Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Patterns and Related Influencing Factors of Multiple Chronic Diseases Among the Elderly Based on Data Mining in Shanxi, China

**Chichen Zhang** ( zhangchichen@sina.com )
Southern Medical University https://orcid.org/0000-0003-1095-9939
**Shujuan Xiao**
Southern Medical University
**Lei Shi**
Southern Medical University
**Xiao Zheng**
Southern Medical University
**Fang Dong**
Southern Medical University
**Yaqing Xue**
Southern Medical University
**Jiayu Zou**
Southern Medical University
**Jiachi Zhang**
Southern Medical University

**Research**

**Keywords:** Multiple chronic diseases, Data mining, Association rule, Decision tree model, Elderly, Related influencing factors

**Posted Date:** October 23rd, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-94449/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

# Abstract

**Background:** The presence of multiple chronic diseases has become a public health issue worldwide. The study aimed to explore the patterns and related influencing factors of multiple chronic diseases among the elderly so as to provide a scientific reference for the health management strategy, which is dedicated to the prevention and intervention for multiple chronic diseases.

**Methods:** A cross-sectional study, which used a multi-stage random sampling way, was conducted among 3266 elderly aged 60 years and above in Shanxi. A chi square test was used to compare the sociodemographic factors. The data mining methods, including association rule and decision tree, were performed to explore the patterns and related influencing factors of multiple chronic diseases in the elderly.

**Results:** The prevalence of multiple chronic diseases was 30.31%. According to the results of the association rules, there were 10 association patterns screened out. The first five multiple chronic disease patterns were atherosclerosis and hypertension, coronary heart disease and hypertension, diabetes and hypertension, stroke and hypertension, eye disease and hypertension. The decision tree model selected 5 related influencing factors of multiple chronic diseases, including age, smoking, family genetic history, fresh fruit intake frequency and education level, simultaneously generating 9 influence rules.

**Conclusions:** Multiple chronic diseases were prevalent among the elderly, which patterns mainly consisted of two or three chronic diseases. Regarding the unmodifiable related influencing factors, including an older age, a family genetic history, and low education level, relevant departments should help the elderly to establish prevention knowledge and raise awareness of strongly related chronic diseases, so as to reduce the incidence and adverse consequences.


# 1. Introduction

### 1.1 The prevalence and harm of multiple chronic diseases

In recent decades, life expectancy has increased and mortality has declined steadily [1], leading to the prevalence of chronic diseases worldwide. What is particularly worrying is that many people have more than one chronic disease. Approximately one-third of all adults worldwide are diagnosed with multiple chronic diseases [2]. Multiple chronic diseases are defined as two or more chronic diseases within one person in a specific period of time [3]. Compared with single chronic disease, the treatment difficulty, medical consumption, economic burden, and risk of death of patients with multiple chronic diseases increase, driving intense demand for effective clinical interventions and health management support for these patients [4-7]. Although the established interventions for the management of chronic diseases have the potential to improve the care of the elderly, they are only focused on a single disease [8]. At present, the health management and clinical treatment of elderly patients with multiple chronic diseases have become one of the major challenges faced by the medical and health system worldwide [9].

### 1.2 Strengths and limitations of previous studies

Previous studies had focused on comorbidities of various chronic diseases, including type 2 diabetes, chronic rheumatic, musculoskeletal diseases and so on [10-12]. In addition, Donna et al. [13] have pointed out a taxonomy to inform care for patients with multiple chronic conditions, which is generalizable across multiple types of comorbidities. Although comorbidities have been the focus of research in the past few decades, multiple chronic diseases have gradually become research hotspots in recent years. Research on multiple chronic diseases mainly focuses on prevalence, health care utilization and health-related outcomes [14-17]. Recently, there has been an effort to describe the full patterns of multiple chronic diseases to provide complete information. Noe et al. [18] described multimorbidity patterns in low, middle, and high-income countries. Nevertheless, there were limited studies on how chronic diseases existed in some patterns of multiple chronic diseases, which may be meaningful for improving the management of multiple chronic diseases. This study, however, was limited to the 12 chronic diseases included. Bauer et al. pointed out that chronic disease burden largely results from a short list of risk factors [19]. To the best of our knowledge, few studies were identifying potential factors underlying multiple chronic diseases.

### 1.3 Association Rule and Decision Tree

Data mining technique can extract potential knowledge and patterns that will be of interest from an extremely abundance of data. As the data mining technology, association rule and decision tree all can help researchers extract valuable knowledge from huge data. The applied rule mining approach was the Apriori algorithm, which mined valuable patterns in large unordered data as association rules. Another applied data mining technology was the decision tree, which was easy to implement and interpret. The major purpose of this method is to make a predictive model for the target variable [20]. The main advantage of decision tree analysis is that it can transform the complex interaction of various variables into an organized flow chart, which can clearly identify related influencing factors. The CHAID algorithm in the decision tree is based on the self-stratification of dependent variables, which automatically merges the independent variable categories to maximize their significance. In clinical and practice, it is of great significance for obtaining a stratified tool of related influencing factors to better manage multiple chronic diseases. Therefore, this study adopted the CHAID algorithm to define new models for predicting the related influencing factors of multiple chronic diseases in the elderly.

### 1.4 Aim and hypotheses

To the best of our knowledge, there is few study using data mining methods to simultaneously explore the patterns and their related influencing factors in combination with multiple chronic diseases. Therefore, this study aimed to explore the common patterns and related influencing factors further of 26 chronic diseases, employing association rules and the decision tree, which provided a scientific reference for the health management strategy.

Based on the above, we proposed two hypotheses:

1. The most common pattern of 26 diseases is a combination of two or more conditions, among which the most prevalent dyads and triads include hypertension, diabetes and coronary heart disease.
2. At the same time, age, sex or other factors are the main related factors affecting the presence of multiple chronic diseases.

# 2. Methods

## 2.1 Sample and participants

The selected survey site was Shanxi Province in Central China. Using the stratified random cluster sampling method, we recruited older adults (aged 60 years and above) from Shanxi Province's 11 cities. The sampling method was as follows: In the first stage, according to the order of districts (counties) on the government's website, each district (county) in every city was numbered. In the second stage, two (districts) counties in each city were selected using the random number table, and then two communities (administrative villages) were drawn from each district (county) in the same way. In the fourth stage, considering the different scales of each community (administrative village), once again select 1 to 2 residential communities (natural villages) from each community (administrative village). Finally, in the selected residential communities (natural villages), we randomly selected the elders who meet the criteria in this study.

The inclusion criteria for this study were (1) being 60 years and older; (2) having clear awareness and barrier-free communication skills with the investigator; and (3) volunteering to participate in the survey. Meanwhile, those who have cognitive dysfunction, unwilling to cooperate, were excluded. The study was conducted from June to August, 2019, involving 3266 elders, of which 3250 completed the questionnaire effectively, so the effective response rate was 99.51%.

All study procedures were approved by the university ethics committee. All participants were informed of the purpose and procedure of the research upon their recruitment, and assured of their right to refuse to participate. Their anonymity and confidentiality were guaranteed. After signing the consent, participants were invited to conduct questionnaires to collect data by trained investigators.

## 2.2 Instruments

The questionnaire comprised two sections: the self-made general information and prevalence of chronic diseases. The self-made general information includes (1) socio-demographic characteristics of the research subjects, including living area, gender, age, body mass index (BMI), occupation, education level, monthly family income, marital status and empty nest situation; and (2) lifestyle behaviors, including dietary habits and smoking and drinking status. The questionnaire on prevalence of chronic diseases includes 26 chronic diseases diagnosed by doctors, including obesity, hypertension, diabetes, coronary heart disease, stroke, arrhythmia, atherosclerosis, tuberculosis, respiratory diseases, Parkinson's disease, chronic obstructive Pneumonia, sciatica, rheumatoid or rheumatoid arthritis, hypothyroidism,

hyperthyroidism, gout, osteoporosis, hearing impairment, eye disease, hepatitis, chronic nephritis, tuberculosis, mental illness, dementia, digestive system diseases, and cancer.

### 2.3 Statistical analysis

The data were analyzed by SPSS Version 24.0 Statistical software. Descriptive statistical analyses were carried out for basic socio-demographic characteristics. The chi-square test was used to compare the differences in the prevalence of multiple chronic diseases under different general demographic information among the elderly. Confidence level with $P < 0.05$ was considered statistically significant.

The Apriori algorithm in SPSS Modeler Version 18.0 software was employed to analyze common patterns of multiple chronic diseases. Three kernel values are involved with association rule analysis, including support, confidence, and lift [21]. Based on this study, the support of A→B was the probability of the simultaneous occurrence of chronic disease A and B. The confidence was the conditional probability of suffering from chronic disease B under the premise of suffering from chronic disease A. The degree of lift reflects the influence of the consequent B by the antecedent A compared to the overall. Therefore, when the degree of lift $L_{A\rightarrow B}>1$, A→B can be considered as a directional association. In this study, set the minimum conditional support to 3.0%, the minimum rule confidence to 30%, and the maximum number of preceding items to 5.

With the help of the CHAID algorithm of the decision tree in SPSS Version 24.0 Statistical software, the related influencing factors affecting multiple chronic diseases of the elderly were analyzed. Taking the prevalence of multiple chronic diseases as the dependent variable (0=without chronic disease, 1=having 1 chronic disease, 2=having 2 or above chronic diseases), and the variables with statistical significance of univariate analysis as independent variables, built a decision tree model. The maximum tree depth of the decision tree was set to 3. The minimum sample size of the parent node was set to 100 and the minimum sample size of the child node was set to 50. The growth direction of the decision tree was from top to bottom.

## 3. Results

### 3.1 Sociodemographic characteristics

In total, 3250 participants completed the questionnaire effectively. Of all elderly in this survey, there were 1515 males (46.62%), and 1735 females (53.38%); 1453 (44.71%) lived in urban areas and 1797 (55.29%) in rural areas ; 577 (17.75%) had no formal education, 1062 (32.68%) with primary school education, 850 (26.15%) with junior high school education, 474 (14.58%) with high school/secondary school education, 144 (4.43%) with junior college/high school education, 143 people (4.40%) with bachelor degree or above; 166 (5.11%) had a family genetic history. The average age of the respondents was 69.65 years old. In this study, age was categorized into three categories (60 to 69, 70 to 79, 80 and above years). Among them, there were 1769 (54.43%) in the 60-69 age group, 1164 (35.82%) in the 70-79 age group, and 317 (9.75%) in the 80 and above age group.

### 3.2 The prevalence of chronic diseases in the elderly

Among the 3250 elderly people surveyed, 1901 (58.49%) had chronic diseases and 985 (30.31%) had multiple chronic diseases. The number of multiple chronic diseases ranged from 2 to 9. The coexistence of 2, 3, and 4 chronic diseases was relatively common, accounting for 54.11%, 26.19%, and 11.17%, respectively. The top ten prevalent chronic diseases were hypertension, diabetes, rheumatoid or rheumatoid arthritis, hearing impairment, digestive system diseases, osteoporosis, coronary heart disease, eye diseases, respiratory diseases (bronchitis, emphysema, asthma) and stroke. Table 1 showed the gender differences in the top 10 chronic diseases.

### 3.3 Analysis of association rules for multiple chronic diseases

The most common algorithm for association rule mining is the Apriori algorithm [22]. We employed the Apriori algorithm to obtain association rules between common chronic diseases in this study. Finally, 10 association patterns with strong association strength were selected. They were both patterns of coexistence of two chronic diseases, 8 of which were related to hypertension (Table 2). The most strongly associated comorbid pair was hypertension and atherosclerosis. Using the network graphics provided in SPSS Modeler 18.0 software, drew a network graph of the correlation strength between the 11 chronic diseases, which finally included 9 chronic diseases (Figure 1). The thickness of the line between the two points reflected the strength of the chronic disease [23]. The thicker the line between the two points, the stronger the correlation between chronic diseases was.

### 3.4 Univariate analysis of related influencing factors of multiple chronic diseases

The chi-squared test showed that there were significant differences in multiple chronic diseases between the elderly in age, gender, education level, living area, family genetic history, empty nest situation, marital status, smoking, drinking, diet, fresh fruit intake frequency, BMI ($P < 0.05$). However, there was no statistically significant difference in family monthly income ($P=0.125$), regularity of three meals ($P=0.524$), and fresh vegetable intake frequency ($P=0.127$) in multiple chronic diseases.

### 3.5 Results of Decision tree analysis

The decision tree model had a total of 15 nodes with 9 terminal nodes, which depth was 3 (Fig 2). Five main variables were affecting the presence of multiple chronic diseases, among which the first variable was age, and followed by smoking, family genetic history, fresh fruit intake frequency, and education level, with the significance of $P < 0.05$. The result had important practical implications, which suggested that it would be important to target these five factors in health management aimed to reduce the presence of multiple chronic disease. Modifiable lifestyle factors including smoking, and fresh fruit intake frequency affected the presence of chronic diseases. Unmodifiable factors included age, family genetic history and education level. Other variables did not reach the significance of 0.05 and were not included in the model such as gender, living area, empty nest situation, marital status, drinking, diet, and BMI.

In this study, the following 9 decision rules associated with different prevalence of multiple chronic diseases were extracted through decision tree model. Sort by multiple chronic disease rate from high to low as follows:

Rule 1: If the elderly aged 70-79 years old, had never smoked/were smoking and had a family genetic history, then the multiple chronic disease rate was 62.7%.

Rule 2: If age ≥80 years old, fresh fruit intake frequency was not daily, but at least once a week/not weekly, but at least once a month/not monthly, but sometimes / rarely or never eat fresh fruit, then the multiple chronic disease rate was 55.6%.

Rule 3: If the elderly aged 60-69 years, had a family genetic history, then the multiple chronic disease rate was 50.6%.

Rule 4: If the elderly aged 70-79 years old, had quit smoking, then the multiple chronic disease rate was 45.5%.

Rule 5: If age ≥80 years old, fresh fruit intake frequency was almost every day, then the multiple chronic disease rate was 31.8%.

Rule 6: If the elderly aged 70-79 years old, without family genetic history, had never smoking/was smoking, then the multiple chronic disease rate was 31.5%.

Rule 7: If the elderly aged 60-69 years, without family genetic history, and the education level was primary school/no formal education, then the multiple chronic disease rate was 28.1%.

Rule 8: If the elderly aged 60-69 years, without family genetic history, and education level was junior high school/high school/secondary school/high vocational/junior college, then the multiple chronic disease rate was 20.4%.

Rule 9: If the elderly aged 60-69 years, without family genetic history, education level was undergraduate and above, then the multiple chronic disease rate was 6.9%.

# 4. Discussion

### 4.1 Prevalence of multiple chronic diseases

In this study, the prevalence of multiple chronic diseases was 30.31% in the elderly aged 60 and above. In the United States, approximately one-third of the elderly suffer from multiple chronic diseases [24]. The high incidence of multiple chronic diseases of the elderly not only endangered the physical health, but also imposed a heavy threat to the mental health [25]. Therefore, it is of great significance to mine patterns and related influencing factors of multiple chronic diseases through the cross-sectional study data.

### 4.2 Patterns and association rules of multiple chronic diseases

This study found that patterns in 10 association rules all coexisted with two chronic diseases. For example, hypertension was associated with atherosclerosis, coronary heart disease, diabetes, stroke, and eye disease (cataracts, glaucoma). While sometimes hypertension was reported as genetic and early onset, for others it came on later as a result of another illness [26]. As mentioned above, there was a wide range of associations between various common chronic diseases. It suggests that the elderly may suffer from hypertension at the same time as they suffer from the above diseases. Additionally, this study also found that there were association rules between rheumatoid or rheumatoid arthritis and osteoporosis. Rheumatism or rheumatoid arthritis was a chronic disabling disease that was associated with increased localized and generalized osteoporosis [27]. This study also found association rules between coronary heart disease and atherosclerosis, suggesting that elderly people with atherosclerosis should screen for coronary heart disease as early as possible to reduce their incidence. In health management, staff should strengthen health education on the concept and risk of multiple chronic diseases, and take timely intervention measures as early as possible to address the challenges of multiple chronic diseases. In clinical, medical staff while treating patients who have been diagnosed with chronic diseases, should advise patients to screen for chronic diseases that are strongly associated with them as soon as possible, so that early detection, early diagnosis, and early treatment can be achieved. In other words, the knowledge of comprehensive chronic disease patterns should be taken into account when providing care for multimorbid patients.

### 4.3 Related influencing factors of multiple chronic diseases

The decision tree model showed that age was the root node, indicating that age was the primary and most important related influencing factor for multiple chronic diseases among independent variables. Among them, the prevalence rate of multiple chronic diseases (47.3%) aged 80 years and above was the highest.

In the elderly aged 80 and above, the main related influencing factor selected was the fresh fruit intake frequency. The multiple chronic disease rate (31.8%) of the elderly who ate fresh fruits almost daily was significantly lower than that of the elderly who consumed other frequencies (55.6%). It suggests that eating fresh fruit every day is a protective factor against multiple chronic diseases, which is consistent with the previous study [28].

Among the elderly aged 70-79 years, the multiple chronic disease rate (45.5%) who had quit smoking was higher than the combined elderly who had never smoked or were smoking. This study found that smoking cessation was associated with multiple chronic diseases in the elderly. It may be that some elderly people surveyed appeared to quit smoking after they were diagnosed with multiple chronic diseases. We speculate that the occurrence of chronic diseases will promote the establishment of healthy behavioral lifestyles for the elderly.

The elderly who had never smoked or were smoking, were divided into two categories according to family genetic history. The multiple chronic disease rate (62.7%) with a family genetic history was almost twice that without a family genetic history (31.5%). It suggests that the elderly population with a family genetic history is a key target for the prevention and treatment of chronic diseases. Some studies have reported the association between the family genetic history and risk of chronic disease [29, 30]. Among the elderly aged 60-69, without family genetic history, the higher the education level of the elderly, the lower the risk of disease was. The people with more education are more likely to receive health knowledge and develop a healthy behavioral lifestyle, and thus the rate of multiple chronic diseases is also lower [31, 32].

Drinking is a recognized risk factor for chronic diseases [33-35]. However, it was worth noting that this study did not find that drinking was associated with multiple chronic diseases in the elderly. The first reason was that this study was a cross-sectional study, which cannot prove causality. It only understood the current drinking status of the elderly at the time of the survey, and could not observe the relationship between them. The second reason was that the number of females was higher than males in this study. However, the drinking population was mostly male, and only accounting for only 16.4%. The choice of research sample may lead to a weak association between drinking and multiple chronic diseases. Based on this, this study still believed that drinking alcohol was a risk factor for multiple chronic diseases. However, it was speculated that the occurrence of multiple chronic diseases would prompt the elderly to actively change unhealthy behaviors and lifestyles. The health management department needs to make people aware of the harm caused by these modifiable life behavior factors for multiple chronic diseases through health education and health promotion activities, so as to actively change their behaviors and lifestyles. Regarding the unmodifiable related influencing factors, including age, family genetic history and education level, in clinical, clinicians should focus on such elderly people and take preventive measures.

### 4.4 Limitations

Although this study had certain value for the prevention and control of multiple chronic diseases in the elderly, it also had limitations. It was because that cross-sectional studies may cause information deviations when investigating and analyzing behavioral lifestyles. Some subjects had changed their behavior and lifestyle after suffering from chronic diseases. To a certain extent, the influence of the behavioral lifestyle on multiple chronic diseases was weakened. In addition, the prevalence of chronic diseases in this study was self-reported, and thus there may be information bias. On this basis, this study would further explore the association and related influencing factors of multiple chronic diseases through longitudinal research to obtain more scientific and objective research results.

## 5. Conclusion

Generally speaking, the challenges of multiple chronic diseases are serious, which has threatened their physical and mental health and placed a significant strain on society. Therefore, identifying the patterns and related influencing factors of multiple chronic diseases by data mining method is of great

significance to address the challenges of multiple chronic diseases. The most common patterns of 26 diseases are two condition combinations and three condition combinations. The health management department needs to make people aware of the harm caused by the modifiable life behavior factors for multiple chronic diseases through health education and health promotion activities, so as to actively change their behaviors and lifestyles. Regarding the unmodifiable related influencing factors, including an older age, a family genetic history, and low education level, relevant departments should help them to establish prevention knowledge and raise awareness of strongly related chronic diseases, so as to reduce the incidence and adverse consequences.

# List Of Abbreviations

BMI: Body mass index

# Declarations

### Ethics approval and consent to participate

All study procedures were approved by the Ethic Committee of Shanxi Medical

University.

### Consent for publication

Not applicable.

### Availability of data and materials

Please contact author for data requests.

### Competing interests

The authors declare no competing interests.


### Funding

This work was supported by the National Natural Science Foundation of China (grant number: 71874104), China Postdoctoral Science Foundation Project (grant number: 2019M652133), Philosophy and Social Sciences of Guangdong College for the project of “Public Health Policy Research and Evaluation” Key Laboratory (grant number:2015WSYS0010), Theoretical Research Project of Aging Policy (grant number:2020029).

The sponsor had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

### Authors' contributions

Chichen Zhang and Shujuan Xiao conceived of the study and drafted the manuscript. Lei Shi did the statistical analysis. Xiao Zheng and Yaqing Xue helped revise the manuscript. Jiachi Zhang and Jiayu Zou checked the grammar of the manuscript. Fang Dong helped with formatting of this manuscript. All authors have seen and approved the final manuscript.

### Acknowledgements

We are extremely grateful to all the members who took part in this study, the elderly for their help in recruiting them, and investigators for their help in collecting data.

# References

1. Global, regional, and national under-5 mortality, adult mortality, age-specific mortality, and life expectancy, 1970-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017, 390:1084-1150.
2. Foo KM, Sundram M, Legido-Quigley H: Facilitators and barriers of managing patients with multiple chronic conditions in the community: a qualitative study. BMC Public Health 2020, 20:273.
3. Pruchno RA, Wilson-Genderson M, Heid AR: Multiple Chronic Condition Combinations and Depression in Community-Dwelling Older Adults. J Gerontol A Biol Sci Med Sci 2016, 71:910-915.
4. Hopman P, Schellevis FG, Rijken M: Health-related needs of people with multiple chronic diseases: differences and underlying factors. Qual Life Res 2016, 25:651-660.
5. Islas-Granillo H, Medina-Solís CE, de Lourdes Márquez-Corona M, de la Rosa-Santillana R, Fernández-Barrera M, Villalobos-Rodelo JJ, Hernández-Martínez CT, de Jesús Navarrete-Hernández J, Mendoza-Rodríguez M: Prevalence of multimorbidity in subjects aged ≥60 years in a developing country. Clin Interv Aging 2018, 13:1129-1133.
6. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, Meinow B, Fratiglioni L: Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev 2011, 10:430-439.
7. Skinner HG, Coffey R, Jones J, Heslin KC, Moy E: The effects of multiple chronic conditions on hospitalization costs and utilization for ambulatory care sensitive conditions in the United States: a nationally representative cross-sectional study. BMC Health Serv Res 2016, 16:77.
8. Kastner M, Cardoso R, Lai Y, Treister V, Hamid JS, Hayden L, Wong G, Ivers NM, Liu B, Marr S, et al: Effectiveness of interventions for managing multiple high-burden chronic diseases in older adults: a systematic review and meta-analysis. CMAJ 2018, 190:E1004-E1012.
9. Ioakeim-Skoufa I, Poblador-Plou B, Carmona-Pírez J, Díez-Manglano J, Navickas R, Gimeno-Feliu LA, González-Rubio F, Jureviciene E, Dambrauskas L, Prados-Torres A, Gimeno-Miguel A: Multimorbidity Patterns in the General Population: Results from the EpiChron Cohort Study. Int J Environ Res Public Health 2020, 17.

10. Fried TR, O'Leary J, Towle V, Goldstein MK, Trentelange M, Martin DK: The effects of comorbidity on the benefits and harms of treatment for chronic disease: a systematic review. PLoS One 2014, 9:e112593.
11. Khan A, Uddin S, Srinivasan U: Comorbidity network for chronic disease: A novel approach to understand type 2 diabetes progression. Int J Med Inform 2018, 115:1-9.
12. van Oostrom SH, Picavet HS, van Gelder BM, Lemmens LC, Hoeymans N, van Dijk CE, Verheij RA, Schellevis FG, Baan CA: Multimorbidity and comorbidity in the Dutch population - data from general practices. BMC Public Health 2012, 12:715.
13. Zulman DM, Slightam CA, Brandt K, Lewis ET, Asch SM, Shaw JG: "They are interrelated, one feeds off the other": A taxonomy of perceived disease interactions derived from patients with multiple chronic conditions. Patient Educ Couns 2020, 103:1027-1032.
14. Jiang X, Morgenstern LB, Cigolle CT, Claflin ES, Lisabeth LD: Multiple Chronic Conditions and Functional Outcome after Ischemic Stroke: A Systematic Review and Meta-Analysis. Neuroepidemiology 2020, 54:205-213.
15. Mokraoui NM, Haggerty J, Almirall J, Fortin M: Prevalence of self-reported multimorbidity in the general population and in primary care practices: a cross-sectional study. BMC Res Notes 2016, 9:314.
16. van Oostrom SH, Picavet HS, de Bruin SR, Stirbu I, Korevaar JC, Schellevis FG, Baan CA: Multimorbidity of chronic diseases and health care utilization in general practice. BMC Fam Pract 2014, 15:61.
17. Walker V, Perret-Guillaume C, Kesse-Guyot E, Agrinier N, Hercberg S, Galan P, Assmann KE, Briancon S, Rotonda C: Effect of Multimorbidity on Health-Related Quality of Life in Adults Aged 55 Years or Older: Results from the SU.VI.MAX 2 Cohort. PLoS One 2016, 11:e0169282.
18. Garin N, Koyanagi A, Chatterji S, Tyrovolas S, Olaya B, Leonardi M, Lara E, Koskinen S, Tobiasz-Adamczyk B, Ayuso-Mateos JL, Haro JM: Global Multimorbidity Patterns: A Cross-Sectional, Population-Based, Multi-Country Study. J Gerontol A Biol Sci Med Sci 2016, 71:205-214.
19. Bauer UE, Briss PA, Goodman RA, Bowman BA: Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 2014, 384:45-52.
20. Vallee A, Safar ME, Blacher J: Application of a decision tree to establish factors associated with a nomogram of aortic stiffness. J Clin Hypertens (Greenwich) 2019, 21:1484-1492.
21. Hsieh PC, Cheng CF, Wu CW, Tzeng IS, Kuo CY, Hsu PS, Lee CT, Yu MC, Lan CC: Combination of Acupoints in Treating Patients with Chronic Obstructive Pulmonary Disease: An Apriori Algorithm-Based Association Rule Analysis. Evid Based Complement Alternat Med 2020, 2020:8165296.
22. Hong J, Tamakloe R, Park D: Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. Accid Anal Prev 2020, 142:105497.
23. Luo X, Yan X, Luo M, Liu Y, Guo J, Wang H, Shi J, Sun Z: [Association rule regarding chronic disease-relevant risk factors for the adults in Haidian District]. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2017,

42:570-574.

24. Krist AH, O'Loughlin K, Woolf SH, Sabo RT, Hinesley J, Kuzel AJ, Rybarczyk BD, Kashiri PL, Brooks EM, Glasgow RE, et al: Enhanced care planning and clinical-community linkages versus usual care to address basic needs of patients with multiple chronic conditions: a clinician-level randomized controlled trial. Trials 2020, 21:517.
25. Iovino P, De Maria M, Matarese M, Vellone E, Ausili D, Riegel B: Depression and self-care in older adults with multiple chronic conditions: A multivariate analysis. J Adv Nurs 2020, 76:1668-1678.
26. Thomas-Purcell KB, Jacobs RJ, Seidman TL, Acevedo A, Waldrop-Valverde D, Ownby RL: A mixed analysis approach to elucidate the multiple chronic condition experience of English- and Spanish-speaking older adults. Clin Interv Aging 2019, 14:407-418.
27. Raterman HG, Lems WF: Pharmacological Management of Osteoporosis in Rheumatoid Arthritis Patients: A Review of the Literature and Practical Guide. Drugs Aging 2019, 36:1061-1072.
28. Adams ML, Grandpre J, Katz DL, Shenson D: The impact of key modifiable risk factors on leading chronic conditions. Prev Med 2019, 120:113-118.
29. Carroll JC, Campbell-Scherer D, Permaul JA, Myers J, Manca DP, Meaney C, Moineddin R, Grunfeld E: Assessing family history of chronic disease in primary care: Prevalence, documentation, and appropriate screening. Can Fam Physician 2017, 63:e58-e67.
30. Welch BM, Dere W, Schiffman JD: Family health history: the case for better tools. Jama 2015, 313:1711-1712.
31. Nie P, Ding L, Sousa-Poza A, Alfonso Leon A, Xue H, Jia P, Wang L, Diaz Sanchez ME, Wang Y: Socioeconomic position and the health gradient in Cuba: dimensions and mechanisms. BMC Public Health 2020, 20:866.
32. Zhang C, Zhu R, Lu J, Xue Y, Hou L, Li M, Zheng X, Yang T, Zheng J: Health promoting lifestyles and influencing factors among empty nesters and non-empty nesters in Taiyuan, China: a cross-sectional study. Health Qual Life Outcomes 2018, 16:103.
33. Knox J, Scodes J, Witkiewitz K, Kranzler HR, Mann K, O'Malley SS, Wall M, Anton R, Hasin DS: Reduction in World Health Organization Risk Drinking Levels and Cardiovascular Disease. Alcohol Clin Exp Res 2020, 44:1625-1635.
34. Liu Y, Wang K, Maisonet M, Wang L, Zheng S: Associations of lifestyle factors (smoking, alcohol consumption, diet and physical activity) with type 2 diabetes among American adults from National Health and Nutrition Examination Survey (NHANES) 2005-2014. J Diabetes 2017, 9:846-854.
35. Pinto TJP, Mendonça EP, Bloch KV, Cunha GM, Coutinho ESF: Aggregation of behavioral risk factors to noncommunicable chronic diseases: A national school-based study with Brazilian adolescents. Prev Med 2020, 137:106128.

# Tables

Table 1
Prevalence and gender differences of the top 10 chronic diseases

| Chronic disease | N (%) | Male (n/%) | Female (n/%) | $\chi^2$ | P |
|---|---|---|---|---|---|
| Hypertension | 957(29.45) | 422(44.10) | 535(55.90) | 3.459 | 0.063 |
| Diabetes | 352(10.83) | 159(45.17) | 193(54.83) | 0.331 | 0.565 |
| Rheumatoid or rheumatoid arthritis | 334(10.28) | 126(37.72) | 208(62.28) | 11.824 | 0.001 |
| Hearing impairment | 263(8.09) | 133(50.57) | 130(49.43) | 1.799 | 0.180 |
| Digestive system diseases | 211(6.49) | 95(45.02) | 116(54.98) | 0.230 | 0.632 |
| Osteoporosis | 209(6.43) | 71(33.97) | 138(66.03) | 14.350 | <0.001 |
| Coronary heart disease | 204(6.28) | 79(38.73) | 125(61.27) | 5.445 | 0.020 |
| Eye disease | 182(5.60) | 59(32.42) | 123(67.58) | 15.617 | <0.001 |
| Respiratory diseases | 145(4.46) | 81(55.86) | 64(44.14) | 0.040 | 0.842 |
| Stroke | 130(4.00) | 77(59.23) | 53(40.77) | 8.660 | 0.003 |

Table 2
Analysis results of association rules for chronic diseases

| Consequent | Antecedent | Support (%) | Confidence (%) | Lift |
|---|---|---|---|---|
| Hypertension | Atherosclerosis | 3.02% | 60.20% | 2.05 |
| Hypertension | Coronary heart disease | 6.28% | 53.92% | 1.83 |
| Hypertension | Diabetes | 10.83% | 52.84% | 1.79 |
| Hypertension | Stroke | 4.00% | 49.23% | 1.67 |
| Hypertension | Eye disease | 5.60% | 46.15% | 1.57 |
| Hypertension | Osteoporosis | 6.43% | 39.71% | 1.35 |
| Hypertension | Hearing impairment | 8.09% | 39.16% | 1.33 |
| Hypertension | Rheumatoid or rheumatoid arthritis | 10.28% | 34.43% | 1.17 |
| Rheumatoid or rheumatoid arthritis | Osteoporosis | 6.43% | 33.49% | 3.26 |
| Coronary heart disease | Atherosclerosis | 3.02% | 32.65% | 5.20 |

# Figures

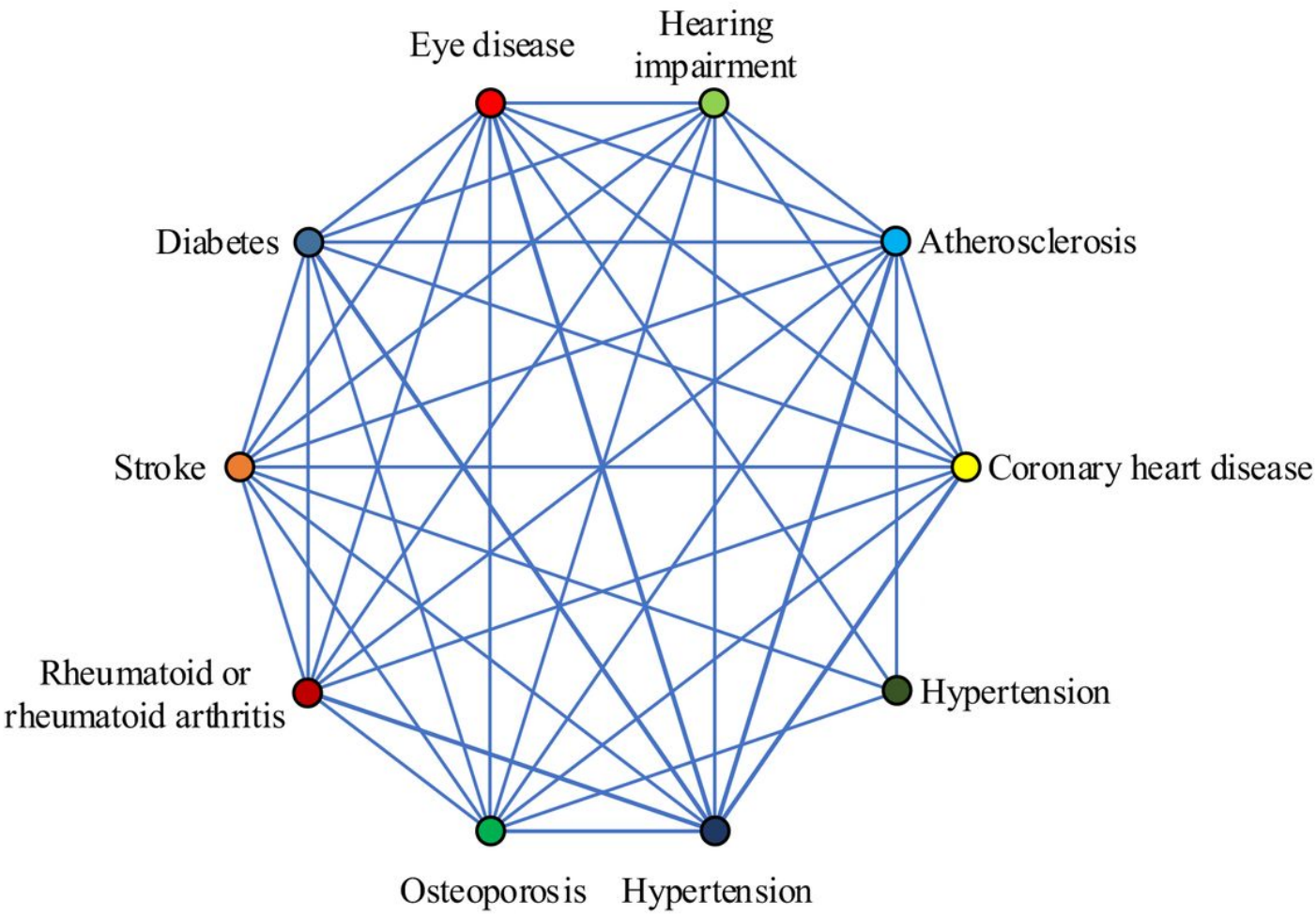
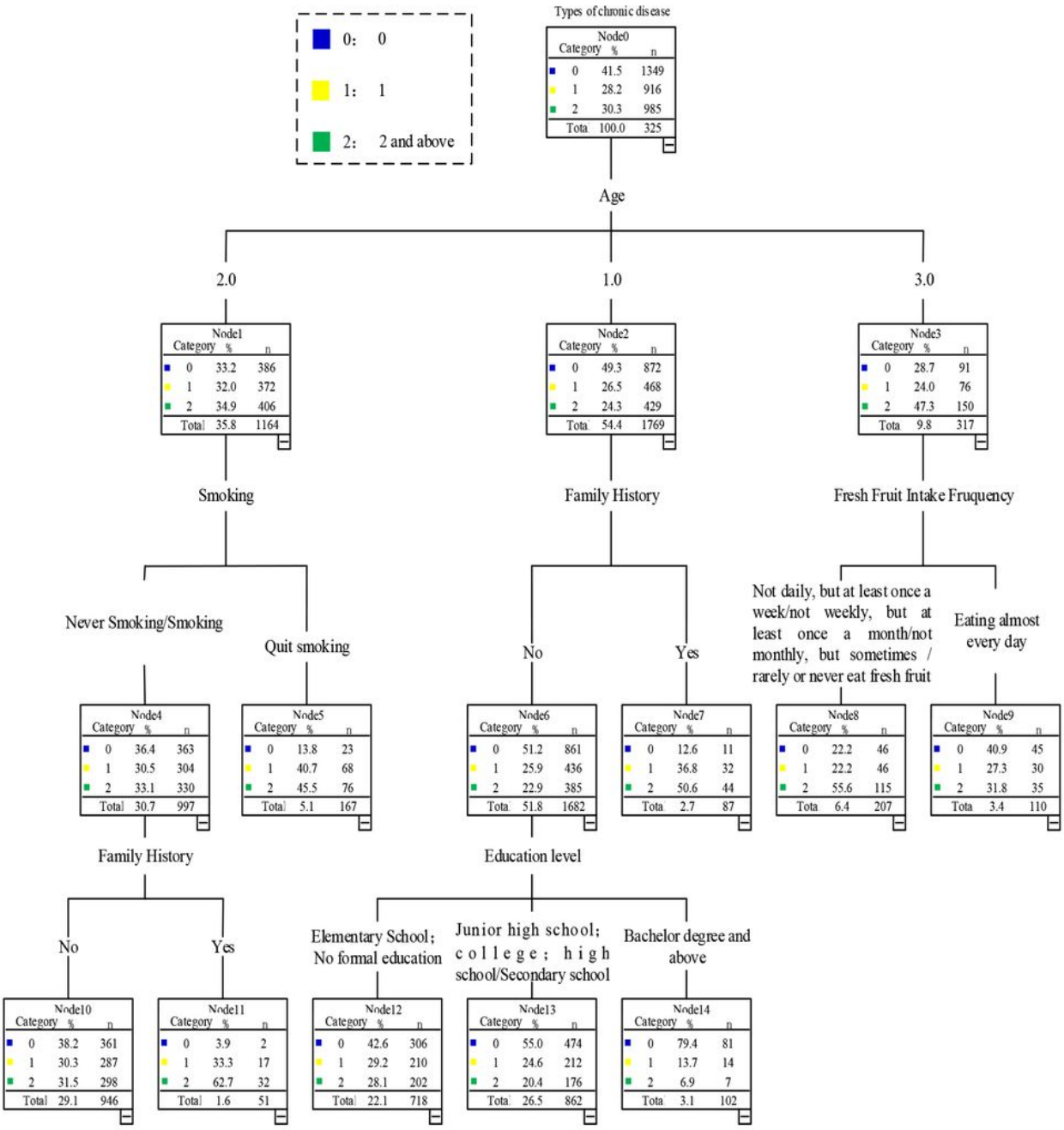



**Figure 1**

Related chronic disease network diagram




## Figure 2

Decision tree model of related influencing factors for multiple chronic diseases