Endocrine (2019) 63:52–61
https://doi.org/10.1007/s12020-018-1718-x

ORIGINAL ARTICLE

Analysis of risk factors of metabolic syndrome using a structural equation model: a cohort study

Zhimin Ma[1,2] · Ditian Li[3] · Siyan Zhan[4] · Feng Sun[4] · Chaonan Xu[1,2] · Yunfeng Wang[1,2] · Xinghua Yang[1,2]

Received: 2 February 2018 / Accepted: 10 August 2018 / Published online: 21 August 2018

Abstract

Purpose We aimed to use a structural equation model (SEM) to determine the interrelations between various risk factors, including latent variables, involved in the development of metabolic syndrome(MetS).

Methods This study used data derived from the MJ Longitudinal Health Check-up Population Database for participants aged 20 to 70 years, who were asymptomatic for MetS at enrollment and were followed up for 5 years. A SEM was applied to investigate the attributions of MetS and the interrelations between different risk factors.

Results Socioeconomic status (SES), living habits, components of metabolic syndrome (COMetS), and blood pressure had a diverse impact on the onset of MetS, directly and (or) indirectly. When investigating the latent risk factors and the interrelations between different risk factors. The standardized total effect (the sum of the direct and indirect effects, β_t) of SES, living habits, blood pressure and COMetS on the onset of MetS was 0.084, −0.179, 0.154, and 0.353, respectively. SES, as a distal risk factor, directly influenced living habits, blood pressure, and COMetS with standardized regression coefficients (β_r) of −0.079 ($P < 0.001$), 0.200 ($P < 0.001$), and −0.163 ($P < 0.001$) respectively. Unfavorable living habits exerted an inverse effect on blood pressure and COMetS ($\beta_r = -0.101$, $P < 0.001$; $\beta_r = -0.463$, $P < 0.001$), which was an important path way for developing MetS.

Conclusions These results demonstrate that individuals with a higher level of SES are susceptible to high blood pressure and are at increased risk for MetS. Additionally, there is a decrease in exercise and an increase in smoking and consumption of alcohol corresponded to an increase in metabolic risk factors.

Keywords Metabolic syndrome · Risk factors · Structural equation model · Socioeconomic status

Introduction

Metabolic syndrome (MetS), a common disease worldwide, is a cluster of inter-related risk factors for cardiovascular disease and type 2 diabetes mellitus [1, 2]. With the increasing prevalence of obesity and sedentary lifestyle, MetS has become a main public health issue globally, with an estimated prevalence of 20–25% [3]. However, beyond lifestyle interventions, available treatments address essentially a single risk factor and an unmet medical need persists [4]. Additionally, The recognized risk factors of MetS include central obesity, raised blood pressure, dyslipidemia, and elevated plasma glucose [5]. Nevertheless, the interrelations underlying these factors have remained uncertain.

When conducting an etiological analysis, pathogenic risk factors are often divided into distal and proximal risk factors. Distal risk factors include factors such as the sociocultural environment, socioeconomic status (SES), and environmental factors. Proximal risk factors can be

These authors contributed equally to this article and should be regarded as first joint author:Zhimin Ma and Ditian Li.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12020-018-1718-x) contains supplementary material, which is available to authorized users.

✉ Xinghua Yang
xinghuay@sina.com

1 School of Public Health, Capital Medical University, Beijing, China

2 Beijing Municipal Key Laboratory of Clinical Epidemiology, Beijing, China

3 Mailman School of Public Health, Columbia University, New York, NY, USA

4 Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China

classified into two types: uncontrollable factors, such as age, sex, and heredity, as well as controllable factors including living habits and dietary behavior. The action time and function of distal and proximal factors differ during the pathogenesis of chronic disease, like MetS, since they are linked to different processes involved in the emergence and development of disease. As endocrine and metabolic diseases are induced by an aggregation of risk factors [6], it is preferable to study a number of variables related to latent factors (a set of variables that have a strong relationship), rather than the effect of a single variable. This is accomplished by studying observed variables and analysing latent variables, which cannot be directly observed. Structural equation model (SEM) is a useful model, which is well-suited to investigate latent variables and causal relationships [7].

Additionally, previous studies have tested the construct validity of MetS, and found latent variables including four latent risk factors for coronary heart disease: high blood pressure (systolic and diastolic blood pressure), obesity (body mass index (BMI) and waist-to-hip ratio), insulin resistance (fasting insulin, high fasting blood glucose), and raised blood lipids (triglycerides (TG), and low levels of high-density lipoprotein (HDL-C)) [8–13]. General logistic regression models do not consider the interaction between these factors, which are considered parallel [14]. However, complex SEMs can be used to analyse attributions, which are associated and interrelate, such as those involved in the development of disease [15, 16]. In this study, we aimed to use an SEM to determine the interrelations between various risk factors, including latent variables, involved in the development of MetS to provide a foundation for complex strategies and measures to prevent and treat the condition.

Materials and methods

Study subjects and measurements

We selected all participants from the 1998 cohort of the MJ Longitudinal Health Check-up Population Database who were aged between 20 and 74 years and did not have a diagnosis of MetS at baseline. Details of the screening program have been described previously [17–20]. Participants were followed up until 2002 (five years from baseline) at which time incidence of MetS was calculated. Exclusion criteria included being lost to follow-up and missing information relating to primary variables. Data of the remaining 15,452 participants were included in the study.

Outcomes and variables of interest

In the present study, we used information from the National Cholesterol Educational Program (the revised NCE-P-ATPIII for Asians in 2004) to define MetS [21], which requires the presence of three or more of the following five MetS components: (1) insulin resistance (fasting plasma glucose (FPG) ≥ 100 mg/dL or taking hypoglycaemic drugs); (2) high blood pressure: (SBP ≥ 130 mmHg or DBP ≥ 85 mmHg or taking antihypertensive drugs); (3) hypertriglyceridemia (TG ≥ 150 mg/dL or taking statins); (4) low (HDL-C < 40 mg/dL in men and < 50 mg/dL in women or taking statins); (5) abdominal obesity (according to the Asian crowd standard, waist circumference (WC ≥ 90 cm in men and ≥ 80 cm in women)).

The MJ Health Personal Questionnaire was a self-administered questionnaire that gathered data of demographic characteristics, living habits, and physical condition. Professional variables were categorized into six classes by ranking from low to high professional reputation. Research variables and assignments are available in Supplementary Table 1. We conducted data cleaning for estimated analysis variables. Missing values were imputed using full information maximum likelihood (FIML) within SEM.

Statistical analysis

First, we conducted an exploratory factor analysis (EFA) to identify the latent variables underlying the observation variables [11]. Latent factors were extracted by principal component analysis and orthogonally underwent varimax rotation [12]. The number of the extracted factors was chosen based on the factor eigenvalue (>1) and screeplot [22]. Second, we performed a confirmatory factor analysis (CFA) to assess whether extracted latent factors indicated observation variables and we constructed an initial SEM. Then we used path graph, which converted a hypothesis into a mathematical model by collecting survey data to verify the hypothesis. This allowed direct relationships to be observed between the variables, including latent and residual variables, and also provided a visual representation of indirect relationships between these variables. Finally, mediation analysis was performed to test these relationships and calculate confidence intervals of standardized direct, indirect, and total effects by using a bootstrap procedure with 5000 iteration [14]. We used Amos (version 22.0), SPSS (version 22.0) and SAS (version 9.4) for the statistical analysis. All P values were two-sided, and statistical significance was defined as $P < 0.05$.

Table 1 Characteristics of study participants at baseline

	Female	Males	*P* value
Population, *n* (%)	8294 (53.7)	7158 (46.3)	
Age (years)	38.09 ± 10.90	37.79 ± 10.59	**<0.001**
Educational level, *n* (%)			
Illiteracy	226 (2.7)	39 (0.5)	**<0.001**
Elementary school	911 (11.0)	427 (6.0)	
Junior high school	477 (5.8)	331 (4.6)	
Senior high school	2060 (24.8)	1339 (18.7)	
Junior college	2226 (26.8)	2018 (28.2)	
University	1964 (23.7)	2048 (28.6)	
Graduate school and above	430 (5.2)	956 (13.4)	
Profession, *n* (%)			
Peasants	65 (0.8)	192 (2.7)	**<0.001**
Households	2204 (26.6)	42 (0.6)	
Laborers	703 (8.5)	2084 (29.1)	
Businessmen	3352 (40.4)	3173 (44.3)	
Civil servants and free lancers	1115 (13.4)	977 (13.6)	
Teachers, soldiers, and students	855 (10.3)	690 (9.6)	
Income (RMB), *n* (%)			
No-income	1714 (20.7)	306 (4.3)	**<0.001**
<400,000	2123 (25.6)	911 (12.7)	
400,000–800,000	3204 (38.6)	3176 (44.4)	
8,000,000–1,200,000	928 (11.2)	1850 (25.8)	
1,200,000–1,600,000	161 (1.9)	432 (6.0)	
1,600,000–2,000,000	78 (0.9)	210 (2.9)	
>2,000,000	86 (1.0)	273 (3.8)	
Smoking (cigarettes), *n* (%)			
No	7790 (93.9)	4366 (61)	**<0.001**
<5	259 (3.1)	715 (10.0)	
5–10	126 (1.5)	560 (7.8)	
10–20	98 (1.2)	1080 (15.1)	
20–40	21 (0.3)	437 (6.1)	
Drinking, *n* (%)			
No	7801 (94.1)	5171 (72.2)	**<0.001**
Used o drink	110 (1.3)	312 (4.4)	
1–2 drinks/week	263 (3.2)	1143 (16.0)	
3–4 drinks/week	68 (0.8)	357 (5.0)	
Every day	52 (0.6)	175 (2.4)	
Exercise, *n* (%)			
No or < 1 h/week	4561 (55)	3063 (42.8)	**<0.001**
1–2 h/week	2059 (24.8)	2000 (27.9)	
3–4 h/week	893 (10.8)	1098 (15.3)	
>5 h/week	781 (9.4)	997 (13.9)	
WC (cm)	69.24 ± 6.70	79.62 ± 7.37	**<0.001**
HDL-C (mg/dL)			**<0.001**
	59.95 ± 14.22	48.996 ± 12.19	
TG (mg/dL)	83.56 ± 42.04	112.74 ± 65.28	**<0.001**
UA (mg/dL)	5.09 ± 1.17	6.89 ± 1.38	**<0.001**
BMI (kg/m^2)	21.35 ± 2.89	23.06 ± 2.71	**<0.001**
SBP (mmHg)	112.42 ± 15.35	118.86 ± 13.89	**<0.001**
DBP (mmHg)	67.92 ± 9.87	72.78 ± 9.73	**<0.001**
FPG (mg/dL)	92.87 ± 11.08	96.35 ± 14.27	**<0.001**
Incidence of MetS after five years, *n* (%)	468 (5.6)	869 (12.1)	**<0.001**

Bold values indicate statistically significant results from the analyses

n number of subjects, *WC* waist circumference, *HDL-C* high-density lipoprotein cholesterol, *TG* triglyceride, *UA* uric acid, *BMI* body mass index, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *FPG* fasting plasma glucose, *MetS* metabolic syndrome

Results

Characteristics of study participants

The general characteristics of the study participants are shown in Table 1. Of the participants, 53.7% were women. The incidence of MetS in women and men after five years was 5.6% and 12.1%, respectively.

Extraction results of latent analysis variables

EFA of the study sample was conducted and the Kaiser-Meyer-Olk statistic was determined to be 0.717, which indicated that the data were suitable for factor analysis. This revealed four sets of latent factors plus exercise with eigenvalue greater than 1. These four factors included SES, living habits, components of metabolic syndrome (COMetS), and blood pressure. The extraction results of the latent variables are shown in Table 2 and screeplot is illustrated in Supplementary Fig. 1.

Then exercise was regarded as a component of living habits to simplify indicators. Finally, we were able to extract four main sets of factors: SES, living habit factors, COMetS, and blood pressure. CFA model was established to investigate the correlation between these factors and resulted in a relatively good fit with the data: chi-square value (x^2) = 4908.341, the ratio of x^2 to the degrees of freedom = 58.433, root mean square error of approximation (RMSEA) = 0.058, comparative fit index (CFI) = 0.782, tuker-lewis index (TLI) = 0.714. Table 3 illustrates that

Table 2 Extraction results of latent analysis variables (Varimax orthogonal rotation)[a]

Latent variables	Indicator variables	Loading coefficient	Cronbach's coefficient
SES	Education	0.815	0.631
	Profession	0.735	
	Income	0.702	
Living habit	Smoking	0.792	0.504
	Drinking	0.822	
COMetS	WC	0.790	0.018
	TG	0.624	
	UA	0.579	
	BMI	0.755	
	HDL-C	−0.658	
	Disease	0.524	
	FPG	0.108	
Blood pressure	SBP	0.901	0.780
	DBP	0.893	
Exercise	Exercise	−0.802	

[a]The result of exploratory factor analysis: Kaiser-Meyer-Olk in measure of sampling adequacy (KMO) = 0.717, Bartlett's test of sphericity approx: Chi-square = 55737.916, $P < 0.001$)

SES socioeconomic status, *COMetS* components of metabolic syndrome, *WC* waist circumference, *TG* triglyceride, *UA* uric acid, *BMI* body mass index, *HDL-C* high-density lipoprotein cholesterol, *Disease* incidence of metabolic syndrome after five years, *FPG* fasting plasma glucose, *SBP* systolic blood pressure, *DBP* diastolic blood pressure

Table 3 Standardized factor loading of the confirmatory factor analysis

Latent variables	Indicator variables	Parameter estimation
SES	Education	0.878
	Profession	0.508
	Income	0.465
Living habit	Smoking	0.626
	Drinking	0.588
	Exercise	0.092
COMetS	WC	0.968
	TG	0.390
	UA	0.506
	BMI	0.821
	HDL-C	−0.405
	Disease	0.321
	FPG	0.162
Blood pressure	SBP	0.886
	DBP	0.779

SES socioeconomic status, *COMetS* components of metabolic syndrome, *WC* waist circumference, *TG* triglyceride, *UA* uric acid, *BMI* body mass index, *HDL-C* high-density lipoprotein cholesterol, *Disease* incidence of metabolic syndrome after five years, *FPG* fasting plasma glucose, *SBP* systolic blood pressure, *DBP* diastolic blood pressure

SES possessed the highest loading for education (0.878). Out of all the COMetS, the standardized factor loading of the WC (0.968) and BMI (0.821) was relatively high. Additionally, HDL-C was negatively loaded, whereas others positively on COMetS.

Construction of path diagram of the SEM

Initial SEM was constructed based on the results of EFA and CFA to measure the interrelation between risk factors and explore how these affect the development of MetS. Education, occupation, and income were indicators of SES, which influenced individual living habits. The individual living habits that were measured included smoking, drinking, and exercise. We designated the COMetS as dependent variables when investigating the influence of SES and living habits.

Model modification and final model

The final model was constructed after correlation coefficients of WC and BMI, HDL and TG were added as free parameters, based on the initial SEM (model 1).The fit indices improved significantly compared with the model 1, with the standardized residuals of the final model more concentrated in [−2, 2] intervals. Of the four different models, model 3 was determined to be the most suitable model and identified as the final model, since there were irrational associations between income and SES in model 4 (shown in Table 4). Figure 1 illustrates that standardized

Table 4 Comparison of the MetS structural equation model fit indices

Model	Content	*Df*	x^2	RMSEA	CFI	TLI
Model 1	Initial model	81	6843.152	0.074	0.878	0.842
Model 2	Add BP- > COMetS[a]	82	12477.748	0.099	0.777	0.715
Model 3	Add WC-BMI and HDL-TG[b]	79	5167.460	0.065	0.909	0.878
Model 4	Add education level[c] -income[c]	80	6839.207	0.074	0.879	0.841

MetS metabolic syndrome, *Df* degrees of freedom, x^2 Chi-square values, *RMSEA* root mean squared error of approximation, *CFI* comparative fit index, *TLI* tuker-lewis index, *BP* blood pressure, *COMetS* components of metabolic syndrome *WC* waist circumference, *BMI* body mass index, *HDL* high-density lipoprotein cholesterol, *TG* triglyceride

[a]Unlike the model 1, the model 2 considered the interrelation between blood pressure and COMetS

[b]In the model 1, it was assumed that WC and BMI were not directly related, but they might be correlated in the model 3. Meanwhile, the relationship between HDL and TG was also indirectly, so they were correlated in the model 3

[c]In the model 1, it was assumed that education and income were not directly related, but they might be correlated in the model 4

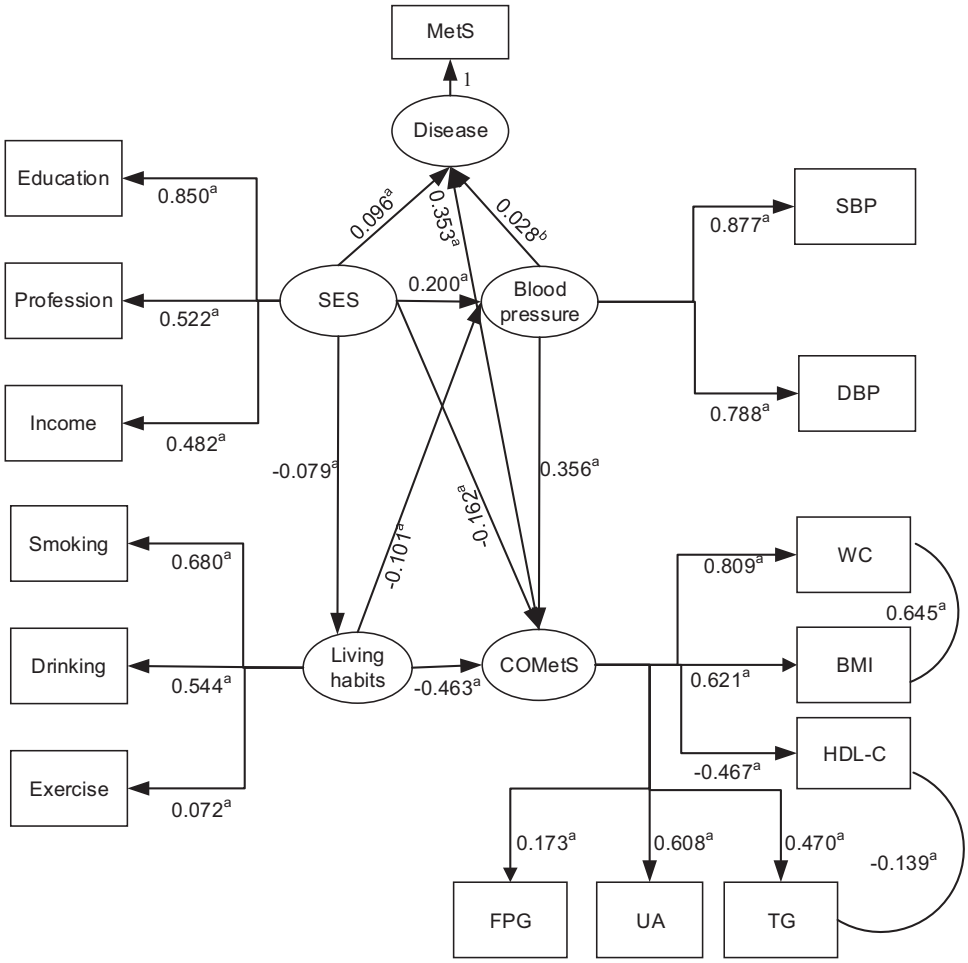

Fig. 1 Structural Equation Modeling for the Associations between SES, living habits, blood pressure, COMetS and MetS [a] *P* values < 0.001, [b] *P* values < 0.05, Oval shapes represent latent factors, and rectangular ones connote directly measured variable, *SES* socioeconomic status, *MetS* metabolic syndrome, *Disease* incidence of metabolic syndrome after fiveyears, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *WC* waist circumference, *BMI* body mass index, *HDL-C* high-density lipoprotein cholesterol, *TG* triglyceride, *UA* uric acid, *FPG* fasting plasma glucose, *COMetS* components of metabolic syndrome

coefficient (β_r) of SES on living habits, blood pressure factors, and COMetS was −0.079 ($P < 0.001$), 0.200 ($P < 0.001$), and −0.163 ($P < 0.001$), respectively. Living habits were inversely associated with blood pressure and COMetS ($\beta_r = -0.101$, $P < 0.001$; $\beta_r = -0.463$, $P < 0.001$), which was an important pathway leading to MetS. Table 5 shows that the standardized total effect (the sum of the direct and indirect effects, β_t) of SES, living habits, blood pressure, and COMetS on the onset of MetS was 0.084 (95% confidence interval (CI) 0.063, 0.104), −0.179 (95% CI −0.192, −0.166), 0.154 (95% CI 0.135, 0.174) and 0.353 (95% CI 0.333, 0.375), respectively. Besides, there were still significant indirect effects and direct effects after mediation testing (shown in Table 5).

Discussion

This study has demonstrated that individuals with a higher level of SES were susceptible to high blood pressure and at increased risk for MetS by using SEM model. With development of the medical research field, it has become important to study and analyse the relationship between social and behavioral factors that affect people's health. This is resulting in a gradual increase in use of SEMs in medicine [23–31]. In previous studies, CFA has been used to study the construct validity of MetS. Those results state that MetS is a latent variable and includes four potential risk factors for coronary heart disease: hypertension (e.g., SBP and DBP), obesity (e.g., BMI and WC), insulin resistance (e.g., FINS and FPG), and hyperlipidaemia (e.g., TG and HDL-C) [8–13]. Therefore, we constructed the latent variable path graph and the SEM model based on past theories. Compared with those previous works, the present study analysed the influence of different factors on the onset of MetS, which focused on the direct and indirect effects.

Additionally, previous studies on model fitting evaluation index have reported that the standard Chi-square value is applicable for $N < 200$. When $N > 200$, Chi-square tests are not suitable [32]. In our study, our sample included more than 200 participants; hence, the Chi-square value was not suitable for model fitting for our sample, although other fitting tests unrestricted by the sample size are relatively good. The fitting model results indicated that the main fit

Table 5 The effect of of MetS structural equation model[a]

Association	Standardized direct effect	Standardized indirect effect	Standardized total effect [b]
Associations between latent variables			
SES → disease	0.096 (0.076, 0.117)	−0.012 (−0.024, −0.001)	0.084 (0.063, 0.104)
Living habits → disease	NA	NA	−0.179 (−0.192, −0.166)
Blood pressure → disease	0.356 (0.336, 0.377)	NA	0.356 (0.336, 0.377)
COMetS → disease	0.353 (0.333, 0.375)	NA	0.353 (0.333, 0.375)
SES → blood pressure	0.200 (0.175, 0.222)	0.008 (0.004, 0.012)	0.208 (0.183, 0.230)
Living habits → blood pressure	−0.101 (−0.132, −0.071)	NA	−0.101 (−0.132, −0.071)
SES → living habits	−0.079 (−0.111, −0.048)	NA	−0.079 (−0.111, −0.048)
SES → COMetS	−0.162 (−0.185, −0.139)	0.111 (0.093, 0.128)	−0.052 (−0.080, −0.025)
Blood pressure → COMetS	0.356 (0.336, 0.377)	NA	0.356 (0.336, 0.377)
Living habits → COMetS	−0.463(−0.488, −0.438)	−0.036 (−0.046, −0.026)	−0.499 (−0.526, −0.472)
Associations between latent variable with its indicator variables			
SES → education	0.850 (0.820, 0.883)	NA	0.850 (0.820, 0.883)
SES → profession	0.522 (0.500, 0.541)	NA	0.522 (0.500, 0.541)
SES → income	0.482 (0.458, 0.505)	NA	0.482 (0.458, 0.505)
Blood pressure → SBP	0.877 (0.857, 0.897)	NA	0.877 (0.857, 0.897)
Blood pressure → DBP	0.788 (0.769, 0.806)	NA	0.788 (0.769, 0.806)
Living habits → smoking	0.680 (0.647, 0.716)	NA	0.680 (0.647, 0.716)
Living habits → drinking	0.544 (0.516, 0.572)	NA	0.544 (0.516, 0.572)
Living habits → exercise	0.072 (0.044, 0.098)	NA	0.072 (0.044, 0.098)
COMetS → WC	0.809 (0.796, 0.822)	NA	0.809 (0.796, 0.822)
COMetS → BMI	0.621 (0.605, 0.638)	NA	0.621 (0.605, 0.638)
COMetS → HDL-C	−0.467 (−0.483, −0.452)	NA	−0.467 (−0.483, −0.452)
COMetS → TG	0.470 (0.451, 0.490)	NA	0.470 (0.451, 0.490)
COMetS → UA	0.608 (0.594, 0.622)	NA	0.608 (0.594, 0.622)
COMetS → FPG	0.173 (0.152, 0.195)	NA	0.173 (0.152, 0.195)
Associations between latent variable with other variables			
SES → SBP	NA	0.182 (0.159, 0.204)	0.182 (0.159, 0.204)
SES → DBP	NA	0.164 (0.146, 0.180)	0.164 (0.146, 0.180)
SES → smoking	NA	−0.054 (−0.075, −0.033)	−0.054 (−0.075, −0.033)
SES → drinking	NA	−0.043 (−0.062, −0.026)	−0.043 (−0.062, −0.026)
SES → exercise	NA	−0.006 (−0.009, −0.003)	−0.006 (−0.009, −0.003)
SES → WC	NA	−0.042 (−0.065, −0.020)	−0.042 (−0.065, −0.020)
SES → BMI	NA	−0.032 (−0.049, −0.015)	−0.032 (−0.049, −0.015)
SES → HDL-C	NA	0.024 (0.011, 0.038)	0.024 (0.011, 0.038)
SES → TG	NA		

Table 5 (continued)

Association	Standardized direct effect	Standardized indirect effect	Standardized total effect [b]
		−0.024 (−0.038, −0.012)	−0.024 (−0.038, −0.012)
SES → UA	NA	−0.031 (−0.049, −0.015)	−0.031 (−0.049, −0.015)
SES → FPG	NA	−0.009 (−0.014, −0.004)	−0.009 (−0.014, −0.004)
Living habits → SBP	NA	−0.089 (−0.115, −0.062)	−0.089 (−0.115, −0.062)
Living habits → DBP	NA	−0.080 (−0.105, −0.055)	−0.080 (−0.105, −0.055)
Living habits → WC	NA	−0.404 (−0.426, −0.381)	−0.404 (−0.426, −0.381)
Living habits → BMI	NA	−0.310 (−0.328, −0.293)	−0.310 (−0.328, −0.293)
Living habits → HDL-C	NA	0.233 (0.219, 0.247)	0.233 (0.219, 0.247)
Living habits → TG	NA	−0.235 (−0.251, −0.219)	−0.235 (−0.251, −0.219)
Living habits → UA	NA	−0.304 (−0.321, −0.286)	−0.304 (−0.321, −0.286)
Living habits → FPG	NA	−0.086 (−0.098, −0.075)	−0.086 (−0.098, −0.075)
Blood pressure → WC	NA	0.288 (0.272, 0.306)	0.288 (0.272, 0.306)
Blood pressure → BMI	NA	0.221 (0.207, 0.235)	0.221 (0.207, 0.235)
Blood pressure → HDL-C	NA	−0.166 (−0.177, −0.156)	−0.166 (−0.177, −0.156)
Blood pressure → TG	NA	0.168 (0.156, 0.179)	0.168 (0.156, 0.179)
Blood pressure → UA	NA	0.217 (0.203, 0.230)	0.217 (0.203, 0.230)
Blood pressure → FPG	NA	0.062 (0.053, 0.071)	0.062 (0.053, 0.071)

[a]The 95% confidence intervals within the parentheses

[b]Standardized total effect is the sum of standardized direct and indirect effects

SES socioeconomic status, *COMetS* components of metabolic syndrome, *NA* no analysis, *Disease* incidence of metabolic syndrome after five years, *WC* waist circumference, *BMI* body mass index, *HDL-C* high-density lipoprotein cholesterol, *FPG* fasting plasma glucose, *UA* uric acid, *TG* triglyceride, *SBP* systolic blood pressure, *DBP* diastolic blood pressure

index CFI and TLI were more than 0.800, with a RMSEA value of 0.065, suggesting a relative good fit.

Individuals with high level of SES should pay more attention to prevent the occurrence of MetS, since SES, as a distal risk factor, directly influenced living habits, blood pressure, and COMetS through this pathway and indirectly prevented metabolic diseases. These results suggested that the influence of individual living habits and COMetS negatively correlated with increasing SES. However, the correlation between BP and MetS was positive. Overall, these results indicated that participants increasingly suffered from MetS who had a higher-level SES, which differed from the results of previous studies that reported an inverse association between SES and developing MetS [33–35]. This might be explained by the general logistic regression model that was used by those previous works, disregarding the interrelation between risk factors and the latent factors.

Unfavorable living habits were inversely associated with blood pressure and COMetS, which was an important pathway leading to MetS. Our results also illustrated that factors associated with living habits had an important influence on COMetS. We found that less exercise, more smoking, and increased consumption of alcohol corresponded with an increase in metabolic risk factors.

When investigating the latent risk factors and the interrelations between different risk factors, SES, blood pressure, and COMetS positively correlated with the onset of MetS. However, unfavorable living habits correlated negatively with the development of MetS. Specifically, HDL-C on COMetS was measured by concentration (mg/dL) means reverse coded was inappropriate, and negative coefficient suggested that HDL-C worked against the COMetS latent variable, which coincided with previous research [11, 36, 37]. Overall, blood pressure and COMetS had a strong

correlation with the onset of MetS. Additionally, SES and living habits, as distal and proximal factors respectively, had a different effect on the development of MetS.

Our study reveals several advantages of SEM application for MetS compared with the traditional analysis methods. First, when studying the pathogenesis of MetS, a more thorough analysis is possible if we allocate latent variables for different stages of development of the disease. When studying determinants of health, it is easy to generate conceptual bias using single-measurement social indicators, particularly with SES, for which there are multiple indices. Second, an SEM can be used to construct a network of latent variables for each of the multiple indicators for pathogenesis and to study the relationship between these variables [11]. Third, the multicollinearity of independent variables can be investigated after introducing latent variables. Fourth, the multicollinearity of independent variables can be investigated using latent variables, as can the measurement error, to improve the accuracy of parameter estimation [38]. Last, the SEM is a robust model because it uses a multiple sequence correlation coefficient matrix to fit the data.

However, there are certain problems that need to be discussed. For example, unless the model is specifically designed to explore the causal effect between variables, it cannot be proved that there is a causal relationship [39]. In 1990, Marsh [40] suggested that when using longitudinal data there should be at least two time measurement points, that multiple indicators should be used to calculate latent variables, and that when the sample size was large enough, the significance of different models and the meaning of index error correlation should be considered. We used longitudinal data in our fitting, so we can confirm that there were some causal relationships. Another limitation in our research is that there are some disadvantages when interpreting the results. The introduction of latent variables is the major advantage of the SEM, which can solve a problem of measuring of a single variable. However, whether the measure really reflects the latent variables, or the observed variables that correspond to the latent variables, is still a matter of debate [41, 42]. In addition, although CFA was conducted based on theories in this study, the low standardized estimates in some of the standardized factor loadings suggested that perhaps there could be a better specification of the measurement model for the SEM.

In conclusion, the present study demonstrated that the onset of MetS was related to SES, living habits, COMetS, and blood pressure factors, which affected the development of MetS directly and(or) indirectly. Additionally, SES was found to have a negative correlation with living habits and COMetS but a positive association with blood pressure through this pathway, indirectly influencing the onset of MetS. Furthermore, unfavorable living habits were found to be inversely associated with blood pressure and COMetS, which was a crucial pathway leading to MetS. These results raised awareness of the relationship between SES and MetS, and suggested the need for a comprehensive management of MetS and its treatment.

Acknowledgements All authors thank MJ Health Management Institution for making their large dataset available to us. We thanks all people participating in the MJ Longitudinal Health study.

Compliance with ethical standards

Ethical approval The Peking University Institutional Review Board approved this study. Since it eliminated all identifiable personal information, it does not belong to studies involving human beings. Consequently, this Board waived the requirement for informed consent and ethical review of the study. (Authorization Code: MJHRFB2014003C).

Informed consent The requirement for informed consent was waived for this study.

Conflict of interest The authors declare that no competing interests exist.

References

1. K.G. Alberti, R.H. Eckel, S.M. Grundy et al.. Harmonizing the metabolic syndrome: A joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Preventiony. Circulation **120**, 1640–1645 (2009)
2. J. Kang, Y.M. Song, Metabolic syndrome and its components among Korean submariners: a retrospective cross-sectional study. Endocrine **59**, 614–621 (2018). https://doi.org/10.1007/s12020-017-1518-8
3. M.Z.I. Chowdhury, A.M. Anik, Z. Farhana et al.. Prevalence of metabolic syndrome in Bangladesh: A systematic review and meta-analysis of the studies. BMC. Public. Health **18**, 308 (2018). https://doi.org/10.1186/s12889-018-5209-z
4. D. Junquero, Y. Rival, Metabolic syndrome: Which definition for what treatment(s)? Med. Sci. **21**, 1045–1053 (2005). https://doi.org/10.1051/medsci/200521121045
5. N. Adler, A. Singh-Manoux, J. Schwartz, J. Stewart, K. Matthews, M.G. Marmot, Social status and health: A comparison of British civil servants in Whitehall-II with European-and African-Americans in CARDIA. Soc. Sci. Med. **66**, 1034–1045 (2008). https://doi.org/10.1016/j.socscimed.2007.11.031
6. R. Karns, P. Succop, G. Zhang et al.. Modeling metabolic syndrome through structural equations of metabolic traits, comorbid diseases, and GWAS variants. Obesity **21**, E745–E754 (2013). https://doi.org/10.1002/oby.20445
7. K.A. Bollen, M.D. Noble, Structural equation models and the quantification of behavior. Proc. Natl. Acad. Sci. USA **108**, 15639–15646 (2011). https://doi.org/10.1073/pnas.1010661108
8. M.M. Smits, P. Woudstra, K.M. Utzschneider et al.. Adipocytokines as features of the metabolic syndrome determined using confirmatory factor analysis. Ann. Epidemiol. **23**, 415–421 (2013). https://doi.org/10.1016/j.annepidem.2013.03.001
9. C.M. Stein, Y. Song, R.C. Elston, G. Jun, H.K. Tiwari, S.K. Iyengar, Structural equation based genome scan for the metabolic syndrome. BMC Genet. **4**, S99 (2003). https://doi.org/10.1186/1471-2156-4-S1-S99

10. J.C. Chan, J.C. Cheung, E.M. Lau, J. Wooà, A.Y. Chan, R. Swaminathan, C.S. Cockrama, The metabolic syndrome in Hong Kong Chinese. The interrelationships among its components analyzed by structural equation modeling. Diabetes Care **19**, 953–959 (1996). https://doi.org/10.2337/diacare.19.9.953
11. C.L. Cheung, K.C. Tan, K.S. Lam, B.M. Cheung, The relationship between glucose metabolism, metabolic syndrome, and bone-specific alkaline phosphatise: A structural equation modelling approach. J. Clin. Endocrinol. Metab. **98**, 3856–3863 (2013). https://doi.org/10.1210/jc.2013-2024
12. J.E. Stevenson, B.R. Wright, A.S. Boydstun, The metabolic syndrome and coronary artery disease: A structural equation modeling approach suggestive of a common underlying patho-physiology. Metabolism **61**, 1582 (2012). https://doi.org/10.1016/j.metabol.2012.04.010
13. S. Novak, L.M. Stapleton, J.R. Litaker, K.A. Lawson, PCV18, a confirmatory factor analysis evaluation of the coronary heart disease risk factors of metabolic syndrome and the effectiveness of the current ATP III guidelines for identification. Value Health **6**, 312–313 (2003). https://doi.org/10.1046/j.1463-1326.2003.00289.x
14. J.E. Given, M.J. O'Kane, V.E. Coates, A. Moore, B.P. Bunting, Comparing patient generated blood glucose diary records with meter memory in type 2 diabetes. Diabetes Res. Clin. Pract. **104**, 358–362 (2014). https://doi.org/10.1016/j.diabres.2014.03.003
15. R. Song, S. Ahn, H. Oh, A structural equation model of quality of life in adults with type 2 diabetes in Korea. Appl. Nurs. Res. **26**, 116–120 (2013). 10.1016/j.apnr. 2013.04.001
16. C. Conti, G.D. Francesco, L. Fontanella et al.. Negative affectivity predicts lower quality of life and metabolic control in type 2 diabetes patients: A structural equation modeling approach. Front. Psychol. **8**, 831 (2017). https://doi.org/10.3389/fpsyg.2017.00831
17. Y.H. Shen, W.S. Yang, T.H. Lee, L.T. Lee, C.Y. Chen, K.C. Huang, Bright liver and alanine aminotransferase are associated with metabolic syndrome in adults. Obes. Res. **13**, 1238–1245 (2005). https://doi.org/10.1038/oby.2005.147
18. P.F. Hsu, S.Y. Chuang, H.M. Cheng, S.T. Tsai, P. Chou, C.H. Chen, Clinical significance of the metabolic syndrome in the absence of established hypertension and diabetes: A community-based study. Diabetes Res. Clin. Pract. **79**, 461–467 (2008). https://doi.org/10.1038/oby.2005.147
19. C.P. Wen, T.Y. Cheng, M.K. Tsai et al.. All-cause mortality attributable to chronic kidney disease: A prospective cohort study based on 462 293 adults in Taiwan. Lancet **371**, 2173 (2008). https://doi.org/10.1016/S0140-6736(08)60952-6
20. C.P. Wen, P. Wai, T. Minkuang et al.. Minimum amount of physical activity for reduced mortality and extended life expectancy: A prospective cohort study. Lancet **378**, 1244–1253 (2011). https://doi.org/10.1016/S0140-6736(11)60749-6
21. S.M. Grundy, J.I. Cleeman, S.R. Daniels et al.. Diagnosis and management of the metabolic syndrome: An AHA/NHLBI Scientific Statement. Curr. Opin. Cardiol. **21**, 1–6 (2006)
22. V. Edefonti, F. Bravi, W. Garavello, et al., Nutrient-based dietary patterns and laryngeal cancer: Evidence from an exploratory factor analysis. Cancer Epidemiol. Biomarkers Prev. 19, (2010). https://doi.org/10.1158/1055-9965
23. L.A. Hayduk, Shame for disrespecting evidence: The personal consequences of insufficient respect for structural equation model testing. BMC. Med. Res. Methodol. **14**, 124 (2014). https://doi.org/10.1186/1471-2288-14-124
24. K.R. Conner, D. Gunzler, W. Tang, X.M. Tu, S.A. Maisto, Test of a clinical model of drinkingand suicidal risk. Alcohol. Clin. Exp. Res. **35**, 60 (2011). https://doi.org/10.1111/j.1530-0277.2010.01322.x
25. C.Y. Huang, C.W. Lu, Y.L. Liu, C.H. Chiang, L.T. Lee, K.C. Huang, Relationship between chr-onic hepatitis B and metabolic syndrome: A structural equation modeling approach. Obesity **24**, 483 (2016). https://doi.org/10.1002/oby.21333
26. E. Fulu, S. Miedema, T. Roselli, et al., Pathways between childhood trauma, intimate partner violence, and harsh parenting: Findings from the UN Multi-country Study on Men and Viole-nce in Asia and the Pacific. Lancet. Health **5**, e512–e522 (2017). https://doi.org/10.1016/S2214-109X(17)30103-1
27. E. Long, S. Xu, Z. Liu et al.. Construction and implications of structural equation modeling network for pediatric cataract: A data mining research of rare diseases. BMC Ophthalmol. **17**, 74 (2017). https://doi.org/10.1186/s12886-017-0468-5
28. S.K. Mama, P.M. Diamond, S.A. Mccurdy, A.E. Evans, L.H. Mcneill, R.E. Lee, Individual, social and environmental correlates of physical activity in overweight and obese African American and Hispanic women: A structural equation model analysis. Prev. Med. Rep. **2**, 57–64 (2015). https://doi.org/10.1016/j.pmedr.2015.01.001
29. F. Ødegaard, P. Roos, Measuring worksite health promotion programs: An application of structural equation modeling with ordinal data. Eur. J. Health Econ. **14**, 639–653 (2013). https://doi.org/10.1007/s10198-012-0409-4
30. M.M. Belvederi, S. Mamberto, L. Briatore, C. Mazzucchelli, M. Amore, R. Cordera, The inter- play between diabetes, depression and affective temperaments: A structural equation model. J. Affect Disord. **219**, 64–71 (2017). https://doi.org/10.1016/j.jad.2017.05.018
31. L. Fisher, D. Hessler, W. Polonsky et al.. Emotion regulation contributes to the development of diabetes distress among adults with type 1 diabetes. Patient Educ. Couns. **101**, 124–131 (2018). https://doi.org/10.1016/j.pec.2017.06.036
32. K.J. Preacher, A.F. Hayes, Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav. Res. Methods **40**, 879–891 (2008). https://doi.org/10.3758/BRM.40.3.879
33. A. Ala'A, S. Nicolas, L. Marie-Lise, A. Adelin, Dietary, behavioural and socio-economic determinants of the metabolic syndrome among adults in Luxembourg: Findings from the ORISCAV-LUX study. Public Health Nutr. **15**, 849–859 (2012). S1368980011002278
34. A. Goyal, D.L. Bhatt, P.G. Steg et al.. Attained educational level and incident atherothrombotic events in low- and middle-income compared with high-income countries. Circulation **122**, 1167–1175 (2010). https://doi.org/10.1161/CIRCULATIONAHA.109.919274
35. M.A. Winkleby, D.E. Jatulis, E. Frank, S.P. Fortmann, Socio-economic status and health: How education, income, and occupation contribute to risk factors for cardiovascular disease. Am. J. Public Health **82**, 816–820 (1992). https://doi.org/10.2105/AJPH.82.6.816
36. W. Lu, K. Song, Y. Wang et al.. Relationship between serum uric acid and metabolic syndrome: An analysis by structural equation modeling. J. Clin. Lipidol. **6**, 159–167 (2012). https://doi.org/10.1016/j.jacl.2011.11.006
37. Lakka Hanna-Maaria, E.Laaksonen David, TimoA. Lakka et al.. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA **288**, 2709–2716 (2002). https://doi.org/10.1001/jama.288.21.2709
38. F.Y. Shi, W.F. Gao, E.X. Tao, H.Q. Liu, S.Z. Wang, Metabolic syndrome is a risk factor for nonalcoholic fatty liver disease: Evidence from a confirmatory factor analysis and structural equation modeling. Eur. Rev. Med. Pharmacol. Sci. **20**, 4313 (2016)
39. M. Santiagotorres, Y. Cui, A.K. Adams, D.B. Allen et al.. Structural equation modeling of the associations between the home environment and obesity-related cardiovascular fitness and

insulin resistance among Hispanic children. Appetite **101**, 23–30 (2016). https://doi.org/10.1016/j.appet.2016.02.003

40. A. Schmitt, A. Reimer, N. Hermanns et al.. Depression is linked to hyperglycaemia via subop-timal diabetes self-management: A cross-sectional mediation analysis. J. Psychosom. Res. **94**, 17 (2017). https://doi.org/10.1016/j.jpsychores.2016
41. C.M. Rebholz, M.E. Grams, Y. Chen et al.. Serum levels of 1,5-anhydroglucitol and risk of incident end-stage renal disease. Am. J. Epidemiol. **186**, 952–960 (2017). https://doi.org/10.1093/aje/kwx167
42. C.I. Mercado, Q. Yang, E.S. Ford, E. Gregg, A.L. Valderrama, Gender-and race-specific metabolic score and cardiovascular disease mortality in adults: A structural equation modelling approach-United States, 1988-2006. Obesity **23**, 1911–1919 (2015). https://doi.org/10.1002/oby.21171