of this study were to explore the interaction of ACE genotype and salt intake on hypertension among Chinese Kazakhs, who have a high prevalence of hypertension and a high dietary salt intake, and to compare the logistic model and the GPLTR model in interaction applications.

METHODS

Study participants

The design scheme, sampling process and data collection of the Chinese Xinjiang Altay Kazakh Heart Study have been presented elsewhere.[16 17] Hong Dun town is located in the urban rural fringe of Altay prefecture in North Xinjiang, China. From October 2012 to February 2013, a total of 1805 individuals aged 30 years or over were recruited, and 1668 participants from 601 families completed the baseline survey and examination under standard procedures. Participants were required to have at least three paternal/maternal biological generations living in the same region without a history of intermarriage. Non-consanguineous spouses without consanguinity and only one member of biological relatives from each family were extracted to ensure independence (figure 1). Approval for the study was given by the Ethics Committee of the Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, and written informed consent was obtained from all participants prior to enrollment.

Study variables

We complied with current Chinese Guidelines for Blood Pressure Measurement.[18] A quiet room in a village clinic was selected for BP measurements and all people in the room were required to keep quiet. Participant were comfortably seated, with legs uncrossed and the back and arm supported, such that the middle of the cuff on the upper arm was at the level of the right atrium (the mid-point of the sternum). BP was measured with an appropriate arm cuff and a mercury column sphygmomanometer on the left arm after a resting period of at least 10 min in the supine position. All measurements used the same standard procedure and calibrated mercury sphygmomanometers, and were conducted by trained and qualified staff. Currently, a large-size arm cuff for a mercury column sphygmomanometer is not available in the market of

China, so radial artery auscultation on the upper forearm was used for obese individuals with large upper arms, as recommended by BP measurement guidelines.[18] BP was defined as the mean of two readings, or three readings if there was a difference of more than 5 mmHg between the initial readings. Participants were classified as hypertensive if they met at least one of the following three criteria: a) a systolic BP ≥ 140 mmHg; b) a diastolic BP ≥ 90 mmHg; c) the use of antihypertensive medications within the last 2 weeks.

Venous blood samples were drawn after an overnight fast of at least 10 hours and centrifuged within 20 minutes, then divided into two. One sample was used for biochemistry assays performed within 4 hours, and the other was stored in a portable refrigerator at -20°C then transferred as soon as possible to Beijing, where it was kept at -80°C for genetic analysis. D and I ACE alleles were identified by polymerase-chain-reaction (PCR) amplification of the respective fragments from ACE intron 16 and size fractionation and visualization by electrophoresis. The ACE genotype was classified as II, ID or DD depending on the presence of each allele, with a 99.76% genotyping success rate.

The second urine sample after waking was collected, and urinary sodium was measured with a Caretium XI-921CT automatic electrolyte analyzer (Shenzhen, China). Salt intake was assessed by urinary sodium excretion per day from the amounts of urinary sodium, urine creatinine (Cr), and the 24-h urinary Cr excreted as estimated from height, body weight and age,[19] for which each 43 mmol of sodium is approximately equivalent to 1 g of sodium or 2.5 g of salt (sodium chloride).

Other covariates were also collected using standard procedures. Demographic information (including age, sex, occupation and education), and information about diet, alcohol

consumption, cigarette smoking and medical history were obtained from a face-to-face questionnaire survey. Body weight and height were measured in light indoor clothes and without shoes. Body mass index (BMI) was calculated as body weight in kilograms divided by height in meters squared (kg/m^2). Fasting blood glucose (FBG) was collected according to standardized protocols by a BECKMAN COULTER AU2700 clinical chemistry analyzer (CA, USA).

The GPLTR model and selection criteria

Y was denoted as the outcome of interest, X were the confounding variables to be modeled linearly, and G were the group of risk factors whose interaction effect was to be approximated by the tree-based structure. The main purpose of this GPLTR model was to test the combined effect of G on Y while adjusting for the confounders X; the model was expressed as

$$g(E(Y|X,Z)) = X'\theta + \beta_T F(T(Z))$$

where $g(.)$ is a specific link function (generalized linear model), such as the logit link function $g(\mu) = \log(\frac{\mu}{1-\mu})$ for a binary outcome, and $F(T(Z))$ is the vector of indicator variables signifying the leaves of the tree $T(Z)$.

In this study, we used Y to represent the presence of hypertension, G to be the ACE genotype and the estimated salt intake, and X to be the environmental risk factors having an apparent impact on the hypertension rate. First, we constructed a maximal tree with successive steps of initialization, iteration and ending conditions. The goodness of a candidate split was assessed for each node by the deviance of a generalized linear model, which equals the test deviance of the parent node minus the sum of the deviance of the two child nodes.

Second, the sequence of nested candidate subtrees was created, and Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to determine the best tree in the sequence.

Statistical analysis

All analyses were conducted separately by sex. Basic characteristics were compared with regard to the status of hypertension (table 1). Variables significantly different between hypertension and normotension were chosen as confounders, therefore smoking and family history of hypertension were not included as confounders in further analyses. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the associations of the ACE polymorphism with hypertension (table 2). To examine the presence of an interaction, associations of salt intake with hypertension were illustrated as stratified by ACE genotype as II and ID+DD (table 3). To compare the effect of the interaction, two models were applied (table 4), a logistic regression model, and a GPLTR model adjusting for the same covariates to extract the high-risk group of hypertension using AIC and BIC (figure 2). Description and preliminary analyses were carried out by SAS Version 9.3 software (SAS Institute, Cary, NC, USA), and GPLTR models were conducted by R Version 3.1.3 with the 'GPLTR' package. One-sided hypotheses were established for trend tests based on the logistic model, and all p-values were two-tailed and considered significant at ≤ 0.05.

RESULTS

This study enrolled 916 independent participants (342 men and 574 women) who had complete data for all study variables. The prevalence of hypertension was 49.4% for men and

35.9% for women. Participant characteristics are shown in table 1. Briefly, participants with hypertension had increased age, BMI, and FBG compared with normotensive participants for both sexes.

Table 1 Basic characteristics of enrolled 916 participants by sex and hypertension status

	Men (N=342)			Women (N=574)		
	Hypertension (n=169)	Normotension (n=173)	p	Hypertension (n=206)	Normotension (n=368)	p
Age, years	50.1 ± 12.1	42.8 ± 10.5	<0.001	51.0 ± 11.2	39.6 ± 7.8	<0.001
Occupation, n (%)						
Nomad	68 (40.2)	67 (38.7)	0.900	87 (42.2)	114 (31.0)	0.001
Farmer	75 (44.4)	81 (46.8)		94 (45.6)	169 (45.9)	
City worker	26 (15.4)	25 (14.5)		25 (12.2)	85 (23.1)	
Education, n (%)			0.028			<0.001
Primary school and below	61 (36.1)	40 (23.1)		75 (36.4)	63 (17.1)	
Junior middle school	69 (40.8)	89 (51.5)		86 (41.8)	161 (43.8)	
High school and above	39 (23.1)	44 (25.4)		45 (21.8)	144 (39.1)	
BMI, kg/m^2	26.9 ± 4.8	24.6 ± 3.4	<0.001	28.7 ± 5.4	25.8 ± 4.5	<0.001
Smoking, n (%)	116 (69.8)	133 (76.9)	0.140	12 (5.8)	11 (3.0)	0.097
Drinking, n (%)	53 (31.4)	38 (22.0)	0.049	0 (0.0)	0 (0.0)	—
SBP, mm Hg	152.3 ± 19.2	122.2 ± 9.5	<0.001	154.1 ± 21.7	117.8 ± 9.8	<0.001
DBP, mm Hg	95.7 ± 11.2	77.9 ± 7.2	<0.001	94.1 ± 11.4	75.5 ± 7.7	<0.001
Family history of hypertension, n (%)	92 (57.1)	94 (56.3)	0.876	116 (60.1)	213 (58.2)	0.663
FBG, mmol/L	5.4 ± 0.8	5.2 ± 0.4	<0.001	5.3 ± 0.8	5.0 ± 0.5	<0.001
Salt intake, g/day	19.5 ± 26.7	18.3 ± 15.9	0.599	17.1 ± 7.4	15.9 ± 9.7	0.129

BMI: body mass index; SBP: systolic blood pressure; DBP: diastolic blood pressure; FBG: fasting blood glucose.
Values are means ± SD and frequency (percent).

Independent associations

The independent associations of the ACE genotype and salt intake with hypertension are shown in table 2. The D allele had a protective effect on hypertension in women, with the DD genotype having an OR=0.47 (95% CI: 0.25-0.88) compared with the II genotype. This protective effect was not seen in men, with the DD genotype having an OR=1.27 (95% CI: 0.66-2.46) compared with the II genotype. Salt intake by quartiles did not correlate with hypertension for either sex (p for trend=0.095 for men and 0.152 for women).

Table 2 Association of ACE genotype and salt intake with risk of hypertension

	ACE Genotype				Quartiles of salt intake with median (interquartile range), g/day				
	II (n=356)	ID (n=399)	DD (n=161)	p	Q1	Q2	Q3	Q4	p for linear trend
Men	128	148	66		11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
HTN, %	46.9	49.3	54.6		48.8	47.1	44.2	57.7	
Crude OR (95% CI)	1.00	1.10 (0.69, 1.77)	1.36 (0.75, 2.47)	0.327	1.00	0.93 (0.51, 1.70)	0.83 (0.46, 1.51)	1.43 (0.78, 2.61)	0.334
Age-adjusted OR (95% CI)	1.00	1.08 (0.66, 1.78)	1.31 (0.70, 2.46)	0.416	1.00	0.95 (0.50, 1.79)	0.88 (0.47, 1.66)	1.79 (0.95, 3.39)	0.105
Multi-adjusted OR*(95% CI)	1.00	1.01 (0.60, 1.72)	1.24 (0.64, 2.40)	0.567	1.00	0.98 (0.50, 1.95)	0.90 (0.46, 1.76)	1.94 (0.97, 3.90)	0.099
Women	228	251	95		10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
HTN, %	39.9	35.1	28.4		28.5	36.4	35.0	43.8	
Crude OR (95% CI)	1.00	0.81 (0.56, 1.18)	0.60 (0.36, 1.00)	0.048	1.00	1.44 (0.87, 2.36)	1.35 (0.82, 2.23)	1.95 (1.20, 3.19)	0.013
Age-adjusted OR (95% CI)	1.00	0.71 (0.46, 1.09)	0.50 (0.27, 0.92)	0.017	1.00	1.59 (0.89, 2.85)	1.25 (0.69, 2.26)	1.87 (1.06, 3.30)	0.073
Multi-adjusted OR* (95% CI)	1.00	0.71 (0.45, 1.13)	0.47 (0.25, 0.87)	0.014	1.00	1.36 (0.74, 2.51)	1.11 (0.59, 2.09)	1.64 (0.89, 2.99)	0.178

HTN: hypertension prevalence;

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose; Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

Interaction effect

Stratified analysis

Associations between salt intake and hypertension stratified by ACE II genotype and ID+DD genotypes were evaluated (table 3). No correlation was observed in the II genotype group for either men (p for trend=0.967) or women (p for trend=0.904). An increase in hypertension with a higher salt intake was observed in the ID+DD genotypes group for both men (p for trend=0.028) and women (p for trend=0.045).

Logistic model

Results from the logistic regression model adjusted for multiple confounders are shown in table 4. For men, no main effects of the ACE genotype (OR=1.33, p=0.426) or salt intake (OR=1.00, p=0.415) were observed, and there was no interaction effect (OR=0.99, p=0.464). For women, no main effect of salt intake (OR=0.99, p=0.523) was observed, but a main effect of the ACE genotype (OR=0.20, p=0.003) and an interaction effect (OR=1.07, p=0.027) were observed to be present in opposite directions. Recessive model was also built as sensitivity analysis, and ORs were turned as reciprocals as expected, which confirms the robustness of the findings.

GPLTR model

All iterative procedures converged after 15 iterations, and trees selected by AIC and BIC with different degrees of pruning resulted in 10 and seven leaves for men, and nine and three leaves for women, respectively (figure 2). BIC trees were chosen for the final interpretation based on their stability. Risk assessments of trees are shown in table 4. For men, leaves are identified from the left as:

(1) Node 1: 57 participants characterized by salt intake ≥ 19.5 g/day and ACE II genotype;

(2) Node 2: 40 participants characterized by salt intake ≥ 19.5 g/day and ACE ID+DD genotype;

(3) Node 3: 127 participants characterized by 15.3 g/day ≤ salt intake < 19.5 g/day;

(4) Node 4: 13 participants characterized by 14.6 g/day ≤ salt intake < 15.3 g/day;

(5) Node 5: 57 participants characterized by 12.2 g/day ≤ salt intake < 14.6 g/day;

(6) Node 6: 17 participants characterized by salt intake < 9.6 g/day;

(7) Node 7: 31 participants characterized by 9.6 g/day ≤ salt intake < 12.2 g/day.

Split variables for the selected tree included both salt intake and ACE genotype as II and ID+DD. After adjusting for the same confounders in the multivariable logistic model, the final tree showed that Node 2 had 3.99-fold (p=0.004) higher risk for hypertension compared with Node 1, and the Node 3 had an OR=0.93 (p=0.860). For women, leaves were identified from the left as:

(1) Node 1: 185 participants characterized by salt intake < 20.1 g/day and ACE II genotype;

(2) Node 2: 288 participants characterized by salt intake < 20.1 g/day and ACE ID+DD genotype;

(3) Node 3: 101 participants characterized by salt intake ≥ 20.1 g/day.

Split variables also included both salt intake and ACE genotype as II and ID+DD. Node 2 had an OR=0.55 (*p*=0.014) and Node 3 had an OR=1.26 (*p*=0.444) compared with Node 1.

Table 3 Multivariable-adjusted odds ratios (95% confidence interval) in the ACE genotypes of II and ID+DD by different salt intake level in men and women

Multi-adjusted OR* (95% CI)	Quartiles of salt intake with median (interquartile range), g/day				
	Q1	Q2	Q3	Q4	p for linear trend
Men	11.9 (10.5-13.0)	15.7 (14.8-16.4)	18.3 (17.5-19.1)	23.2 (21.4-25.9)	
II (n=128)	1.00	2.04 (0.66, 6.30)	1.59 (0.51, 5.01)	1.09 (0.37, 3.19)	0.967
ID+DD (n=214)	1.00	0.59 (0.25, 1.41)	0.68 (0.30, 1.58)	3.15 (1.25, 7.93)	0.028
Women	10.4 (9.5-11.3)	13.6 (12.8-14.3)	16.5 (15.9-17.3)	21.4 (19.6-24.7)	
II (n=228)	1.00	1.19 (0.47, 2.99)	0.75 (0.28, 2.04)	1.15 (0.44, 3.00)	0.904
ID+DD (n=346)	1.00	1.55 (0.67, 3.59)	1.53 (0.66, 3.52)	2.35 (1.06, 5.24)	0.045

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

Table 4 ACE genotype and salt intake interaction explorations by logistic model and generalized partially linear tree-based regression model

Logistic model			GPLTR model		
Terms in the model	Multi-adjusted OR* (95% CI)	p	Nodes in the model	Multi-adjusted OR* (95% CI)	p
Men (ROC=0.747)			Men (ROC=0.809)		
Salt intake, g/day	1.00 (0.99, 1.03)	0.415	Node 1: Salt intake ≥ 19.5 and ACE(II)	1.00	
ACE(ID+DD vs. II)	1.33 (0.66, 2.65)	0.426	Node 2: Salt intake ≥ 19.5 and ACE(ID+DD)	3.99 (1.55, 10.26)	0.004
ACE(ID+DD vs. II)×Salt intake	0.99 (0.96, 1.02)	0.464	Node 3: 15.3 ≤ Salt intake < 19.5	0.93 (0.41, 2.09)	0.860
Women (ROC=0.839)			Women (ROC=0.839)		
Salt intake, g/day	0.99 (0.94, 1.03)	0.523	Node 1: Salt intake < 20.1 and ACE(II)	1.00	
ACE(ID+DD vs. II)	0.20 (0.07, 0.59)	0.003	Node 2: Salt intake < 20.1 and ACE(ID+DD)	0.55 (0.34, 0.89)	0.014
ACE(ID+DD vs. II)×Salt intake	1.07 (1.01, 1.13)	0.027	Node 3: 20.1 ≤ Salt intake	1.26 (0.70, 2.25)	0.444

*Men adjusted for age, education, body mass index, drinking, and fasting blood glucose;
Women adjusted for age, education, body mass index, occupation and fasting blood glucose.

DISCUSSION

There are three main findings from the present study. First, we observed that the association between ACE gene polymorphism and hypertension was sex-specific. Second, stratified analysis indicated that an interaction between ACE genotype and salt intake existed for both sexes. Third, an application comparison of the two models suggested that the traditional logistic model was not appropriate in this situation, and that the GPLTR model could effectively determine the high-risk characteristics of hypertension.

The distribution of the ACE genotype observed in this study is similar to that reported for an isolated group of Kazakhs in 2003,[20] with a DD genotype frequency of less than 20%, which differs from that observed in other studies on Kazakhs[21] and in studies of other ethnicities.[4-7] This difference may arise from the diverse structure of the population and racial heterogeneity. Our current results are consistent with previous studies that reported ACE genotype to be a sex-specific candidate gene for hypertension.[5 6 22 23] No positive association was detected between the D allele and hypertension in men, but a protective effect of the D allele was observed for women, presumably through sexes, ethnicities and other factors including gene-gene or gene-environment interactions.[6] Nonlinear association between salt intake and risk of hypertension was confirmed by restricted cubic spline with automatically computed 'knots' (data not shown). With regard to the interaction of ACE genotype and salt intake on hypertension, a previous study that examined 284 Japanese men using logistic regression analysis[10] reported a negative association for ACE (II+ID vs. DD: OR=0.1, p=0.024) and a positive association for interaction (ACE(II+ID vs. DD)× salt intake level: OR=3.6, p=0.047) in opposite directions, which was similar to our current results when the II

genotype was set as the reference.

Strong evidence from association and linkage studies suggests that ACE has a profound effect on the variance of plasma ACE levels,[24] which maintain BP homeostasis. Among the major factors that activate RAAS, dietary sodium/potassium balance is the most important, and is mainly influenced by dietary salt intake.[3] The sodium status has been reported to affect the ACE I/D polymorphism phenotype, and to enhance BP and renal function responses, while the conversion of aldosterone to angiotensin I with the DD genotype is blunted by low a sodium intake.[8] These reports suggest that the ACE polymorphism interacts with dietary salt intake, with subsequent effects on the regulation of BP. Gene-targeting experiments that inactivate ACE in mice have provided evidence for the sex-specific effect of ACE gene on blood pressure.[25] Population-based research also supports the sex-specific influence of ACE on BP variability.[4 26] Assessment of the dietary habits of individuals in the current study (data not shown) revealed average salt intake of 17.3 g/day, which is far greater than that of other populations and World Health Organization recommendations.[27] Given the high daily salt intake for this population, we suggest that the potential harmful effects of high salt intake on the development of hypertension should be emphasized to Chinese Kazakhs. Previous studies into the relationship between ACE and salt sensitivity indicated possession of the ACE D allele by men may lead to an increased response to a high salt intake;[4 28] thus carriers of the D allele may trigger hypertension if they consume a high salt diet. For women, the mechanism behind the interaction is uncertain; however, the combination of the ACE D allele and a lower salt intake may have a synergistic effect on hypertension.[28]

Previous discussions into gene–environment interactions have historically progressed

through three main stages, including eugenics, IQ, and the serotonin transporter gene controversy.[29] This has resulted in an apparent divide between variation-partitioning and mechanism-elucidation approaches. The former approach interpreted interactions as strictly statistical phenomena, as suggested by Fisher,[11] while the latter emphasized the need to probe and understand the causal mechanisms behind interaction responsible for the developmental process, as suggested by Hogben.[11] For complex diseases like hypertension, failure to pay proper attention to the methodological checks that must be undertaken for cG×E has resulted in some studies making dubious claims.[30 31] Logistic regression requiring main effects and linear conditions is usually restricted to low-order interactions, and exploration relying on cross-product terms to measure the cause of the variation responsible for the interaction can be particularly problematic. We have built a full model using three main factors (i.e., gender, ACE genotype, and salt intake), in which the three main effects and the second-order interaction term are positive, but the three first-order interaction terms are all negative. Moreover, explanatory and confounding variables are entered equally into logistic regression models, which may disperse or conceal the actual main and interactive effects.[32] For example, the main protective effects from the ACE genotype and the risk interaction effects of the ACE genotype and salt intake in women of the current study cannot be readily explained because the ACE genotype has no reported independent effect on hypertension, and rarely shows the opposite effect after interacting with other factors.[8] Therefore, new models suitable for complicated high-order interactions and appropriate confounder adjustments, designed to be distinguished from generalized linear regression models (GLMs), are warranted to reveal gene-environment interactions for complex diseases.[33] The GPLTR model is semi-parametric

and integrates the advantages of the tree structure and GLMs. Its purpose is to decompose a data space recursively into smaller areas defined by a set of explanatory variables with adjustment for confounding variables,[15 34 35] where explanatory and confounding variables are not treated in the same way. The modified procedure [15] can explicitly identify sub-groups of individuals at high risk for hypertension with fewer computational demands, where highlighted features of subgroup provide clues for the mechanisms of disease development. Taking the results of the GPLTR model in the current study as an example, the ACE D allele was suggested to be the trigger for hypertension among men with a high salt intake, and female carriers of the D allele combined with a lower salt intake may offer help protection from the development of hypertension. Our calculation of the 10-fold cross-validation prediction error (data not shown) of the logistic regression model and GPLTR model for both men and women revealed the overall improved accuracy of GPLTR model.

The high-quality study design, a response rate of 92.4% and high aggregation of Chinese Kazakhs in the current study provide good internal validity of the results. Sex-stratified analyses were performed to reveal potential differences in mechanisms behind the effect of the ACE genotype and salt intake interaction on hypertension. The high prevalence of hypertension (49.4% for men and 35.9% for women) allowed for robust comparisons between individuals with hypertension and non-hypertensive controls. Application comparisons between the logistic model and the GPLTR model was based on the same confounding variables by sex, which were cautiously chosen according to baseline results. The study should also be considered in light of its limitations. The cross-sectional design cannot realize investigation of the causal pathway of the association between the ACE genotype and salt

intake interaction with hypertension, so confirmatory studies in other large population-based samples are warranted. The relatively small sample size may cause statistical instability. However, all enrolled participants did not have consanguinity, which is necessary for gene-environment interaction research. Salt intake was assessed by urinary sodium excretion so might be underestimated because salt can be removed from the body through other means, for example, sweat. Nevertheless, this sodium excretion estimation method was shown to be closely correlated with the value determined from 24-h pooled urine.[19] Additionally, because of field conditions, this method was deemed superior to an estimation from a dietary questionnaire.

CONCLUSION

The interaction of ACE genotype and salt intake on hypertension was observed for both men and women among Chinese Kazakhs, albeit in different ways. Application comparison of the two models should be of interest to epidemiologists and geneticists for cG×E exploration, and revealed that the GPLTR model might be more suitable to help elucidate the etiology of complex diseases. Further research is needed to determine the causes for the observed differences behind the interaction mechanisms for men and women.

Acknowledgements We thank the government of Altay City and Hong Dun Town and the Health Bureau of Altay City who supported this project, the village head and staff of each sample village, clinical doctors and nurses of town hospital and people hospital in Altay region, Kazakh college students from Xinjiang Medical University who participated in the surveys, and all participants of this study.

Funding This work was supported by National Natural Science Foundation of China (81273181).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Committee of Institute of Basic Medical Sciences Chinese Academy of Medical Sciences.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Contributors YW and BZ abstracted data, performed data analysis and interpreted results with substantial contributions from the other co-authors, and drafted the manuscript, which was critically revised by all co-authors. JJ and LH conceived the study and assisted in data analysis, design of figures and tables. WH, FX and YW assisted in data analysis and interpretation, added important background knowledge and improved the manuscript by repeated readings and rephrasing as well as critical discussions of the intellectual content. MZ, YT, SL, WW, KA and CQ contributed to the data collection, quality control, general editing purposes and comments regarding medical terminology and critical review of the manuscript. ZW, YH and LW mainly performed statistical analyses, designed tables and figures and reviewed the manuscript. All authors have read and approved the final version of this manuscript for submission.

REFERENCES

1. Padmanabhan S, Caulfield M, Dominiczak AF. Genetic and Molecular Aspects of Hypertension. *Circ Res* 2015;116(6):937-59.
2. Te RL, van Esch JH, Roks AJ, et al. Hypertension: renin-angiotensin-aldosterone system alterations. *Circ Res* 2015;116(6):960-75.
3. Adrogué HJ, Madias NE. Sodium and potassium in the pathogenesis of hypertension. *N Engl J Med* 2007;356(19):1966-78.
4. Sayed-Tabatabaei FA, Oostra BA, Isaacs A, et al. ACE polymorphisms. *Circ Res* 2006;98(9):1123-33.
5. Ned RM, Yesupriya A, Imperatore G, et al. The ACE I/D polymorphism in US adults: limited evidence of association with hypertension-related traits and sex-specific effects by race/ethnicity. *Am J Hypertens* 2012;25(2):209-15.
6. O'Donnell CJ, Lindpaintner K, Larson MG, et al. Evidence for association and genetic linkage of the angiotensin-converting enzyme locus with hypertension and blood pressure in men but not women in the Framingham Heart Study. *Circulation* 1998;97(18):1766-72.
7. Ji LD, Zhang LN, Shen P, et al. Association of angiotensinogen gene M235T and angiotensin-converting enzyme gene I/D polymorphisms with essential hypertension in Han Chinese population: a meta-analysis. *J Hypertens* 2010;28(3):419-28.
8. Calculator B. Hypertension and the ACE gene. http://www.gbhealthwatch.com/GND-Hypertension-ACE.php (accessed 5 Jul 2016).
9. Fg VDK, de Jong PE, Henning RH, et al. Enhanced responses of blood pressure, renal function, and aldosterone to angiotensin I in the DD genotype are blunted by low sodium intake. *J Am Soc Nephrol* 2002;13(4):1025-33.
10. Zhang L, Miyaki K, Araki J, et al. Interaction of angiotensin I-converting enzyme insertion-deletion polymorphism and daily salt intake influences hypertension in Japanese men. *Hypertens Res* 2006;29(10):751-8.
11. Tabery J. Debating interaction: the history, and an explanation. *Int J Epidemiol* 2015;44(4):1117-23.
12. Dick DM, Agrawal A, Keller MC, et al. Candidate gene-environment interaction research: reflections and recommendations: Brookings Institution 37-59 p.
13. Aliev F, Latendresse SJ, Bacanu SA, et al. Testing for Measured Gene-Environment Interaction:

Problems with the use of Cross-Product Terms and a Regression Model Reparameterization Solution. *Behav Genet* 2014;44(2):165-81.

14. Ganzach Y. Misleading interaction and curvilinear terms. *Psychol Methods* 1997;2(3):235-47.
15. Mbogning C, Perdry H, Toussile W, et al. A novel tree-based procedure for deciphering the genomic spectrum of clinical disease entities. *J Clin Bioinforma* 2014;4(1):1-11.
16. Jiang J, Zhang B, Zhang M, et al. Prevalence of conventional cardiovascular disease risk factors among Chinese Kazakh individuals of diverse occupational backgrounds in Xinjiang China. *Int J Cardiol* 2015;179:558-60.
17. Hou L, Zhang M, Han W, et al. Influence of Salt Intake on Association of Blood Uric Acid with Hypertension and Related Cardiovascular Risk. *PLoS One* 2016;11(4).
18. Wang W, Weizhong Z, Ninglin S, et al. Chinese Guidelines for Blood Pressure Measurement. *Chinese Journal of Hypertension* 2011, 6(12):1101-15.
19. Kawano Y, Tsuchihashi T, Matsuura H, et al. Report of the Working Group for Dietary Salt Reduction of the Japanese Society of Hypertension: (2) Assessment of salt intake in the management of hypertension. *Hypertens Res* 2007;30(10):887-93.
20. Wang X, Wang S, Lin R, et al. Relationship of I/D polymorphism of angiotensin converting enzyme gene with hypertension in Xinjiang Kazakh isolated group. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2003;20(3):253-5.
21. Zhang B, Simayi A, Sun Y, et al. Interaction Between Genetic and Environmental Risk Factors on Essential Hypertension in Kazakh Population. *J Med Res* 2009.
22. Stankovic A, Zivkovic MD. Angiotensin I-converting enzyme gene polymorphism in a Serbian population: a gender-specific association with hypertension. *Scandinavian Journal of Clinical & Laboratory Investigation* 2009;62(6):469-75.
23. Chiang FT, Chern TH, Lai ZP, et al. Age- and gender-dependent association of the angiotensin-converting enzyme gene with essential hypertension in a Chinese population. *J Hum Hypertens* 1996;10(12):823-6.
24. Busjahn A, Knoblauch H, Knoblauch M, et al. Angiotensin-Converting Enzyme and Angiotensinogen Gene Polymorphisms, Plasma Levels, Cardiac Dimensions A Twin Study A Twin Study. *Hypertension* 1997;29(2):165-70.
25. Krege JH, John SWM, Langenbach LL, et al. Male–female differences in fertility and blood pressure in

ACE-deficient mice. *Nature* 1995;375(6527):146-8.

26. Fornage M, Amos CI, Kardia S, et al. Variation in the region of the angiotensin-converting enzyme gene influences interindividual differences in blood pressure levels in young white males. *Circulation* 1998;97(97):1773-9.

27. O'Donnell M, Mente A, Yusuf S. Sodium intake and cardiovascular health. *Circ Res* 2015;116(6):1046-57.

28. Giner V, Poch E, Bragulat E, et al. Renin-Angiotensin System Genetic Polymorphisms and Salt Sensitivity in Essential Hypertension. *Hypertension* 2000;35(1 Pt 2):512-7.

29. Caspi A, Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 2010;301(5631):386-9.

30. Rutter M. Some of the complexities involved in gene-environment interplay. *Int J Epidemiol* 2015;44(4):1128-9.

31. Chair MR. Chapter 12. Gene–Environment Interaction in Complex Diseases: Asthma as an Illustrative Case: John Wiley & Sons, Ltd; 2008. 184-97.

32. Keller MC. Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biol Psychiatry* 2014;75(1):18-24.

33. Bastone L, Reilly M, Rader DJ, et al. MDR and PRP: a comparison of methods for high-order genotype-phenotype associations. *Hum Hered* 2004;58(2):82-92.

34. Chen J, Yu K, Hsing A, et al. A partially linear tree ‐ based regression model for assessing complex joint gene–gene and gene– environment effects. *Genet Epidemiol* 2007;31(3):238–51.

35. Mbogning C, Perdry H, Broët P. A Bagged, Partially Linear, Tree-Based Regression Procedure for Prediction and Variable Selection. *Hum Hered* 2015;79(3-4):182-93.

Figure legends

Figure 1 Flow chart for the 916 enrolled independent participants.

Figure 2 Optimal trees obtained from 342 men and 574 women using AIC and BIC selection criteria. Leaves are denoted by rectangles and the number in each node represents the number of participants falling in this leaf.

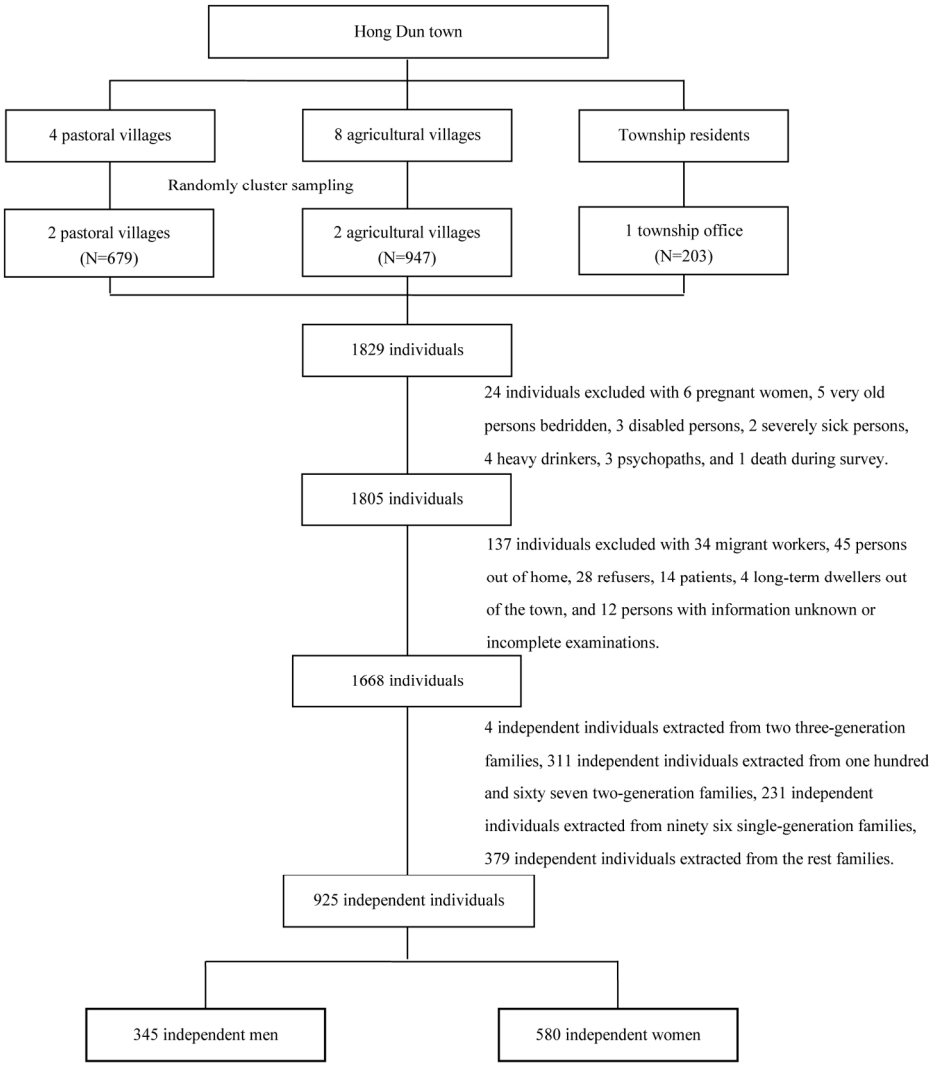

Figure 1 Flow chart for the 916 enrolled independent participants.

167x181mm (300 x 300 DPI)

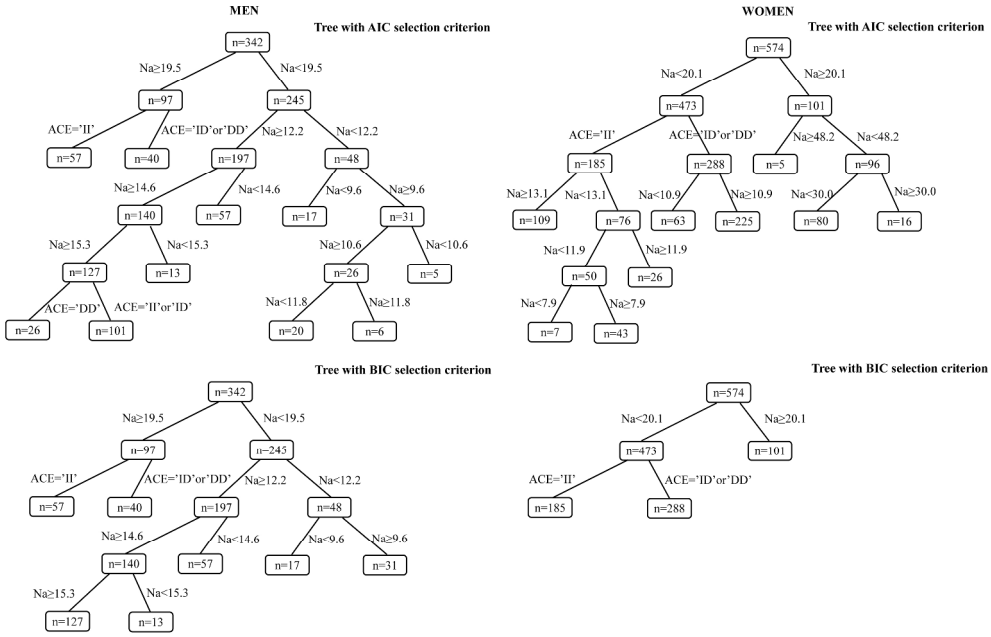

Figure 2 Optimal trees obtained from 342 men and 574 women using AIC and BIC selection criteria. Leaves are denoted by rectangles and the number in each node represents the number of participants falling in this leaf.

292x188mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1-2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	5
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5-6
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	5-6
Bias	9	Describe any efforts to address potential sources of bias	5-6
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5-7
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8
		(*b*) Describe any methods used to examine subgroups and interactions	8
		(*c*) Explain how missing data were addressed	
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	
		(*e*) Describe any sensitivity analyses	

Results			**Page No**
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	5
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	8
		(b) Indicate number of participants with missing data for each variable of interest	8
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	8-9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	8-13
		(*b*) Report category boundaries when continuous variables were categorized	8-13
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11-13
Discussion			
Key results	18	Summarise key results with reference to study objectives	14
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14-17
Generalisability	21	Discuss the generalisability (external validity) of the study results	17-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	18

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.