

# Factor Analysis of Metabolic Syndrome and Its Relationship with the Risk of Cardiovascular Disease in Ethnic Populations in Rural Xinjiang, China

Yu Ren<sup>1,\*</sup>  
Bin Wei<sup>2,\*</sup>  
Yanpeng Song<sup>2</sup>  
Heng Guo<sup>1</sup>  
Xianghui Zhang<sup>1</sup>  
Xinping Wang<sup>1</sup>  
Yizhong Yan<sup>1</sup>  
Jiaolong Ma<sup>1</sup>  
Kui Wang<sup>1</sup>  
Mulatibieke Keerman<sup>1</sup>  
Jingyu Zhang<sup>1</sup>  
Rulin Ma<sup>1</sup>  
Jia He<sup>1,3</sup>  
Shuxia Guo<sup>1,3</sup>

<sup>1</sup>Department of Public Health, Shihezi University School of Medicine, Shihezi, Xinjiang, People's Republic of China;

<sup>2</sup>Department of Social Work, The First Affiliated Hospital of Shihezi University Medical College, Shihezi, Xinjiang, People's Republic of China; <sup>3</sup>NHC Key Laboratory of Prevention and Treatment of Central Asia High Incidence Diseases (First Affiliated Hospital, School of Medicine, Shihezi University), Shihezi, Xinjiang, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Shuxia Guo; Jia He  
Department of Public Health, Shihezi University School of Medicine, North 2th Road, Shihezi, Xinjiang, People's Republic of China  
Tel +8618009932625  
Fax +8609932057153  
Email gsxshzu@sina.com; hejia123.shihezi@163.com

**Background:** This cohort study created a risk equation of CVD for the Uyghur and Kazakh ethnic groups with metabolic syndrome (MetS) in Xinjiang and its associated factors, evaluated the model's feasibility, and provided theoretical support for the prevention and early diagnosis of CVD.

**Methods:** A total of 5655 participants from Xinyuan and Jiashi counties in Xinjiang from 2010 to 2012 were selected, including 3770 and 1885 training and validation samples, respectively. A factor analysis was performed on 975 patients with MetS in the training sample, whereas potential factors related to CVD were extracted from 21 MetS biomarkers. Cox regression was used to create and verify a CVD-risk prediction model based on training samples. The receiver operating characteristic curve was drawn to evaluate the model's prediction efficiency.

**Results:** The cumulative incidence of CVD was 9.20% (training sample, 9.12%; validation sample, 9.36%). Nine potential factors were extracted from the training sample population with MetS to predict the CVD risk: lipid (hazard ratio [HR], 1.205), obesity (HR, 1.047), liver function (HR, 1.042), myocardial enzyme (HR, 1.008), protein (HR, 1.024), blood pressure (HR, 1.027), liver enzyme (HR, 1.012), renal metabolic (HR, 1.015), and blood glucose (HR, 1.010). The area under the curve of the training and validation samples was 0.841 (95% confidence interval [CI], 0.821–0.861) and 0.889 (95% CI, 0.870–0.909), respectively.

**Conclusion:** The CVD prediction model created with nine potential factors in patients with MetS in Kazakh and Uyghur has a good predictive power.

**Keywords:** factor analysis, metabolic syndrome, cardiovascular disease, ethnic populations

## Introduction

Cardiovascular disease (CVD) is a leading cause of death and disability worldwide. In 2018, the number of patients with CVD in China reached 290 million, and CVD accounted for >40% of deaths, ranking the first in the spectrum of death causes, causing serious diseases and socioeconomic burden.<sup>1</sup> Metabolic syndrome (MetS), a group of metabolic disorder syndromes including insulin resistance, central obesity, dyslipidemia, and hypertension,<sup>2</sup> is associated with a series of risk factors that increased the CVD risk. It has been known that CVD has a long incubation period, high mortality rate, and serious complications and cannot be cured completely. Therefore, early risk assessment methods can be used to identify and manage high-risk individuals. A previous report has shown that MetS can be used as a risk assessment tool for CVD.<sup>2</sup>

The American Heart Association currently recommends the use of the Framingham Risk Score (FRS) to predict the 10-year absolute CVD risk in MetS.<sup>3</sup> However, when comparing the ability of MetS score and FRS to predict CVD, studies in China and other countries show inconsistent conclusions due to the presence of many risk factors for CVD that vary among models, affecting their CVD predictive performance. Some studies found that identifying and adding new predictors can effectively improve the CVD prediction ability.<sup>4</sup> However, this model is currently applicable to the Kazakh population only, and its usefulness for other ethnic groups remains unclear. Due to differences in regions, races, CVD composition, and intensity of each risk factor, the CVD risk assessment formula developed by different countries or regions is often not suitable for other populations.

The Uyghur and Kazakhs are the two ethnic groups with the largest populations widely distributed in Xinjiang. Due to their special diet, living environment, and genetic characteristics, their prevalence of chronic diseases, such as MetS, hypertension, and obesity, differ from those of other ethnic groups.<sup>5-7</sup> Therefore, this study selected the Xinjiang Uyghur and Kazakh populations as the research participants, used the factor analysis method to extract factors related to CVD from patients with MetS, established a CVD-risk prediction model suitable for this population, which has important reference significance for the future development of comprehensive prevention and control of CVD in ethnic groups in rural areas of Xinjiang, and provided theoretical support for the prevention and early diagnosis of CVD.

## Materials and Methods

### Study Population

In this study, a multistage stratified cluster random sampling method was used to select populations from Xinyuan County (Kazakh) and Jiashi County (Uyghur) as the study participants. The baseline survey was conducted in 2010 and 2012. In 2013, 2016, and 2017, the follow-up rate was 87.56%, with the average follow-up time of 6.14 (standard deviation [SD], 1.39) years. Those with incomplete information and severe diseases, who were unconscious, unable or unwilling to cooperate, and pregnant, and the floating population were excluded from the study. Finally, 5655 people were included. Two-thirds of the study participants were randomly selected as training samples (3770) and

one-third as validation samples (1885). Factor analysis was performed on 975 participants with MetS of the training samples at the baseline to explore the aggregation pattern of MetS components. After excluding 136 patients with CVD at baseline, Cox regression was used to create the CVD-risk prediction model using the training (3346), and validation samples (1665) were used for verification. All participants were followed up for >5 years and provided written informed consent. The investigation was approved by the ethical review committee of the First Affiliated Hospital of Medical College in Shihezi University (no. shz2010ll01), and the study was conducted in accordance with the relevant guidelines.

### Data Collection

A questionnaire survey was administered to all study participants. The follow-up survey questionnaire was the same as the baseline survey questionnaire, administered face-to-face by uniformly trained investigators. The content included personal information, disease history, family history, lifestyle behavior, height, weight, waist circumference (WC), hip circumference (hip), and blood pressure according to standard methods, and body mass index (BMI) = weight (kg)/(height [m])<sup>2</sup>, and blood serum samples were collected from the participants' elbow veins after an 8-h fast. Indicators, such as total cholesterol (TC), triglyceride (TG), apolipoprotein B (APOB), fasting plasma glucose (FPG), fructosamine (FMN), total protein (TP), albumin (ALB), total bilirubin (TBIL), indirect bilirubin (IBIL), uric acid (UA), serum creatinine (Scr), blood urea nitrogen (urea),  $\alpha$ -hydroxybutyrate dehydrogenase ( $\alpha$ -HBDH), lactate dehydrogenase (LDH), alanine aminotransferase (ALT), and aspartate aminotransferase (AST), were tested using an Olympus 2007 automatic biochemical analyzer in the laboratory of the First Affiliated Hospital of Medical College in Shihezi University.

**Blood pressure:** Investigators used the standard mercury cuff sphygmomanometer to measure systolic blood pressure (SBP) and diastolic blood pressure (DBP) of the upper arm in sitting position three times with international standard method and recorded the readings accurately with an interval of at least 1 min.

**Smoking:** Smoking more than one cigarette per day for >1 year.

**Drinking:** Drinking at least one day per week for >1 year.

## Definition of MetS and CVD

According to the diagnostic criteria recommended by the Joint Interim Statement, MetS is defined as the presence of three or more of the following five items:<sup>8</sup> 1) WC of  $\geq 85$  cm (male),  $\geq 80$  cm (female); 2) TG,  $\geq 1.70$  mmol/L; 3) HDL-C,  $< 1.00$  mmol/L (male),  $< 1.30$  mmol/L (female); 4) SBP,  $\geq 130$  mmHg or DBP,  $\geq 85$  mmHg; and 5) FPG,  $\geq 5.6$  mmol/L.

According to the diagnostic criteria of CVD recommended by the World Health Organization, the following events were identified as CVD:<sup>9</sup> after collecting the baseline data, the participant was hospitalized due to CVD, underwent coronary intervention (cardiac catheterization and diagnosis of coronary heart disease or coronary artery bypass graft), had angina (nitroglycerin after the cohort study was conducted), and died (ICD10); and according to medical records at discharge, unstable angina, myocardial infarction, congestive heart failure, stroke, transient cerebral ischemia, and peripheral vascular disease (abdominal aortic aneurysm, peripheral vascular surgery or intracarotid artery, peripheral vascular surgery, or carotid intima cut). If the same type of event occurs twice or more in one participant, the first event is considered as the ending event.

## Statistical Analysis

Statistical analysis was performed using the SPSS version 20.0 (SPSS Inc., Chicago, IL, USA). Data are expressed as mean (SD) and compared between groups using the Student's *t*-test. Count data are expressed as rate and constituent ratio and compared between groups using the chi-square test. Logarithmic transformation of TGs was also performed. The factor analysis was based on a principal component analysis to establish factor variables (determine the number of factors according to the characteristic root of  $> 1$ ). We used the maximum variance rotation method to rotate the factor matrix to obtain a reasonable explanation of factor naming, cox regression to construct a CVD risk prediction model, and R software (version 4.0.3; <https://www.r-project.org/>) to draw the receiver operating characteristic (ROC) curve, evaluate the feasibility of the model, use training and verification samples for model verification, evaluate the predictive ability of the model, and verify all statistics. Academic tests were all two-sided, with *P*-values of  $< 0.05$  indicating statistical significance.

## Results

A total of 5655 participants (2612 men [46.19%], mean age, 42.46[SD=15.28] years) were included. The prevalence of MetS was 27.69% (training sample, 25.86%; verification sample, 31.35%); the cumulative incidence of CVD was 9.20% (training sample, 9.12%; validation sample, 9.36%). Whether it was training samples or verification samples, age, BMI, WC, hip, TG, APOB, TC, FPG, ALB, TBIL, IBIL, ALT, SBP, and DBP in the MetS groups were higher than those in the non-MetS groups. Other indicators are generally as shown in Table 1.

According to the factor analysis requirements, the Kaiser-Meyer-Olkin test and Bartlett sphere tests were performed on patients with MetS in the baseline population. The Kaiser-Meyer-Olkin test value was  $P = 0.629$ , whereas the Bartlett sphere test value was  $P < 0.001$  (see [Supplementary Table A. 1](#)), indicating that a strong correlation was found among 21 variables, which was consistent with the factor analysis condition. Principle component analysis was used to extract factors. According to the actual clinical significance of each variable, nine independent MetS factors were named, including obesity (BMI, WC, and hip), liver function (TBIL and IBIL), lipid (APOB, TG, and TC), myocardial enzymes ( $\alpha$ -HBDH and LDH), blood pressure (SBP and DBP), renal metabolic (UA, Scr, and urea), protein (TP and ALB), liver enzymes (ALT and AST), and blood glucose (FMN and FPG) levels. The cumulative contribution rate of factor variance indicated that 77.700% of the variation in patients with MetS was attributable to these nine main factors. The factor load matrix was obtained using the maximum orthogonal rotation method (Table 2).

The factor scoring model was calculated according to the factor scoring coefficients as shown in [Supplementary Table A. 2](#). We constructed a cox regression model to predict the CVD risk based on age, sex, smoking, and drinking, and the nine main extracted factors, and found that age, sex, smoking, obesity, liver function, lipid, myocardial enzyme, blood pressure, protein, liver enzyme, renal metabolic, and blood glucose were risk factors for CVD in the training samples. In the verification sample, age, female sex, lipid, blood pressure, renal metabolic, and blood glucose were risk factors for CVD ( $P < 0.05$ ) (Table 3).

An ROC curve was drawn to test the applicability of the model. The area under the ROC curve (AUC) of the training sample was 0.841 (95% confidence interval [CI],

**Table 1** Baseline Characteristics and CVD Incidence in Xinjiang Ethnic Population

	Training Samples (n=3770)			Validation Samples (n=1885)		
	MetS	Non-MetS	P	MetS	Non-MetS	P
Quantity, n (%)	975 (25.86)	2795 (74.14)	–	591 (31.35)	1294 (68.65)	–
Male, n (%)	392 (40.21)	1328 (47.51)	<0.001	265 (44.84)	627 (48.45)	<0.001
Age (years), mean (SD)	47.86 (14.51)	40.66 (14.99)	<0.001	48.21 (14.71)	39.63 (15.03)	<0.001
BMI (kg/m <sup>2</sup> ), mean (SD)	25.10 (4.13)	22.42 (3.18)	<0.001	25.43 (4.17)	22.42 (3.21)	<0.001
WC (cm), mean (SD)	90.51 (10.16)	80.73 (9.19)	<0.001	91.35 (10.09)	80.63 (8.99)	<0.001
Hip (cm), mean (SD)	98.60 (8.15)	93.50 (6.94)	<0.001	99.05 (7.75)	93.28 (7.11)	<0.001
TG (mmol/L), mean (SD)	0.30 (0.18)	0.60 (0.28)	<0.001	0.25 (0.20)	0.22 (0.02)	<0.001
APOB (g/L), mean (SD)	0.92 (0.31)	0.80 (0.27)	<0.001	0.88 (0.34)	0.79 (0.29)	<0.001
TC (mmol/L), mean (SD)	4.47 (1.13)	4.16 (1.02)	<0.001	4.38 (1.03)	4.11 (0.96)	<0.001
FMN (μmol/L), mean (SD)	222.28 (80.53)	210.07 (55.30)	<0.001	215.04 (55.72)	213.83 (71.04)	0.714
FPG (mmol/L), mean (SD)	5.06 (1.67)	4.44 (0.82)	<0.001	5.24 (1.85)	4.53 (0.92)	<0.001
TP (g/L), mean (SD)	72.84 (6.77)	73.14 (7.05)	0.244	70.65 (7.23)	71.43 (6.29)	0.017
ALB (g/L), mean (SD)	42.96 (6.54)	43.74 (6.53)	0.001	41.30 (7.30)	43.05 (6.29)	<0.001
TBIL (μmol/L), mean (SD)	9.30 (4.54)	10.03 (4.68)	<0.001	10.23 (5.82)	10.77 (5.15)	0.043
IBIL (μmol/L), mean (SD)	6.74 (3.74)	7.40 (3.87)	<0.001	7.60 (4.12)	8.06 (3.97)	0.02
UA (μmol/L), mean (SD)	225.28 (78.44)	217.97 (69.24)	0.006	225.32 (80.88)	218.16 (78.87)	0.07
Scr (mmol/L), mean (SD)	57.59 (14.48)	57.43 (14.81)	0.771	59.25 (13.45)	58.54 (15.06)	0.329
Urea (mmol/L), mean (SD)	4.90 (1.48)	4.78 (1.40)	0.023	4.77 (1.27)	4.68 (1.36)	0.187
α-HBDH (IU/L), mean (SD)	140.58 (44.80)	139.67 (47.87)	0.604	124.19 (30.30)	128.39 (50.98)	0.063
LDH (IU/L), mean (SD)	186.24 (65.46)	185.09 (69.56)	0.653	165.53 (36.76)	170.15 (48.24)	0.039
ALT (IU/L), mean (SD)	18.89 (16.88)	16.67 (21.13)	0.003	18.05 (19.97)	15.87 (12.80)	0.004
AST (IU/L), mean (SD)	26.06 (20.58)	25.13 (16.68)	0.158	24.35 (16.61)	25.03 (20.93)	0.486
SBP (mmHg), mean (SD)	135.60 (22.05)	123.30 (19.71)	<0.001	135.88 (21.75)	122.91 (20.84)	<0.001
DBP (mmHg), mean (SD)	86.18 (13.00)	77.75 (13.04)	<0.001	85.45 (12.84)	77.19 (13.94)	<0.001
Smoking, n (%)	243 (24.92)	662 (23.69)	0.436	152 (25.72)	298 (23.03)	0.204
Drinking, n (%)	56 (5.74)	171 (6.12)	0.672	40 (6.77)	96 (7.42)	0.612
Incidence of CVD, n (%)	139 (14.72)	197 (7.20)	<0.001	66 (11.56)	106 (8.37)	0.03

**Abbreviations:** SD, standard deviation; BMI, Body Mass Index; WC, waist circumference; hip, hip circumference; TG, Triglyceride; APOB, Apolipoprotein B; TC, Total Cholesterol; FMN, Fructosamine; FPG, Fasting Plasma Glucose; TP, Total Protein; ALB, Albumin; TBIL, Total Bilirubin; IBIL, Indirect Bilirubin; UA, Uric acid; Scr, Serum creatinine; urea, Blood urea nitrogen; α-HBDH, α-Hydroxybutyrate dehydrogenase; LDH, Lactate dehydrogenase; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; SBP, Systolic blood pressure; DBP, Diastolic blood pressure.

0.821–0.861) (Figure 1A). The AUC of training sample with MetS was 0.811 (95% CI, 0.776–0.846) (Figure 1B) and without MetS was 0.845 (95% CI, 0.818–0.872) (Figure 1C). We used the validation sample to test the independent validity of the constructed CVD prediction model and found that the AUC is 0.889 (95% CI, 0.870–0.909) (Figure 2A). The AUC of validation sample with MetS was 0.869 (95% CI, 0.828–0.909) (Figure 2B) and without MetS was 0.916 (95% CI, 0.895–0.936) (Figure 2C) (see Supplementary Table A. 3).

## Discussion

This study included ethnic groups aged >18 years in Xinjiang, with an average age of 42.46 years. Results showed that the cumulative incidence of CVD in the study participants was 9.20% (8.76% and 10.09% for the training samples and

verification samples, respectively), which was higher than the survey level of a 35–70-year-old population in China (5.95%).<sup>10</sup> Although the study population is older than this study, our study population has the characteristics of high prevalence of hypertension, hyperlipidemia, and obesity. The cumulative incidence of CVD is lower than that of a 40–75-year-old population study abroad (19.03%),<sup>11</sup> which may be related to the study follow-up time and overall younger age of study subjects participants in this article. The prevalence of MetS was 27.69% (training sample, 26.95%; validation sample, 29.17%), which is higher than the national level (24.50%).<sup>12</sup> Our results suggest that the prevalence of CVD and MetS in this population is not optimistic and that effective prevention and treatment are urgently required. We also found that the CVD risk in patients with MetS was 1.79 times higher than that in non-MetS patients, indicating that MetS can

**Table 2** Principal Component Analysis Was Used to Analyze 21 Biochemical Indexes of Patients with MetS

	Obesity	Liver Function	Lipid	Myocardial Enzyme	Blood Pressure	Renal Metabolic	Protein	Liver Enzyme	Blood Glucose
SBP (mmHg)	0.103	0.002	0.030	0.046	0.912	0.020	-0.029	0.065	0.052
DBP (mmHg)	0.103	0.032	-0.023	0.054	0.910	-0.020	0.027	0.024	-0.029
BMI (kg/m <sup>2</sup> )	0.862	0.040	0.063	-0.024	0.126	0.016	0.047	0.049	-0.017
WC (cm)	0.886	0.078	0.096	0.030	0.082	0.052	0.034	0.071	0.032
Hip (cm)	0.900	0.003	0.017	0.014	0.018	0.027	0.015	0.042	0.007
FMN (μmol/L)	-0.043	0.073	0.058	0.048	-0.075	0.006	-0.018	0.066	0.772
FPG (mmol/L)	0.061	-0.038	0.051	-0.026	0.096	0.069	0.089	0.055	0.739
TP (g/L)	-0.006	-0.031	0.108	0.088	0.065	0.830	0.052	0.052	0.038
ALB (g/L)	0.109	0.266	0.076	0.076	-0.080	0.826	0.023	0.106	0.062
TBIL (μmol/L)	0.051	0.951	0.030	0.043	0.028	0.118	0.101	0.090	0.014
IBIL (μmol/L)	0.062	0.960	0.036	0.013	0.011	0.056	0.069	0.074	0.026
UA (μmol/L)	0.278	0.188	0.069	-0.083	0.091	0.220	0.205	0.649	0.006
Scr (mmol/L)	0.138	0.155	0.036	-0.103	0.092	-0.132	0.102	0.817	0.003
Urea (mmol/L)	-0.135	-0.096	0.066	0.138	-0.052	0.163	-0.100	0.704	0.166
α-HBDH (IU/L)	0.010	-0.029	0.029	0.956	0.054	0.093	0.040	-0.043	0.006
LDH (IU/L)	0.008	0.085	0.032	0.950	0.050	0.070	0.066	0.016	0.021
ALT (IU/L)	0.058	0.014	0.057	0.022	-0.035	0.058	0.883	0.075	0.020
AST (IU/L)	0.025	0.146	-0.014	0.081	0.034	0.008	0.877	0.037	0.060
TC (mmol/L)	0.093	0.065	0.858	0.059	0.048	0.188	0.038	0.061	0.087
TG (mmol/L)	0.030	0.102	0.718	-0.046	-0.003	-0.231	0.063	0.098	0.069
APOB (g/L)	0.065	-0.110	0.780	0.064	-0.040	0.319	-0.057	-0.016	-0.014
Variance explained (%)	11.987	9.704	9.128	9.026	8.286	8.085	7.897	7.883	5.704
Cumulative variance (%)	11.987	21.691	30.819	39.845	48.131	56.216	64.113	71.997	77.700

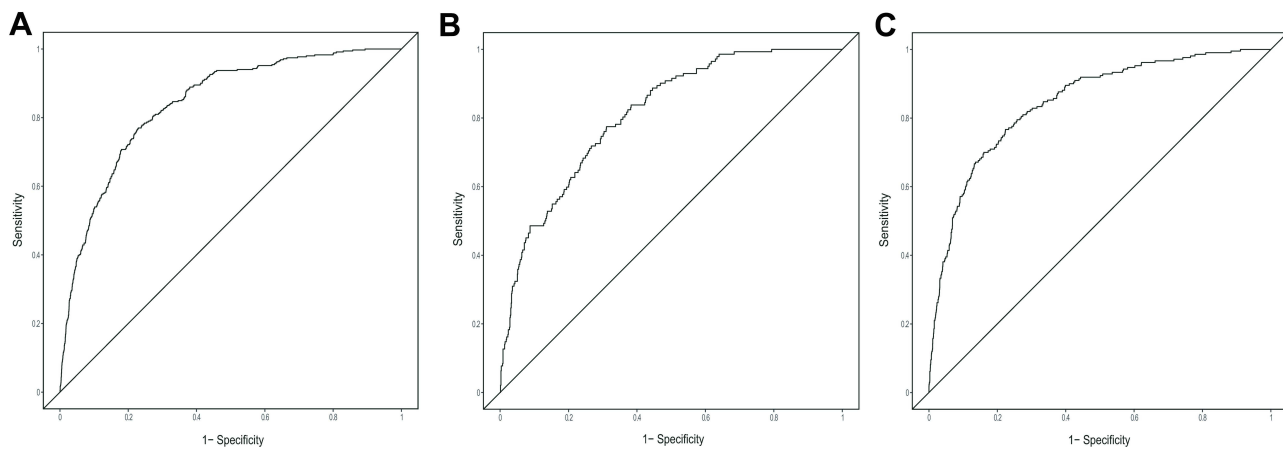
**Notes:** Obesity (BMI, WC, hip), liver function (TBIL, IBIL), lipids (TC, TG, APOB), myocardial enzymes (α-HBDH, LDH), blood pressure (SBP, DBP), renal/metabolic (UA, Ser, urea), protein (TP, ALB), liver enzymes (ALT, AST) and blood glucose (FMN, FPG).

increase the CVD risk, a finding consistent with relevant study results at home and abroad.<sup>13</sup> This study of MetS in the Xinjiang ethnic populations has more important practical significance in predicting the CVD risk.

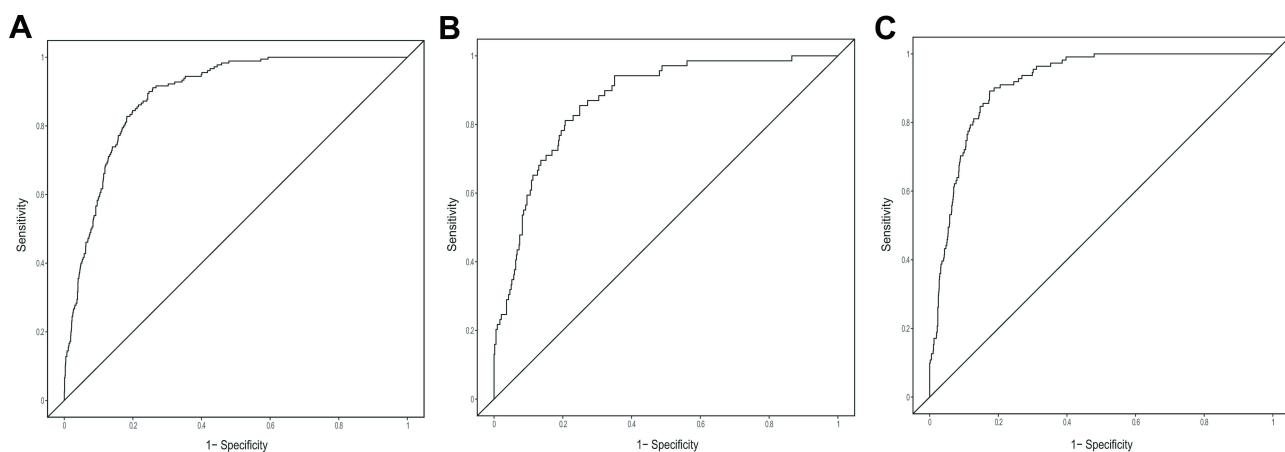
As an effective tool for predicting the incidence of CVD, the CVD overall risk assessment has been widely adopted by domestic and foreign guidelines, such as the FRS in the United States, Systematic Coronary Risk

**Table 3** Cox Regression Analysis Results of Training Samples and Validation Samples

	Training Samples			Validation Samples		
	B	P	HR (95% CI)	B	P	HR (95% CI)
Age	0.041	<0.001	1.042 (1.034–1.050)	0.042	<0.001	1.042 (1.031–1.054)
Sex	0.924	<0.001	2.518 (1.936–3.276)	0.715	<0.001	2.045 (1.431–2.922)
Smoking	0.522	<0.001	1.685 (1.309–2.169)	0.262	0.136	1.300 (0.921–1.835)
Drinking	-0.205	0.399	0.815 (0.506–1.312)	0.002	0.995	1.002 (0.588–1.707)
Obesity	0.045	<0.001	1.047 (1.030–1.063)	0.024	0.052	1.024 (1.000–1.048)
Liver function	0.041	0.001	1.042 (1.017–1.067)	-0.012	0.542	0.988 (0.949–1.028)
Lipid	0.186	0.010	1.205 (1.046–1.388)	0.238	0.046	1.269 (1.004–1.603)
Myocardial enzyme	0.008	0.002	1.008 (1.003–1.013)	-0.030	0.051	0.986 (0.962–1.000)
Blood pressure	0.026	<0.001	1.027 (1.022–1.032)	0.018	<0.001	1.018 (1.011–1.026)
Protein	0.023	0.030	1.024 (1.002–1.045)	0.009	0.561	1.009 (0.979–1.039)
Liver enzyme	0.012	<0.001	1.012 (1.008–1.016)	0.006	0.319	1.006 (0.994–1.018)
Renal metabolic	0.015	0.006	1.015 (1.004–1.026)	0.022	0.018	1.022 (1.004–1.041)
Blood glucose	0.010	0.016	1.010 (1.002–1.018)	0.013	0.047	1.014 (1.000–1.027)



**Figure 1** Predictive ability and validity test of training sample CVD prediction model. **(A)** training sample; **(B)** training sample with MetS; **(C)** training sample without MetS.



**Figure 2** Predictive ability and validity test of validation sample CVD prediction model. **(A)** validation sample; **(B)** validation sample with MetS; **(C)** validation sample without MetS.

Evaluation (SCORE) for European populations, and Pooled Cohort risk Equations (PCEs) derived from the data of Black Americans. However, these models are not suitable for the Chinese population due to differences in ethnic variables and exposure to risk factors of CVD. In 2016, Yang et al constructed a 10-year atherosclerotic CVD-risk prediction model (China-PAR) applicable to the Chinese Han population.<sup>14</sup> Sun compared the above four models based on the Chinese population and found that FRS, SCORE, and PCEs all overestimated the CVD risk in China and determined that the China-PAR may be a more suitable prediction model for Chinese people.<sup>15</sup> However, whether the study is applicable to other ethnic risk predictions remains unclear. However, due to differences in ethnic characteristics and exposure to risk factors of CVD, a predictive model should be established based on population characteristics.

In this study, nine factors were extracted from 21 indicators of patients with MetS based on factor analysis, which explained 77.700% of the original variable information, a value that was higher than those reported by Hajian-Tilaki (men, 65.30%; women, 66.80%).<sup>16</sup> The reason may be that in addition to blood pressure, blood lipids, and blood glucose, the classic components of MetS-related factors, this study also included liver enzymes, liver function, renal metabolism, myocardial enzymes, and other related indicators, which affect the occurrence and development of MetS and are significant to the risk prediction of CVD. A study based on the Dongfeng-Tongji cohort found a U-shaped association among TBIL, IBIL, and CVD incidence.<sup>17</sup> Other studies showed that ALT and AST can be used as independent predictors of CVD.<sup>18</sup> This study included liver function (TBIL and IBIL) and liver enzyme (ALT and AST), and the cumulative variance

contribution rate reached 17.60%. It was reported that the role of liver function indicators in the process of CVD prediction cannot be ignored. Chronic kidney disease is one of the reasons for the CVD prevalence. A national study in South Korea has found a significant correlation between UA and CVD risk.<sup>19</sup> A U-shaped association was observed between male UA levels and CVD risk, whereas a J-shaped association was observed between female UA levels and CVD risk. In this study, the renal metabolism, including UA, Scr, and urea, can more comprehensively reflect the renal metabolism of this population and improve the accuracy of the CVD prediction model. Myocardial enzymes can accurately determine the patient's myocardial injury and have a high diagnostic value for CVD.<sup>20</sup> In the past, factor analysis rarely involved myocardial enzymes. Results of this study not only confirmed the practicability of using MetS to predict CVD but also provided a new idea for the prevention and treatment of CVD.

In this study, cox regression was used to create a CVD prediction model and was found that lipid, blood pressure, renal metabolic, and blood glucose in training and verification samples are all related to the CVD risk. Among them, hyperlipidemia, hypertension, and hyperglycemia are independent risk factors for CVD. The detection rate of abnormal blood lipids in Xinjiang ethnic population is significantly higher than that of Han population,<sup>21</sup> which suggests that the impact of blood lipids on CVD may be more prominent in this population. Chen found that high normal blood pressure would increase the risk of CVD death,<sup>22</sup> and the average systolic blood pressure of the population in this study has reached 127 mmHg, showing that exploring the prediction of hypertension for CVD in this population is of great significance for CVD prevention and treatment. The most common complication of diabetes is CVD, resulting in increased the risk of death. In this study, FPG and FMN were used as blood glucose factors to establish a prediction model. FPG and CVD mortality risk showed a J-type measurement response relationship.<sup>23</sup> FMN can reflect the average blood glucose level of participants within 2–3 weeks, which is very important to predict the CVD risk. These results also showed that obesity, liver function, protein, and liver enzyme were related to CVD risk in the training samples, but no such relationship was found in the validation samples. This is different from Guo's study results.<sup>24</sup> It may be because of the small number of verification samples or may be related

to differences in genetic characteristics and dietary habits among different ethnic groups.

ROC curve is currently a widely used model evaluation method. In this study, the ethnic population was used as the study subject to construct a CVD prediction model, and its predictive power (AUC = 0.84) illustrates that the model is more suitable for CVD risk prediction in the Xinjiang ethnic population. Inappropriate prediction tools will waste many resources, and external crowd verification is a reliable basis for evaluating model feasibility. Jiang conducted external verification of PCE, FRS, and China-PAR based on populations in rural areas of Xinjiang and found that all three models underestimated the 5-year ASCVD risk for women.<sup>25</sup> In this study, validation samples were used to verify the constructed CVD prediction model (AUC = 0.89), indicating that the validation sample validity test results were good. Subsequently, we separately evaluated the predictive performance of the predictive model in training and validation samples regarding patients with and without MetS. All four AUCs are >0.8, indicating that the prediction model is suitable for this population.

## Strengths and Limitations

This study has the following advantages. First, the study participants were Kazakh and Uyghur groups. Results can be applied to both ethnic groups, which is simple and convenient. Second, factor analysis does not simply summarize the number of predictors, to consider the relationship between predictors and results and make study results more practical. Third, this study is based on indicators used in routine physical examination, which are all from the national health examination project supported by the Chinese government. The information bias caused by language communication obstacles is greatly avoided; however, these indicators are easily obtained in actual practice and have better applicability. At the same time, there are some limitations in this study. First, study participants mainly consisted of Uyghur and Kazakh population with low income in rural areas of Southern Xinjiang. Due to their special religious culture and lifestyle, our findings may not be applicable to other groups. Second, this study included only routine physical examination data; not all CVD risk factors were included, which would have decreased the CVD prediction efficiency. Third, this study only analyzed the baseline information of all predictors. However, these factors will be affected by various influencing factors during the follow-up period, such as interventions, pharmacology, and non-pharmacology,

which will affect the CVD results. Lastly, this study lists a variety of CVDs as disease outcomes, with quite distinct etiologies. Targeted disease prediction needs further investigation.

## Conclusion

The incidence of CVD in rural ethnic minorities in Xinjiang is higher than the national level. Nine potential factors related to CVD were extracted from patients with MetS, including lipid, obesity, liver function, myocardial enzymes, protein, blood pressure, liver enzymes, renal metabolism, and blood glucose. The prediction model constructed by these factors can be used to predict the CVD risk in Xinjiang rural ethnic population.

## Abbreviations

$\alpha$ -HBDH,  $\alpha$ -hydroxybutyrate dehydrogenase; ALB, albumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUC, area under the curve; BMI, Body Mass Index; CVD, cardiovascular disease; DBP, Diastolic blood pressure; FMN, fructosamine; FPG, fasting plasma glucose; FRS, Framingham Risk Score; Hip, hip circumference; IBIL, indirect bilirubin; LDH, lactate dehydrogenase; HDL-C, high-density lipoprotein cholesterol; MetS, metabolic syndrome; ROC, receiver operating characteristic; SBP, Systolic blood pressure; Scr, serum creatinine; SD, standard deviation; TBIL, total bilirubin; TC, total cholesterol; TG, Triglyceride; TP, total protein; UA, uric acid; Urea, blood urea nitrogen; WC, waist circumference.

## Data Sharing Statement

The datasets used during the current study are available from the corresponding author on reasonable request.

## Patient Consent for Publication

Patients have consented to publication of anonymised group results.

## Patient and Public Involvement Statement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

## Ethics Approval and Consent to Participate

This study was conducted in accordance with the Declaration of Helsinki and was approved by the

Institutional Ethics Review Board (IERB) of the First Affiliated Hospital of Shihezi University School of Medicine (IERB No. SHZ2010LL01). All of the participants provided their written informed consent prior to the start of the study.

## Acknowledgments

We sincerely thank those who participated in the study. We would also like to acknowledge the clinical laboratory of First Affiliated Hospital of Shihezi University School of Medicine for their work.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas, took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was funded by the National Natural Science Foundation of China (No. 81560551), Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2020-PT330-003) and The Scientific and Technological Research Project in Key Fields of the Corps "Construction of Corps Medical Big Data Platform and Research and Application of Health Management Application Technology" (2018AB017).

## Disclosure

The authors declare that they have no competing interests.

## References

- Hu S, Gao R, Liu L, et al. Summary of the 2018 report on cardiovascular diseases in China. *Chin Circ J*. 2019;34(03):209–220.
- Eckel R, Grundy S, Zimmet P. The metabolic syndrome. *Lancet*. 2005;366(9501):1923. doi:10.1016/S0140-6736(05)67781-1
- Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97(18):1837–1847. doi:10.1161/01.CIR.97.18.1837
- Zhu Z, Liu Y, Zhang C, et al. Identification of cardiovascular risk components in urban Chinese with metabolic syndrome and application to coronary heart disease prediction: a longitudinal study. *PLoS One*. 2013;8(12):e84204. doi:10.1371/journal.pone.0084204
- Ma X, Zhang M, Guo S, et al. Prevalence of hypertension in Uygur, Kazakhs and Han people in rural areas of Xinjiang. *Chin J Hypertens*. 2013;21(12):1164–1168.



6. Qin X, Qiu L, Tang G, Tsoi MF, Cheung BMY. Prevalence of metabolic syndrome among ethnic groups in China. *BMC Public Health*. 2020;20(1). doi:10.1186/s12889-020-8393-6
7. He J, Guo H, Ding Y, et al. Epidemiological study on overweight and obesity among rural adult residents in Hazakh, Uyghur and Han populations in Xinjiang. *Chin J Epidemiol*. 2013;34(12):1164–1168.
8. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation*. 2009;120(16):1640–1645.
9. Investigators WMPP. The World Health Organization MONICA Project (monitoring trends and determinants in cardiovascular disease): a major international collaboration. WHO MONICA Project Principal Investigators. *J Clin Epidemiol*. 1988;41(2):105–114. doi:10.1016/0895-4356(88)90084-4
10. Liu W, Hu B, Dehghan M, et al. Fruit, vegetable, and legume intake and the risk of all-cause, cardiovascular, and cancer mortality: a prospective study. *Clin Nutr*. 2021;40(6):4316–4323. doi:10.1016/j.clnu.2021.01.016
11. Lindbohm J, Sipilä P, Mars N, et al. Association between change in cardiovascular risk scores and future cardiovascular disease: analyses of data from the Whitehall II longitudinal, prospective cohort study. *Lancet Digit Health*. 2021;3(7):e434–e444.
12. Li R, Li W, Lun Z, et al. Prevalence of metabolic syndrome in Mainland China: a meta-analysis of published studies. *BMC Public Health*. 2016;16:296. doi:10.1186/s12889-016-2870-y
13. do Vale Moreira NC, Hussain A, Bhowmik B, Mdala I, Meyer HE. Prevalence of Metabolic Syndrome by different definitions, and its association with type 2 diabetes, pre-diabetes, and cardiovascular disease risk in Brazil. *Diabetes Metab Syndr Clin Res Rev*. 2020;14(5):1217–1224.
14. Yang X, Li J, Hu D, et al. Predicting the 10-year risks of atherosclerotic cardiovascular disease in Chinese Population: the China-PAR Project (Prediction for ASCVD Risk in China). *Circulation*. 2016;134(19):1430–1440. doi:10.1161/CIRCULATIONAHA.116.022367
15. Sun Z, Zhang Q, Xu S, et al. A study on four different risk assessment models to assess the 10-year cardiovascular disease risk of people over 40 years old in Guiyang. *Guizhou Med J*. 2018;42(8):1007–1010.
16. Hajian-Tilaki K. Factor analysis of metabolic syndrome components in an Iranian non-diabetic adult population: a population-based study from the north of Iran. *Int J Endocrinol Metab*. 2018;16(2):e14159. doi:10.5812/ijem.14159
17. Lai XF, Fang Q, Yang LL, et al. Direct, indirect and total bilirubin and risk of incident coronary heart disease in the Dongfeng-Tongji cohort. *Ann Med*. 2018;50(1):16–25. doi:10.1080/07853890.2017.1377846
18. Doganer YC, Rohrer JE, Aydogan U, Agerter DC, Cayci T, Barcin C. Atherosclerosis and liver function tests in coronary angiography patients. *West Indian Med J*. 2015;64(4):333–337.
19. Lee S, Park W, Suh Y, et al. Association of serum uric acid with cardiovascular disease risk scores in Koreans. *Int J Environ Res Public Health*. 2019;16(23):4632.
20. Zhang X, Li Y, Zhang X. Diagnostic value of combined detection of myocardial enzymes, Hcy and PCT in coronary heart disease. *Chin Foreign Med Res*. 2021;19(10):87–89.
21. Zhou J, Yao H, Wang H, et al. Epidemic features of dyslipidemia among Uyghur, Kazakh and Han ethnic groups in Urumqi. *Natl Med J China*. 2016;38(96):3090–3093.
22. Chen X, Barywani SB, Hansson PO, Rosengren A, Fu MJM. High-normal blood pressure conferred higher risk of cardiovascular disease in a random population sample of 50-year-old men: a 21-year follow-up. *Medicine*. 2020;99(17):e19895.
23. Tian G. *Prospective Cohort Study of the Relationship between Blood Pressure and Fasting Blood Glucose with All-Cause and Cardiovascular Disease Mortality* [dissertation]. Zhengzhou University; 2020.
24. Guo S, Mao L, Liao P, et al. Using metabolic related factors constructing a predictive model for the risk of cardiovascular diseases in Xinjiang Kazakh population. *Chin J Endocrinol Metab*. 2020;01:51–57.
25. Jiang Y, Ma R, Guo H, et al. External validation of three atherosclerotic cardiovascular disease risk equations in rural areas of Xinjiang, China. *Bmc Public Health*. 2020;20(1). doi:10.1186/s12889-020-09579-4

## International Journal of General Medicine

### Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>

across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress