



**OPEN** **Patterns of risk factors for cardiovascular diseases in Kheramah PERSION cohort population using a latent class analysis**

**Hadis Salimi<sup>1</sup>, Mohebat Vali<sup>2</sup>, Abbas Rezaianzadeh<sup>3</sup>, Masoumeh Ghoddusi Johari<sup>4</sup> & Haleh Ghaem<sup>2</sup>✉**

Cardiovascular diseases (CVDs) are a leading cause of mortality and morbidity worldwide, with a significant burden in Iran. Limited research has investigated patterns of modifiable CVDs risk factors in Iran. This study aims to address this gap by identifying distinct patterns of modifiable CVDs risk factors among adult aged + 40 and explore the relationship between demographic characteristics and risk factor patterns. This study was conducted using data from the Kherameh cohort study. The participants consisted of 9,422 individuals aged 40–70 years without CVDs. Latent Class Analysis (LCA) was used to identify latent classes of modifiable CVD risk factors. Multinomial logistic regression assessed the relationship between latent classes (LCs) and demographic variables. Three latent classes were identified as follows: low-risk (42%), clinical-risk (52%) and lifestyle-risk (6%) classes. Female gender (Adjusted OR: 13.48, 95% CI: 11.81–15.39), older age (Adjusted OR: 1.16, 95% CI: 0.99–1.35) and rural residence (Adjusted OR: 0.76, 95% CI: 0.67–0.86) had a greater risk of being in clinical-risk class compared to the low-risk class. Moreover, individuals of Fars ethnicity exhibited a significantly elevated risk of being classified in the clinical risk class for CVD (Adjusted OR: 1.28, 95% CI: 1.14–1.43) and they demonstrated a markedly higher risk of belonging to the lifestyle risk class (Adjusted OR: 1.51, 95% CI: 1.11–2.07). The study identified distinct latent classes of modifiable CVD risk factors and provided insights into their associations with demographic characteristics. Understanding risk patterns is crucial for developing effective preventive strategies and providing appropriate health protocols.

**Keywords** Cardiovascular diseases, Risk factor, Latent class

#### Abbreviations

CVD	Cardiovascular diseases
LCA	Latent class analysis
NCD	Non-communicable diseases
YLL	Years of life lost
DALY	Disability-adjusted life years
IHD	Ischemic heart disease
DM	Diabetes mellitus
PERSIAN	Prospective Epidemiological Research Studies in Iran
PCA	Principal component analysis
BIC	Bayesian information criterion
AIC	Akaike information criterion

<sup>1</sup>Student Research Committee, Shiraz University of Medical Science, Shiraz, Iran. <sup>2</sup>Department of Epidemiology, Non-Communicable Diseases Research Center, Research Center for Health Sciences, Institute of Health, School of Health, Shiraz University of Medical Sciences, Shiraz, Iran. <sup>3</sup>Colorectal Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. <sup>4</sup>Breast Diseases Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. ✉email: ghaemh@sums.ac.ir

Cardiovascular diseases (CVDs) are a major public health concern worldwide, causing a significant portion of deaths and disabilities<sup>1</sup>. Ischemic heart disease (IHD) and stroke are the leading contributors to CVD-related deaths, with developing countries experiencing higher rates<sup>2</sup>. In Iran, CVDs are the leading cause of death and disability, with a prevalence exceeding 9% in some regions<sup>3–5</sup>. The burden of CVDs in Iran is expected to double by 2025 compared to 2005, particularly among individuals over 30<sup>6</sup>.

Risk factors for CVDs can be categorized into non-modifiable and modifiable factors. Non-modifiable factors include age, gender, ethnicity, and family history. Modifiable risk factors can be further divided into three domains: clinical/biological factors (e.g., obesity, diabetes, hypertension, dyslipidemia), psychosocial factors (e.g., anxiety, stress, depression), and behavioral/lifestyle factors (e.g., sleep patterns, alcohol consumption, smoking, diet, physical activity)<sup>7–11</sup>.

Understanding the patterns of CVD risk factors within specific populations is crucial for developing effective preventive interventions. Latent Class Analysis (LCA) is an advanced statistical method that can identify subgroups of individuals with similar risk factor profiles. Although several studies have been carried out concerning risk factors for CVD in Iran, few of these used LCA to assess the patterning of these risk factors. It therefore creates a glaring need for in-depth analysis in order to unravel the hidden patterns and provide a background for efficient preventive measures. Thereby, the obtained patterns in Kherameh will be important not only to this population but also to others having similar characteristics in terms of risks.

This study aims to identify distinct patterns of well-established modifiable CVD risk factors among adults aged 40–70 years in the Kherameh cohort study. These risk factors include diabetes mellitus (DM), hypertension, high blood total cholesterol, abdominal obesity, disturbed sleep, smoking status, and alcohol consumption. The study will further characterize the most prominent risk factors within each identified pattern and explore the relationship between demographic characteristics and these risk factor patterns.

## Methods

### Study design

The present findings were extracted from a research project approved by the Ethics Committee of Shiraz University of Medical Sciences (IR.SUMS.SCHEANUT.REC.1402.052) and informed consent was obtained from all the participants. This cross-sectional, analytical study utilized baseline data from the Kherameh cohort study, a population-based sub-cohort of the Prospective Epidemiological Research Studies in Iran (PERSIAN). The PERSIAN Cohort Study, Launched in 2014, encompasses 18 geographical regions across Iran and includes all major Iranian ethnicities and its rationale, goals, and design have already been published<sup>12</sup>. Kherameh, a city in southern Fars province with a population of 61,580, contributes to this study with 10,663 individuals aged 40–70 years. Due to missing values, 10,611 participants were included in the current analysis<sup>12</sup>.

### Inclusion criteria and exclusion criteria

Inclusion criteria mirrored those of the Kherameh cohort study: participants aged 40–70 years, as behaviors and health outcomes like chronic diseases are well-established in this group. Additionally, participants had to reside in the Kherameh area for at least 9 months prior to enrollment to account for environmental and cultural influences. Exclusion criteria included untreated acute illnesses, diagnosed mental health conditions, unwillingness to participate, and inability to attend clinic examinations. Additionally, individuals with a previous diagnosis of cardiovascular disease (CVD), specifically including ischemic heart disease, myocardial infarction (MI), stroke, and cerebrovascular accident (CVA), to focus on assessing CVD risk in those without prior heart disease.

### CVD risk indicators

Our focus is on specific subtypes of CVD including ischemic heart disease, myocardial infarction (MI), stroke, and cerebrovascular accident (CVA). Our analysis included commonly used CVD risk indicators such as hypertension, high blood total cholesterol, DM, obesity, and smoking, based on models like the Framingham risk score<sup>13,14</sup>. Additional indicators included abnormal sleep and alcohol consumption<sup>8,11</sup>. Hypertension was defined as systolic blood pressure  $\geq 130$  mmHg or diastolic blood pressure  $\geq 85$  mmHg<sup>15–17</sup>. High blood cholesterol was defined as total cholesterol level  $\geq 250$  mg/dL<sup>18</sup>. Abdominal obesity was defined as waist circumference  $\geq 102$  cm in men and  $\geq 88$  cm in women<sup>19</sup>. DM, smoking, and alcohol consumption were coded as binary variables. Disturbed sleep was defined as sleep duration  $\geq 9$  h or  $\leq 5$  h<sup>20,21</sup>.

### Demographic characteristics

Demographic characteristics included age (40–49, 50–59, and 60–70 years), gender, ethnicity (Fars and others), employment status (employed or unemployed), area of residence (urban or rural), education level (illiterate, diploma or lower diploma, academic degree), and socioeconomic status (low, medium, or high). Socioeconomic status was determined using principal component analysis (PCA) based on factors like house ownership, house size, number of bathrooms, car ownership and price, travel frequency, and ownership of household items.

### Statistical methods

To understand how modifiable CVD risk factors cluster, we used latent class analysis (LCA) to identify distinct patterns of these factors<sup>22</sup>. Model evaluation criteria included Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), entropy, and interpretability. Lower BIC and AIC values indicate better model fit<sup>23</sup>, while higher entropy values indicate better classification<sup>24</sup>. Multi-nominal logistic regression was used to examine the relationship between identified risk groups and personal variables. LCA was performed using R software version 4.3.1, and StataMP17 was used for regression analysis.

Variable		Frequency	Percentage
Diabetes mellitus	Yes	1335	14.2
	No	8078	58.8
Hypertension	Yes	1917	20.3
	No	7505	79.7
Abdominal obesity	Yes	5140	54.6
	No	4280	45.4
High cholesterol level	Yes	962	10.3
	No	8452	89.7
Sleep disorders	Yes	1991	21.1
	No	7431	78.9
Smoking	Yes	2392	25.4
	No	7030	74.6
Alcohol consumption	Yes	351	3.7
	No	9071	96.3

**Table 1.** Description of observed variables applied in LCA method.

Number of latent class	Number of parameters estimated	G2	df	AIC	BIC	X2	Maximum log-likelihood	Entropy
2	15	139.7734	112	58,877.49	58,984.73	181.7394	-29,423.74	3.126888
3	23	103.8544	104	58,857.57	59,022.01	119.8267	-29,405.78	3.124326
4	31	93.65534	96	58,863.37	59,085.01	110.3749	-29,400.69	3.124024

**Table 2.** Criteria for LCA model fit.

	Latent class		
	Class 1	Class 2	Class 3
Latent class prevalence	<b>0.5225</b>	<b>0.0587</b>	<b>0.4188</b>
Item-response probabilities	<b>0.5197</b>	<b>0.0245</b>	<b>0.4558</b>
High blood pressure	<b>0.2149</b>	<b>0.1704</b>	<b>0.1941</b>
Total cholesterol	<b>0.1343</b>	<b>0.1331</b>	<b>0.0578</b>
Diabetes	<b>0.2004</b>	<b>0.0790</b>	<b>0.0771</b>
Abdominal obesity	<b>0.7869</b>	<b>0.2061</b>	<b>0.2924</b>
Sleep disorder	<b>0.2137</b>	<b>0.2519</b>	<b>0.2024</b>
Smoking	<b>0.0428</b>	<b>0.6916</b>	<b>0.4557</b>
Alcohol consumption	<b>0.0259</b>	<b>0.4030</b>	<b>0.0000</b>

**Table 3.** Predicted probabilities of CVD risks by LCs from latent class analysis.

## Results

### Description

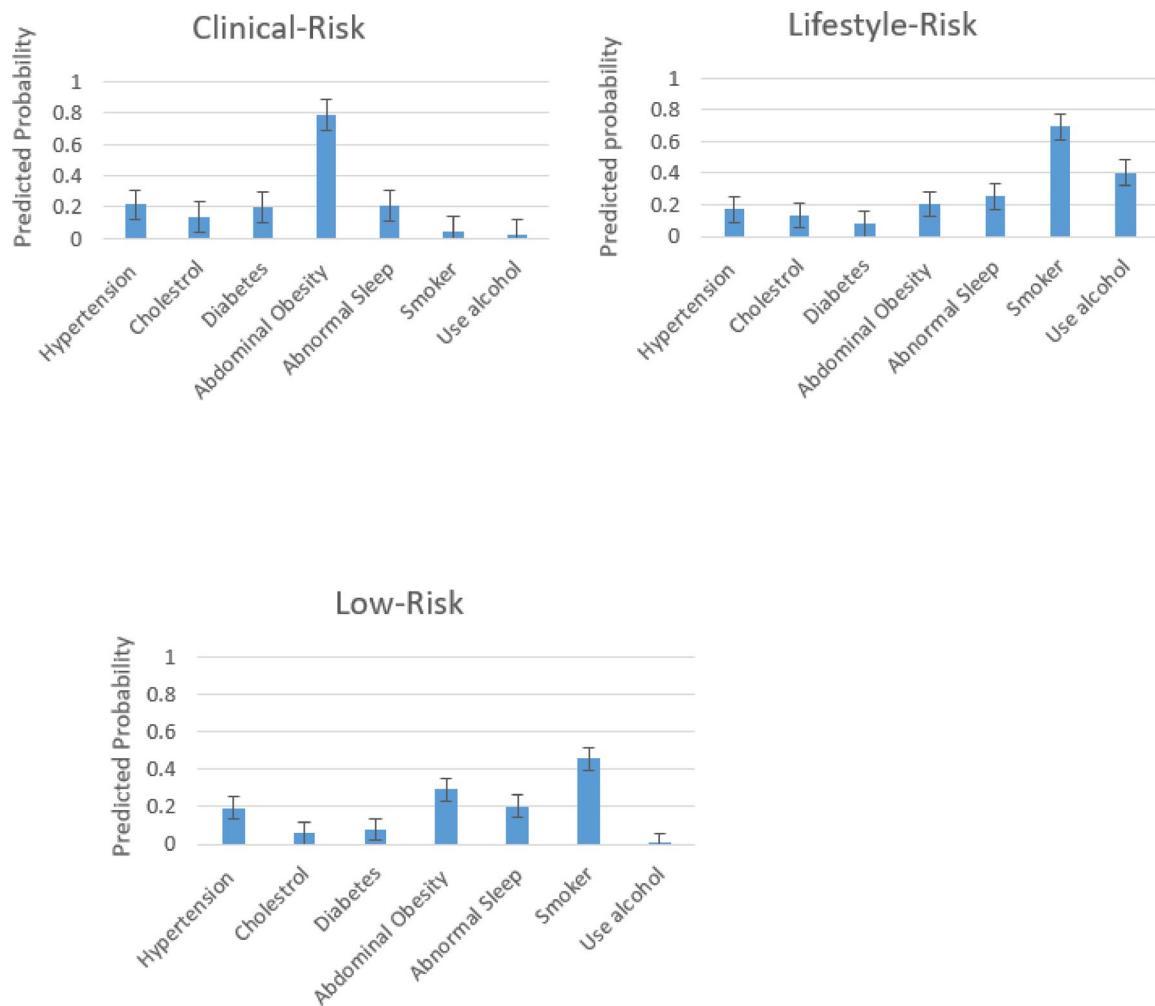
The present study started with 10,611 participants. After excluding 1,189 subjects with CVDs at baseline, 9,422 subjects were included in the analysis. Among the 9,422 participants, 1,917 (20.3%) had hypertension, 962 (10.2%) had high blood total cholesterol, 1,335 (14.2%) had DM, and 5,140 (54.6%) had abdominal obesity. Additionally, 1,991 subjects (21%) had a disturbed sleep, 2,392 (25%) were smokers, and 351 (4%) consumed alcohol (Table 1).

### Number of latent classes

The variables hypertension, high blood Total cholesterol, DM, abdominal obesity, disturbed sleep, smoking, and alcohol consumption were used as indicator variables in the model. To determine the best-fitting number of latent classes (LCs), we began with a two-class model and incrementally increased the number of classes up to four. Model evaluation criteria included BIC, AIC, entropy, and interpretability. Based on these criteria, three LCs were selected as the best fit for individuals at risk of developing CVD (Table 2).

### Identified classes labeling

The largest class was Class 1 (52%), followed by Class 3 (42%) and Class 2 (6%). The predicted probabilities of observed variables for each class are shown in Table 3 and Fig. 1.



**Fig. 1.** Predicted probabilities of CVD risks by LCs from LCA.

- Class 1 (Clinical Risk): High probability of clinical risks but relatively healthy lifestyle.
- Class 2 (Lifestyle Risk): Unhealthy lifestyle with low probability of clinical risks.
- Class 3 (Low Risk): Relatively healthy lifestyle with low probability of clinical risks.

#### The relationship between identified risk groups with demographic variables and socioeconomic status

The findings indicated significant associations between risk group membership and various demographic and socioeconomic factors:

The odds of belonging to the clinical risk group compared to the low-risk group were 13.5 times higher for women than men (Adjusted OR: 13.48, 95% CI: 11.81–15.39).

The odds of being in the clinical risk group versus the low-risk group were 24% higher for rural residents compared to urban residents (Adjusted OR: 0.76, 95% CI: 0.67–0.86).

For the 60–70 age group, the odds of clinical risk group membership compared to the low-risk group were 16% higher than for the 40–49 and 50–59 age groups (Adjusted OR: 1.16, 95% CI: 0.99–1.35).

The odds of being in the clinical risk group compared to the lifestyle risk group (Adjusted OR: 1.28, 95% CI: 1.14–1.43) and the low-risk group (Adjusted OR: 1.51, 95% CI: 1.11–2.07) were 28% and 50% higher, respectively, for the Fars ethnicity compared to other ethnicities (Table 4).

#### Discussion

The current study utilized baseline data from the Kherameh cohort to identify patterns of CVDs risk factors and their relation to demographic characteristics using the LCA model. In this study we found the following: Firstly, three LCs of modifiable risk factors for CVDs were recognized (i.e. clinical risk class, lifestyle risk class, and low risk class). Secondly, Female gender, older age and rural residence had a greater risk of being in clinical-risk class compared to the low-risk class. Lastly, individuals of Fars ethnicity revealed a significantly elevated risk of being classified in the clinical risk class and they recognized an obviously higher risk of belonging to the lifestyle risk class.

Variable		Frequency (%)	Clinical risk class**			Lifestyle class**		
			Adjusted OR	95% CI	P-Value	Adjusted OR	95% CI	P-Value
Gender	Male	4196 (44.5)	References					
	Female	5226 (55.5)	13.48	11.81–15.39	<0.0001*	0.75	0.49–1.14	0.181
Age	40–49	4269 (45.3)	Reference					
	50–59	3289 (34.9)	0.97	0.86–1.09	0.620	0.80	0.58–1.10	0.174
	60–70	1864 (19.8)	1.16	0.99–1.35	0.056*	1.09	0.73–1.61	0.659
Ethnicity	Others	3125 (33.2)	Reference					
	Fars	6297 (66.8)	1.28	1.14–1.43	<0.0001*	1.51	1.11–2.07	0.008*
Marital status	Single	1002 (10.6)	Reference					
	Married	4269 (89.4)	1.13	0.96–1.34	0.129	1.36	0.66–2.78	0.395
Living place	Rural	6044 (64.1)	Reference					
	Urban	3378 (35.9)	0.76	0.67–0.86	<0.0001*	1.28	0.93–1.75	0.118
Employment status	Unemployed	4457 (47.3)	Reference					
	Employed	4965 (52.7)	0.88	0.78–1.00	0.062	0.99	0.68–1.44	0.990
Education level	Illiterate	2920 (31.0)	Reference					
	Lower-diploma and diploma	5991 (63.6)	1.02	0.90–1.16	0.714	1.02	0.72–1.45	0.895
	Academic	511 (5.4)	1.18	0.91–1.52	0.194	0.55	0.25–1.18	0.128
Socioeconomic status	Low-	2375 (25.2)	Reference					
	Medium	2634 (28.0)	0.93	0.81–1.06	0.319	1.12	0.76–1.64	0.549
	High	4403 (46.8)	1.01	0.88–1.15	0.854	1.22	0.86–1.73	0.247

**Table 4.** Association between demographic variables with identified clinical and lifestyle risk classes compared to the low risk class as a reference group. \*Statistically significant. \*\*Reference group: Low risk class.

### Patterns of modifiable CVD risk factors

Three LCs of modifiable risk factors for CVDs were identified: 1. Clinical Risk Class: High probability of DM, hypertension, central obesity, and high blood total cholesterol, 2. Lifestyle Risk Class: Unhealthy lifestyle including disturbed sleep, smoking, and alcohol consumption, 3. Low Risk Class: Relatively healthy lifestyle with a low probability of clinical risks. Based on three LCs (i.e. clinical risk, life-style risk and low risk), the clinical risk class was the largest, followed by the low risk class. To our knowledge, limited studies have used LCA to identify LCs of CVD. Different criteria used by researchers make direct comparisons challenging. One such study, conducted on 8,218 UK adults over 50 years, identified four CVD risk factor classes: low risk, high risk, clinical risk, and lifestyle risk<sup>25</sup>.

Our findings showed that modifiable CVD risk factors divided into two main classes: clinical and lifestyle risks. Individuals with hypertension often had other CVD-related clinical conditions, and sleep disturbances were associated with other lifestyle risks such as smoking and alcohol consumption. This is consistent with previous studies indicating a high prevalence of co-occurring conditions like diabetes, hypertension, dyslipidemia, and obesity, as well as the clustering of health behaviors<sup>26,27</sup>.

### The relationship between identified risk groups with demographic variables and socioeconomic status

We found that membership in the clinical risk class was more common among individuals aged 60–70 years. This aligns with studies showing that the average age in high-risk groups for CVD is around 69 years<sup>28</sup>. One of the reasons why the risk of CVD increases with age can be due to a combination of physiological changes and the accumulation of risk factors over time. It was also stated in the previous study<sup>29</sup>, the elderly population are especially susceptible to cardiovascular diseases. Age is an independent risk factor for CVD in adults, these risks are compounded by other factors including obesity and diabetes. These complex factors are known as age-related CVD risk factors.

Our results also indicated that clinical risk membership was significantly higher in women than men, which is consistent with some studies<sup>14,30–32</sup>. As mentioned, gender is another potential risk factor in older adults, with older women reported to be at greater risk for CVD than age-matched men. However, in both men and women, the risks associated with CVD increase with age, and these risks may be related to a general decline in sex hormones, mainly estrogen and testosterone<sup>29</sup>.

Interestingly, our study found that the clinical risk class was more prevalent in rural areas, contrary to many studies that report higher prevalence of CVD risk factors in urban areas<sup>33–35</sup>. But studies similar<sup>36,37</sup> to ours show a higher prevalence of clinical CVD risk factors in rural people. Several factors may contribute to the higher prevalence of CVD risk factors in rural areas, including inequities in access to health care, low socio-economic status and level of education in rural areas and a lack of grocery stores that offer affordable and healthy foods. But in general, the exact reasons for these rural–urban health inequalities are unclear and need further investigation.

A previous study<sup>38</sup> has shown that there is heterogeneity and diverse distribution of known lifestyle-related risk factors for heart disease among major Iranian ethnic groups. In our study, Fars ethnicity was more likely to be in the clinical and lifestyle risk classes compared to other ethnicities, which supports findings from other

research<sup>39</sup>. This ethnic difference in CVD risks can be due to differences in diet, physical activity, smoking, and other lifestyle behaviors among ethnic groups, which are closely related to the development of hypertension, obesity, diabetes, and dyslipidemia. Also, differences in socioeconomic status, education level, and access to health care resources among ethnic communities could be other reasons for this result. Additionally, potential genetic and biological differences between ethnic groups and the role of cultural practices, traditions, and environmental exposures unique to different ethnic communities may influence CVD risk.

In this study, no relationship was observed between class membership with employment status, education level, marital status, and SES. Meanwhile, a study conducted on 2,380 cases from five cities of Stanford investigated the independent relation between education level and income with a set of CVD risk factors, such as cigarette smoking, systolic and diastolic blood pressure, total cholesterol and HDL, and found that lower education and income were associated with higher risk of CVD development<sup>40</sup>. Studies on lifestyle-related CVD risk factors with SES have shown contradictory results, which may depend on the definition of SES and lifestyle factors as well as population characteristics<sup>41</sup>. It should be noted that previous studies have shown contradictory results regarding the relation between marital status and CVD risk<sup>39,42,43</sup>.

### Strengths and limitations

This study is the first to examine the latent structure of modifiable CVD risk factors in Iran using Latent Class Analysis (LCA), a method with distinct advantages over traditional clustering techniques. LCA, as a cross-sectional latent variable mixture modeling approach, identifies subgroups that are internally homogeneous and externally heterogeneous<sup>44,45</sup>, using maximum likelihood estimation to ensure robust classification<sup>46</sup>. It also provides fit statistics<sup>47</sup>, which allow for the selection of the most appropriate model<sup>48</sup> and hypothesis testing, as well as information on the probability of individuals belonging to specific classes<sup>49,50</sup>. These features make LCA particularly suitable for analyzing complex, categorical data such as the risk factors in our study.

Moreover, this research was conducted with a large sample size and high accuracy, providing a strong basis for the identification of latent classes of modifiable CVD risk factors. However, several limitations should be noted:

The analysis included only individuals aged 40 years or older, making it unclear if the same patterns occur in younger age groups. Additionally, the cross-sectional design prevents establishing causality. Future research should explore how clusters of CVD risk factors evolve over time and which changing behavioral or clinical risk patterns pose the highest risk for CVD.

The cohort was derived from the Kherameh region, which has a rural context and specific demographic characteristics. These unique attributes may limit the generalizability of our findings to other populations, particularly urban settings or populations with different socioeconomic and cultural profiles. Future studies should replicate this analysis in other regional and demographic contexts to validate and expand the patterns identified in this study. This includes conducting similar analyses in urban and mixed populations to test the robustness and applicability of the findings in diverse settings.

Our focus on specific subtypes of CVD, including ischemic heart disease, myocardial infarction (MI), stroke, and cerebrovascular accident (CVA), may also limit the generalizability of our findings to all subtypes of CVD. This specificity means that the risk factor patterns identified might not fully represent other forms of cardiovascular disease not explicitly included in our definition.

Future studies should consider longitudinal designs to track the evolution of CVD risk factors over time and examine the impact of interventions aimed at reducing risk factors in high-risk groups. Moreover, investigating genetic and environmental interactions contributing to CVD risk in diverse populations could provide valuable insights.

By addressing these limitations and expanding research efforts in this area, including replication studies in different regional and demographic contexts, the findings of our study could pave the way for a more comprehensive understanding of modifiable CVD risk factors and their implications for public health strategies.

While our study provides valuable insights into the latent structure of modifiable CVD risk factors, it does not directly assess the predictive ability of these risk factors in forecasting CVD outcomes. The primary focus of this study was on identifying distinct latent classes of risk factors rather than developing predictive models. As such, the applicability of the identified latent classes for direct prediction of CVD events remains an area for future research.

Additionally, the use of Latent Class Analysis (LCA) inherently focuses on subgroup identification based on observed data patterns rather than predicting outcomes. The lack of a labeled outcome variable in our dataset, such as CVD diagnosis or progression, limits the ability to employ supervised machine learning methods to evaluate predictive utility.

Future studies could address this limitation by incorporating longitudinal data or labeled outcomes, enabling the use of machine learning or deep learning models to validate the predictive utility of the identified latent classes. This would provide a deeper understanding of how these latent structures correlate with CVD risk over time and their potential use in clinical risk prediction.

### Implications for practice

These findings enhance our understanding of modifiable CVD risk factors by showing the interaction between different risks and identifying distinct patterns of risk factors using Latent Class Analysis (LCA). This method allowed us to classify individuals into subgroups with unique combinations of modifiable risks, offering valuable insights into how these risks co-occur within a population.

In clinical practice, these insights can inform the development of targeted interventions tailored to specific risk profiles, thereby optimizing resource allocation and improving health outcomes. For example, individuals in high-risk classes could benefit from comprehensive lifestyle modification programs, while those in lower-

risk classes might require less intensive interventions. Additionally, understanding these patterns can aid in designing health protocols that address the unique needs of different demographic groups, particularly in settings with similar rural and demographic characteristics to the Kherameh cohort.

While LCA has provided robust insights, we acknowledge that future studies could explore complementary methods to further validate these findings and investigate how they apply to other populations and settings.

### Policy recommendations

Based on our findings, several policy recommendations can be made to address CVD risk factors in similar populations. Public health interventions should focus on:

**Targeting Key Risk Factors:** Prioritize interventions addressing abdominal obesity and smoking, which are prominent risk factors in our identified groups. This approach can streamline efforts and enhance the effectiveness of health initiatives.

**Increasing Awareness:** Develop tailored awareness campaigns that stress the importance of regular health check-ups, with an emphasis on managing weight and quitting smoking, particularly targeting rural and older populations.

**Implementing Targeted Educational Campaigns:** Focus educational efforts on reducing smoking and alcohol consumption, with special attention to how these behaviors interact with other risk factors like obesity.

**Enhancing Access to Healthcare:** Improve healthcare access in rural areas to detect and manage clinical risk factors early, ensuring services are equipped to address issues like obesity and smoking cessation.

### Data availability

The data that support the findings of this study are available from the corresponding author, [HGH], upon reasonable request.

Received: 28 July 2024; Accepted: 20 June 2025

Published online: 08 July 2025

### References

1. Joseph, P. et al. Reducing the global burden of cardiovascular disease, part 1: the epidemiology and risk factors. *Circ. Res.* **121**, 677–694 (2017).
2. Sarrafzadeegan, N. & Mohammadi, N. Cardiovascular disease in Iran in the last 40 years: Prevalence, mortality, morbidity, challenges and strategies for cardiovascular prevention. *Arch. Iran. Med.* **22**, 204–210 (2019).
3. Ghaemian, A., Nabati, M., Saeedi, M., Kheradmand, M. & Moosazadeh, M. Prevalence of self-reported coronary heart disease and its associated risk factors in Tabari cohort population. *BMC Cardiovasc. Disord.* **20**, 1–10 (2020).
4. Baeradeh, N. et al. The prevalence and predictors of cardiovascular diseases in Kherameh cohort study: A population-based study on 10,663 people in southern Iran. *BMC Cardiovasc. Disord.* **22**, 244 (2022).
5. World Health, O. *Global status report on noncommunicable diseases 2010*. (World Health Organization, 2011).
6. Sarraf-Zadeegan, N. et al. Secular trends in cardiovascular mortality in Iran, with special reference to Isfahan. *Acta Cardiol.* **54**, 327–333 (1999).
7. Hippisley-Cox, J. et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* **336**, 1475–1482 (2008).
8. Yusuf, S. et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): Case-control study. *The Lancet* **364**, 937–952 (2004).
9. Huxley, R., Mendis, S., Zheleznyakov, E., Reddy, S. & Chan, J. Body mass index, waist circumference and waist: Hip ratio as predictors of cardiovascular risk—a review of the literature. *Eur. J. Clin. Nutr.* **64**, 16–22 (2010).
10. Steptoe, A. & Kivimäki, M. Stress and cardiovascular disease: An update on current knowledge. *Annu. Rev. Public Health* **34**, 337–354 (2013).
11. Timmis, A. et al. European society of cardiology: Cardiovascular disease statistics 2019. *Eur. Heart J.* **41**, 12–85 (2020).
12. Poustchi, H. et al. Prospective epidemiological research studies in Iran (the PERSIAN Cohort Study): Rationale, objectives, and design. *Am. J. Epidemiol.* **187**, 647–655 (2018).
13. Conroy, R. M. et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: The SCORE project. *Eur. Heart J.* **24**, 987–1003 (2003).
14. D'Agostino, R. B. Sr. et al. General cardiovascular risk profile for use in primary care: The Framingham Heart Study. *Circulation* **117**, 743–753 (2008).
15. Saugel, B., Dueck, R. & Wagner, J. Y. Measurement of blood pressure. *Best Pract. Res. Clin. Anaesthesiol.* **28**, 309–322 (2014).
16. Ghosh, A., Halder, D. & Mukhopadhyay, K. Comparative assessment of blood pressure, random blood sugar and addiction status among security staffs and scavengers in an Indian medical college hospital. *Int. J. Curr. Microbiol. App. Sci.* **5**, 294–300 (2016).
17. Tayefi, M. et al. Systolic and diastolic blood pressure percentiles by age and gender in Northeastern Iran. *J. Am. Soc. Hypertens.* **12**, e85–e91 (2018).
18. Roth, G. A. et al. High total serum cholesterol, medication coverage and therapeutic control: an analysis of national health examination survey data from eight countries. *Bull. World Health Organ.* **89**, 92–101 (2011).
19. McGee, S. *Evidence-based physical diagnosis e-book* (Elsevier Health Sciences, 2021).
20. Cappuccio, F. P., Cooper, D., D'Elia, L., Strazzullo, P. & Miller, M. A. Sleep duration predicts cardiovascular outcomes: A systematic review and meta-analysis of prospective studies. *Eur. Heart J.* **32**, 1484–1492 (2011).
21. Min, Y. & Slattum, P. W. Poor sleep and risk of falls in community-dwelling older adults: A systematic review. *J. Appl. Gerontol.* **37**, 1059–1084 (2018).
22. Aflaki, K., Vigod, S. & Ray, J. G. Part I: A friendly introduction to latent class analysis. *J. Clin. Epidemiol.* **147**, 168–170 (2022).
23. Hagenaars, J. A. & McCutcheon, A. L. *Applied latent class analysis* (Cambridge University Press, 2002).
24. Clark, S. L. & Muthén, B. (Los Angeles, California, USA, 2009).
25. Bu, F., Steptoe, A. & Fancourt, D. Relationship between loneliness, social isolation and modifiable risk factors for cardiovascular disease: A latent class analysis. *J. Epidemiol. Community Health* **75**, 749–754 (2021).
26. Weycker, D. et al. Risk-factor clustering and cardiovascular disease risk in hypertensive patients. *Am. J. Hypertens.* **20**, 599–607 (2007).
27. Poortinga, W. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Prev. Med.* **44**, 124–128 (2007).

28. Wei, X. et al. Characteristics of high risk people with cardiovascular disease in Chinese rural areas: Clinical indicators, disease patterns and drug treatment. *PLoS ONE* **8**, e54169 (2013).
29. Rodgers, J. L. et al. Cardiovascular risks associated with gender and aging. *J. Cardiovasc. Dev. Dis.* <https://doi.org/10.3390/jcdd6020019> (2019).
30. Choinière, R., Lafontaine, P. & Edwards, A. C. Distribution of cardiovascular disease risk factors by socioeconomic status among Canadian adults. *CMAJ* **162**, S13–S24 (2000).
31. Yu, Z. et al. Associations between socioeconomic status and cardiovascular risk factors in an urban population in China. *Bull. World Health Organ.* **78**, 1296–1305 (2000).
32. Cho, C.-M. & Lee, Y.-M. The relationship between cardiovascular disease risk factors and gender. *Health* **4**, 309–315 (2012).
33. Gu, D. et al. Prevalence of cardiovascular disease risk factor clustering among the adult population of China: Results from the international collaborative study of cardiovascular disease in Asia (InterAsia). *Circulation* **112**, 658–665 (2005).
34. Costache, I. I. et al. Associations between area of residence and cardiovascular risk. *Revista de Cercetare si Interventie Sociala* **49**, 68 (2015).
35. Tareen, M. F. et al. Location of residence or social class, which is the stronger determinant associated with cardiovascular risk factors among Pakistani population? A cross sectional study. *Rural Remote Health* **11**, 32–39 (2011).
36. Joshi, R., Taksande, B., Kalantri, S. P., Jajoo, U. N. & Gupta, R. Prevalence of cardiovascular risk factors among rural population of elderly in Wardha district. *J. Cardiovasc. Dis. Res.* **4**, 140–146. <https://doi.org/10.1016/j.jcdr.2013.03.002> (2013).
37. Turecamo, S. E. et al. Association of rurality with risk of heart failure. *JAMA Cardiol.* **8**, 231–239. <https://doi.org/10.1001/jamacardio.2022.5211> (2023).
38. Babahajani, M. et al. Ethnic differences in the lifestyle behaviors and premature coronary artery disease: A multi-center study. *BMC Cardiovasc. Disord.* **23**, 170. <https://doi.org/10.1186/s12872-023-03192-0> (2023).
39. Ghorashi, S. M. et al. Association between nontraditional risk factors and calculated 10-year risk of atherosclerotic cardiovascular disease in a large general population: Based on the pars cohort study. *J. Tehran Univ. Heart Center* **18**, 24 (2023).
40. Winkleby, M. A., Jatulis, D. E., Frank, E. & Fortmann, S. P. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am. J. Public Health* **82**, 816–820 (1992).
41. Zhang, Y.-B. et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: Two prospective cohort studies. *BMJ* **373**, n604 (2021).
42. Alviar, C. L., Rockman, C., Guo, Y., Adelman, M. & Berger, J. Association of marital status with vascular disease in different arterial territories: A population based study of over 3.5 million subjects. *J. Am. Coll. Cardiol.* **63**, A1328–A1328 (2014).
43. Chen, R. et al. Marital status, telomere length and cardiovascular disease risk in a Swedish prospective cohort. *Heart* **106**, 267–272 (2020).
44. Dumith, S. C., Muniz, L. C., Tassitano, R. M., Hallal, P. C. & Menezes, A. M. B. Clustering of risk factors for chronic diseases among adolescents from Southern Brazil. *Prev. Med.* **54**, 393–396 (2012).
45. Schuit, A. J., van Loon, A. J. M., Tijhuis, M. & Ocké, M. C. Clustering of lifestyle risk factors in a general adult population. *Prev. Med.* **35**, 219–224 (2002).
46. Berlin, K. S., Williams, N. A. & Parra, G. R. An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses. *J. Pediatr. Psychol.* **39**, 174–187 (2014).
47. Weller, B. E., Bowen, N. K. & Faubert, S. J. Latent class analysis: A guide to best practice. *J. Black Psychol.* **46**, 287–311 (2020).
48. Vermunt, J. K. & Magidson, J. Latent class analysis. *Sage Encycl. Soc. Sci. Res. Methods* **2**, 549–553 (2004).
49. Miettunen, J., Nordström, T., Kaakinen, M. & Ahmed, A. O. Latent variable mixture modeling in psychiatric research—A review and application. *Psychol. Med.* **46**, 457–467 (2016).
50. Tegegne, T. K., Islam, S. M. S. & Maddison, R. Effects of lifestyle risk behaviour clustering on cardiovascular disease among UK adults: Latent class analysis with distal outcomes. *Sci. Rep.* **12**, 17349 (2022).

## Acknowledgements

This article was a part of Hadis Salimi's MSc thesis, approved and financially supported by the Research Vice-chancellor of Shiraz University of Medical Sciences (Grant no. 1402-4-4-28337). The Ethics Committee approved this study at Shiraz University of Medical Sciences (IR.SUMS.SCHEANUT.REC.1402.052). This paper was edited using large language models (LLMs), including Bard and ChatGPT. However, the accuracy of all information was independently verified by the authors.

## Author contributions

H.G. is the lead author and guarantor and contributed to interpreting the data and revising the manuscript. H.S. and M.V. planned the study and led the drafting and revising of the manuscript. H.S., M.V., A.R., M.Gh. and H.G. contributed to interpreting the data and drafting and revising the manuscript. All authors approved the submitted version of the manuscript. All authors have contributed to the preparation of the manuscript, have read, and approved the submitted manuscript. All authors listed meet the authorship criteria according to the latest guidelines of the International Committee of Medical Journal Editors and agree with the manuscript. The work is original and not under consideration by any other journal.

## Funding

This article was a part of Hadis Salimi's MSc thesis, approved and financially supported by the Research Vice-chancellor of Shiraz University of Medical Sciences (Grant no. 1402-4-4-28337).

## Declarations

### Competing interests

The authors declare no competing interests.

### Ethical approval

PERSIAN Cohort Study is being performed in 18 geographical regions of Iran. PERSIAN Cohort Study was approved by the Ethics Committee of the Ministry of Health and Medical Education. This study was in agreement with the Helsinki Declaration and Iranian national guidelines for ethics in research (IR.SUMS.SCHEANUT.REC.1402.052). Besides, Informed consent was obtained from all the participants.

### Additional information

Correspondence and requests for materials should be addressed to H.G.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025