DSI DATA SCIENCE AND INFORMETRICS
VOL.4, NO.1, FEB. 2024

RESEARCH ARTICLE

The prevalence and evolution of binary multimorbidity pattern in middle-aged and elderly population in China, from an insight of health metrics

Yu Wang, Xiaojun Hu*

Department of Medical Informatics, Zhejiang University School of Medicine

ABSTRACT

The prevention and management of chronic diseases multimorbidity has become a public health issue that needs to be focused on and addressed in the process of healthy aging. This paper analyzes the evolution of multimorbidity patterns in the middle-aged and elderly population aged ≥45 years old from the perspective of health metrics, based on CHARLS 2011-2020 data, using an association rule CARMA algorithm, in combination with a visual network presentation for the evolution of multimorbidity patterns and that in different classified populations. The results show that from 2011 to 2020, the prevalence of multimorbidity showed an overall increasing trend, with a multimorbidity prevalence of 37.41% in 2011 and 57.47% in 2020, and the top 3 chronic diseases with the highest prevalence of multimorbidity are arthritis or rheumatism, hypertension and digestive diseases. In addition, in five periods we observed, the prevalence of multimorbidity in all age groups showed an overall increasing trend; the prevalence of binary multimorbidity pattern in females is higher than that in males; people living in rural is of higher prevalence of multimorbidity than those living in the urban in most periods; farmers is of higher prevalence of multimorbidity than those with other occupations in most periods. The results suggest that effective prevention and management of multimorbidity is an urgent task in China, and individualized management plans should be formulated according to different age groups, genders, places of residence and personal habits.

KEYWORDS

Multimorbidity; Middle-aged and elderly; Big data; Health metrics

1 Introduction

With the global aging process accelerating, the prevalence of chronic non-communicable diseases is increasing. In 2016, the UK National Institute for Health and Care Excellence published a guidance stating that multimorbidity is usually defined as when a person has two or more long-term health conditions (Kernick et al., 2017). Mohamud et al. (2023) found that in the United States, the prevalence of multimorbidity increased in all race/ethnic groups over the past 20 years. Zhao et al. (2023) and Li et al. (2023) showed that the prevalence of multimorbidity increased significantly with age, and people over 60 years old were about 20 times more likely to suffer from multiple chronic diseases than those between 18 and 29 years old. The problem of multimorbidity is not only noticeable in high-income countries, but also becoming increasingly important in low- and middle-income countries (Asogwa et al., 2022). People with multimorbidity experienced severe psychological distress, their psychological well-being declined significantly as the number of

Corresponding author: xjhu@zju.edu.cn

diseases increased (Sasseville et al., 2019), resulting in increased risks of disability and death associated with multimorbidity (Rivera-Almaraz et al., 2018). Many researches also claimed that multimorbidity results in an increase in healthcare utilisation and medical expenses, physical impairments and decreased quality of life (Pati et al., 2015; Liu et al., 2023; Ni et al., 2023).

Multimorbidity patterns generally refer to combinations of chronic diseases, which are statistically interrelated to varying degrees as indicated by Ingmar Schäfer et al. (2010). Silva et al. (2023) found that multimorbidity patterns appear to share common pathophysiological mechanisms between chronic diseases groups, which interact and overlap in multimorbidity populations. Recognizing patterns of multimorbidity could provide clues for preventing and treating diseases and improving prognosis (Fan et al., 2022), yet, most of the present research on this topic is cross-sectional (Kudesia et al., 2021; Rajoo et al., 2021).

To catch up the actual direction of multimorbidity in China, in this contribution, we will conduct an investigation based on 5-periods data from the China Health and Retirement Longitudinal Study (CHARLS) in 2011, 2013, 2015, 2018 and 2020 waves, determine the prevalence and temporal trend of multimorbidity among people aged 45 years and above, and explore the network evolution of multimorbidity patterns in different groups from the health metrics perspective.

2 Materials and methods

2.1 Data source

Data was provided by CHARLS, a vital population aging research database in China, to gather a collection of high-quality microdata that represents households and individuals of middle-aged and older persons in China who are aged 45 years and above, aiming to encourage interdisciplinary research on aging, and to provide a more scientific basis for the formulation and improvement of China's relevant policies. The CHARLS study adopts a baseline sample tracking interview research model. With the continuation of the survey from year to year, CHARLS is able to record the development and changes of the same group of scientifically abstracted respondents who are representative of the general population of China, and the results of the interviews have immediate scientific research value as well as multidimensional and interdisciplinary historical value. The national baseline survey of CHARLS was conducted in 2011 and has been followed up every two to three years since. The survey project used multistage sampling and data collection was conducted through direct interviewing, covering all regions of China and different population groups. The regular waves of the CHARLS database contain information on individuals' demographic backgrounds, family information, health status and functioning, cognition, health care and insurance, and more.

This study used partial data on demographic backgrounds and data on 14 chronic diseases in health status and functioning of the middle-aged and elderly individuals aged ≥45 years in the waves of 2011, 2013, 2015, 2018 and 2020 of CHARLS to conduct a longitudinal study of multimorbidity patterns. A total of 93,152 individuals were included in the five waves.

2.2 Variable selection and definition

In this study, data on age, sex, place of residence, occupations, smoking, alcohol consumption and condition of 14 chronic diseases were extracted from the demographic backgrounds and health status and functioning modules. The 14 chronic diseases consist of hypertension, dyslipidemia, diabetes or high blood sugar, cancer or malignant tumor (excluding minor skin cancers), chronic lung diseases (excluding tumors, or cancer), liver diseases (except fatty liver, tumors, and cancer), heart diseases, stroke, kidney diseases (except tumor or cancer), stomach or other digestive diseases (except tumor or cancer), emotional, nervous, or psychiatric problems, memory-related disease, arthritis or rheumatism, and asthma. Multimorbidity was defined in this study as individuals with two or more chronic conditions.

2.3 Statistical analysis

Descriptive analysis was used to describe the characteristics of all participants, with continuous variables expressed as means and standard deviations, and categorical variables expressed as

counts and percentages. Chi-squared test was used to compare the prevalence of multimorbidity in subjects classified by age, sex, place of residence, occupation, smoking, and alcohol consumption. $P < 0.05$ was considered statistically significant. All of the above analyses were performed with SAS version 9.4.

2.4 Association rules and algorithm

Association rules were proposed by Rakesh Agrawal and Ramakrishnan Srikant in 1993 to describe the correlation between two or more things (Agrawal et al., 1993), which predicts other things by one or more things, and can be used to obtain connections between valuable data from a large amount of data. The association rule is usually expressed as $X \rightarrow Y$, where X and Y are disjoint sets of terms, X is called antecedent and Y is called consequent. The support of this rule is the support of $X \cup Y$. The confidence of this rule is the fraction of all transactions containing X that also contain Y. *i.e.* the support of $X \cup Y$ divided by the support of X. The discovery process of association rules usually consists of two steps: the discovery of frequent itemsets and the generation of association rules. First, frequent itemsets, *i.e.*, itemsets that occur more frequently than a preset threshold, are found by calculating the support. Then, based on the frequent itemsets, the confidence level is computed and the association rules that satisfy the confidence level requirement are generated. CARMA algorithm is a classical technique used for analyzing association rules, which was proposed by Hidber et al. (1999). The approach utilizes an iterative layer-by-layer method to identify the itemset relationships in a database and to derive the rules formed from such relationships. The principle of the CARMA algorithm is that the more times two items appear in pairs in the transaction data, the more relevant the two items are, and the two items are strongly correlated. CARMA nodes provide construction settings for rule support (support for preconditions and results) rather than for precondition support. CARMA algorithm is easy to understand in terms of the principle of the algorithm and concise and convenient in terms of its implementation.

3 Results

3.1 Prevalence of included diseases

The data of 16,816, 18,244, 19,500, 19,435 and 19,157 people who were included in the follow-up surveys in 2011, 2013, 2015, 2018 and 2020, respectively, were analyzed, and the basic characteristics of the population are shown in Table 1. For the prevalence of the diseases included in the study, the number of people who suffered from at least one disease per year was 11,271 cases (67.03%), 10,037 cases (55.02%), 11,761 cases (60.31%), 14,080 cases (72.47%) and 15,464 cases (80.72%) per year, decreasing and then increasing with the change of time, and the overall prevalence rate is on the rise; the prevalence rates of the 14 chronic diseases are decreasing and then increasing with the change of survey time, and the top three chronic diseases in prevalence rate are arthritis or rheumatism, hypertension, and digestive diseases, with the prevalence rates of 27.38%-38.42%, 19.74%-40.15%, and 18.38%-31.43%, respectively.

Table 1 Basic characteristics of participants in the cross-sectional analysis [*n* (%)]

Characteristics	2011 (*N*=16,816)	2013 (*N*=18,244)	2015 (*N*=19,500)	2018 (*N*=19,435)	2020 (*N*=19,157)
Age (year, mean±SD)	59.38±9.88	60.30±10.02	60.46±10.17	62.12±10.18	61.82±9.85
Age groups					
45-54 years old	5,921 (35.21)	5,925 (32.48)	6,665 (34.18)	5,419 (27.88)	5,312 (27.73)
55-64 years old	6,282 (37.36)	6,698 (36.71)	6,460 (33.13)	6,386 (32.86)	6,524 (34.06)
65-74 years old	3,137 (18.65)	3,745 (20.53)	4,383 (22.48)	5,126 (26.38)	5,119 (26.72)
75-84 years old	1,279 (7.61)	1,595 (8.74)	1,690 (8.67)	2,081 (10.71)	1,860 (9.71)
85- years old	197 (1.17)	281 (1.54)	302 (1.55)	423 (2.18)	342 (1.79)
Sex					
Male	8,137 (48.39)	8,819 (48.34)	9,446 (48.44)	9,220 (47.44)	9,060 (47.29)

Female	8,679 (51.61)	9,425 (51.66)	10,054 (51.56)	10,215 (52.56)	10,097 (52.71)
Residence					
Rural	13,098 (77.89)	14,164 (77.64)	14,709 (75.43)	14,529 (74.76)	12,147 (63.41)
Urban	3,718 (22.11)	4,080 (22.36)	4,791 (24.57)	4,906 (25.24)	7,010 (36.59)
Occupations					
Employed	4,849 (28.84)	4,578 (25.09)	5,612 (28.78)	11,132 (57.28)	4,680 (24.43)
Self-employed	1,255 (7.46)	1,568 (8.59)	1,722 (8.83)	1,153 (5.93)	1,064 (5.55)
Unpaid Family Business	422 (2.51)	495 (2.71)	608 (3.12)	703 (3.62)	366 (1.91)
Farming	9,635(57.30)	7,659(41.98)	8,202(42.06)	6,359(32.72)	6,154(32.12)
Others	655 (3.90)	3,944 (21.62)	3,356 (17.21)	88 (0.45)	6,893 (35.98)
Smoke					
Yes	6,641 (39.49)	7,169 (39.30)	8,041 (41.24)	7,994 (41.13)	7,247 (37.83)
No	10,175 (60.51)	11,075 (60.70)	11,459 (58.76)	11,441 (58.87)	11,910 (62.17)
Drink					
Yes	5,525 (32.86)	6,234 (34.17)	6,850 (35.13)	6,518 (33.54)	6,860 (35.81)
No	11,291 (67.14)	12,010 (65.83)	12,650 (64.87)	12,917 (66.46)	12,297 (64.19)
Diseases prevalence					
Hypertension	4,092 (24.33)	3,601 (19.74)	4122 (21.14)	5524 (28.42)	7692 (40.15)
Dyslipidemia	1,531 (9.10)	1,346 (7.38)	1727 (8.86)	3286 (16.91)	5072 (26.48)
Diabetes	946 (5.63)	835 (4.58)	1015 (5.21)	1865 (9.60)	2844 (14.85)
Cancer	160 (0.95)	141 (0.77)	192 (0.98)	400 (2.06)	487 (2.54)
Chronic Lung Diseases	1,705 (10.14)	1,501 (8.23)	1916 (9.83)	2433 (12.52)	2739 (14.30)
Liver Diseases	652 (3.88)	587 (3.22)	769 (3.94)	1235 (6.35)	1384 (7.22)
Heart Diseases	1,998 (11.88)	1,764 (9.67)	2112 (10.83)	3030 (15.59)	3995 (20.85)
Stroke	383 (2.28)	333 (1.83)	416 (2.13)	1260 (6.48)	1378 (7.19)
Kidney Diseases	1,061 (6.31)	937 (5.14)	1201 (6.16)	1711 (8.80)	656 (3.42)
Digestive Diseases	3,718 (22.11)	3,353 (18.38)	4395 (22.54)	5093 (26.21)	6021 (31.43)
Psychiatric Problems	225 (1.34)	203 (1.11)	258 (1.32)	437 (2.25)	577 (3.01)
Memory-related Diseases	250 (1.49)	209 (1.15)	287 (1.47)	641 (3.30)	1168 (6.10)
Arthritis or Rheumatism	5,569 (33.12)	4,996 (27.38)	6416 (32.90)	6734 (34.65)	7361 (38.42)
Asthma	610 (3.63)	529 (2.90)	727 (3.73)	949 (4.88)	1167 (6.09)

3.2 Trends in the prevalence of multimorbidity

As shown in Table 2, 6,291/16,816 (37.41%) respondents in 2011, 5,587/18,244(30.62%), 7,089/19,500 (36.35%), 9,169/19,435 (47.18%), 11,009/19,157 (57.47%) in 2013, 2015, 2018 and 2020 were multimorbidity, with an overall increasing trend, and binary multimorbidity (had two chronic diseases at the same time) was the most popular, with the percentage in multimorbidity ranging from 35.82% to 51.08%. In terms of gender, the proportion of multimorbidity in female was higher than that in male. In terms of place of residence, the prevalence of multimorbidity in both urban and rural areas showed a general upward trend from 2011 to 2020, with urban areas having a higher prevalence of multimorbidity than rural areas in 2011 and 2020, and rural areas having a higher prevalence of multimorbidity than urban areas in 2013, 2015 and 2018. In terms of age groups, the prevalence of multimorbidity in all age groups showed an overall increasing trend from 2011 to 2020. In 2011, the 45-54 years old group had the lowest prevalence of multimorbidity, followed by the 85- years old group. In 2020, the 45-54 years old group had the lowest prevalence of multimorbidity, followed by the 55-64 years old group. In terms of occupations, the prevalence of multimorbidity in employed, self-employed, unpaid family businesses, farming, and other occupational groups showed an overall increasing trend. And farmers had the highest prevalence of multimorbidity in 2011, 2018, and 2020.

Table 2 Prevalence of multimorbidity among included subjects, 2011-2020 [*n* (%)]

Populations	2011 (*N*=16,816)	2013 (*N*=18,244)	2015 (*N*=19,500)	2018 (*N*=19,435)	2020 (*N*=19,157)
Age groups					
45-54 years old	1,702 (28.75)	1,190 (20.08)	1,482 (22.24)	1,519 (28.03)	2,296 (43.22)
55-64 years old	2,464 (39.22)	2,144 (32.01)	2,500 (38.70)	3,049 (47.75)	3,768 (57.76)
65-74 years old	1,475 (47.02)	1,540 (41.12)	2,174 (49.60)	3,065 (59.79)	3,456 (67.51)
75-84 years old	577 (45.11)	617 (38.68)	807 (47.75)	1,301 (62.52)	1,285 (69.09)
85- years old	73 (37.06)	96 (34.16)	126 (41.72)	235 (55.56)	204 (59.65)
Sex					
Male	2,859 (35.14)	2,508 (28.44)	3,179 (33.65)	4,175 (45.28)	4,979 (54.96)
Female	3,432 (39.54)	3,079 (32.67)	3,910 (38.89)	4,994 (48.89)	6,030 (59.72)
Residence					
Rural	4,825 (36.84)	4,424 (31.23)	5,488 (37.31)	6,978 (48.03)	6,975 (57.42)
Urban	1,466 (39.43)	1,163 (28.50)	1,601 (33.42)	2,191 (44.66)	4,034 (57.55)
Occupations					
Employed	1,674 (34.52)	908 (19.83)	1,412 (25.16)	4,660 (41.86)	2,044 (43.68)
Self-employed	379 (30.20)	374 (23.85)	456 (26.48)	535 (46.40)	518 (48.68)
Unpaid Family Business	155 (36.73)	160 (32.32)	232 (38.16)	301 (42.82)	208 (56.83)
Farming	3,836(39.81)	2,415(31.53)	3,312(40.38)	3,647(57.35)	3,603(58.55)
Others	247 (37.71)	1,730 (43.86)	1,677 (49.97)	26 (29.55)	4,636 (67.26)
Smoke					
Yes	2,453 (36.94)	2,157 (30.09)	2,831 (35.21)	3,757 (47.00)	4,380 (60.44)
No	3,838 (37.72)	3,430 (30.97)	4,258 (37.16)	5,412 (47.30)	6,629 (55.66)
Drink					
Yes	1,812 (32.80)	1,593 (25.55)	2,129 (31.08)	2,684 (41.18)	3,561 (51.91)
No	4,479 (39.67)	3,994 (33.26)	4,960 (39.21)	6,485 (50.21)	7,448 (60.57)
Total multimorbidity	6,291 (37.41)	5,587 (30.62)	7,089 (36.35)	9,169 (47.18)	11,009 (57.47)
Binary multimorbidity*	3,201 (50.88)	2,854 (51.08)	3,432 (48.41)	3,734 (40.72)	3,943 (35.82)

Note: *Figures in parentheses is the percentage of the binary multimorbidity in all subjects with multimorbidity.

3.3 Evolution of binary multimorbidity patterns

Setting the minimum rule support to 1%, the minimum rule confidence to 10% and the maximum rule size to 2 can generate binary association rules for disease combinations that conform to the aforementioned criteria. There are 74, 74, 75, 84 and 85 rules, respectively. Table 3 illustrates the top 10 strongest multimorbidity association results for the five sampling years.

Figures 1-4 show co-disease networks of binary multimorbidity patterns for middle-aged and elderly people from 2011 to 2020, including males and females, residents in urban and rural, people in the 45-54 and 55-64 years old groups that we are focusing on. There are 14 nodes in the network, different nodes represent different types of chronic diseases, the size of the node is proportional to the prevalence of chronic diseases, the larger the node is, the higher the prevalence; the connecting edges between the nodes represent the simultaneous suffering of these two types of chronic diseases; the thickness of the connecting edges responds to the co-prevalence frequency of these two types of chronic diseases; the thicker the connecting edge is, the higher the co-prevalence frequency. Four nodes in the middle of the network graph have the highest prevalence rate and the most cases of multimorbidity. Changes in the main multimorbidity patterns under different conditions can be observed from them.

Table 3 The prevalence rate of the top 10 binary multimorbidity patterns

Year	Order	Disease-associated rules	Support (%)
2011	1	Digestive Diseases + Arthritis or Rheumatism	49.41
	2	Kidney Diseases + Arthritis or Rheumatism	49.41
	3	Diabetes + Hypertension	36.31
	4	Dyslipidemia + Hypertension	36.31
	5	Heart Diseases + Hypertension	36.31
	6	Chronic Lung Diseases + Digestive Diseases	32.99
	7	Heart Diseases + Digestive Diseases	32.99
	8	Kidney Diseases + Digestive Diseases	32.99
	9	Chronic Lung Diseases + Heart Diseases	17.73
	10	Diabetes + Heart Diseases	17.73
2013	1	Kidney Diseases + Arthritis or Rheumatism	49.79
	2	Digestive Diseases + Arthritis or Rheumatism	49.79
	3	Chronic Lung Diseases + Arthritis or Rheumatism	49.79
	4	Dyslipidemia + Hypertension	35.88
	5	Diabetes + Hypertension	35.88
	6	Heart Diseases + Hypertension	35.88
	7	Kidney Diseases + Digestive Diseases	33.41
	8	Chronic Lung Diseases + Digestive Diseases	33.41
	9	Heart Diseases + Digestive Diseases	33.41
	10	Dyslipidemia + Heart Diseases	17.58
2015	1	Kidney Diseases + Arthritis or Rheumatism	54.55
	2	Chronic Lung Diseases + Arthritis or Rheumatism	54.55
	3	Kidney Diseases + Digestive Diseases	37.37
	4	Heart Diseases + Digestive Diseases	37.37
	5	Chronic Lung Diseases + Digestive Diseases	37.37
	6	Dyslipidemia + Hypertension	35.05
	7	Diabetes + Hypertension	35.05
	8	Heart Diseases + Hypertension	35.05
	9	Dyslipidemia + Heart Diseases	17.96
	10	Diabetes + Heart Diseases	17.96
2018	1	Chronic Lung Diseases + Arthritis or Rheumatism	47.81
	2	Kidney Diseases + Arthritis or Rheumatism	47.81
	3	Heart Diseases + Arthritis or Rheumatism	47.81
	4	Stroke + Hypertension	39.22
	5	Diabetes + Hypertension	39.22
	6	Dyslipidemia + Hypertension	39.22
	7	Heart Diseases + Hypertension	39.22
	8	Kidney Diseases + Hypertension	39.22
	9	Liver Diseases + Digestive Diseases	36.16
	10	Chronic Lung Diseases + Digestive Diseases	36.16
2020	1	Hypertension + Stroke	49.83
	2	Hypertension + Diabetes	49.83
	3	Hypertension + Memory-related Diseases	49.83
	4	Hypertension + Asthma	49.83
	5	Hypertension + Kidney Diseases	49.83
	6	Hypertension + Psychiatric Problems	49.83
	7	Arthritis or Rheumatism + Asthma	47.69
	8	Arthritis or Rheumatism + Memory-related Diseases	47.69

9	Arthritis or Rheumatism + Liver Diseases	47.69
10	Arthritis or Rheumatism + Kidney Diseases	47.69

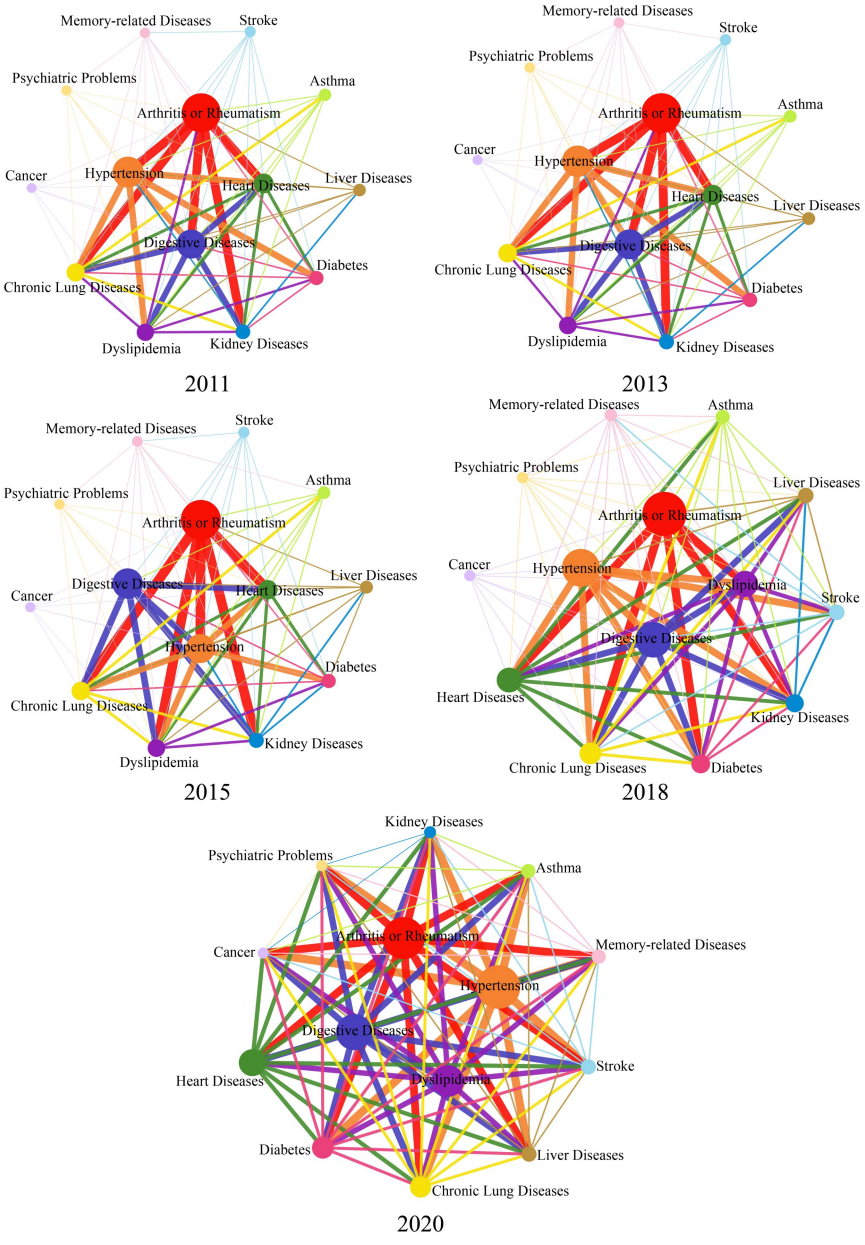
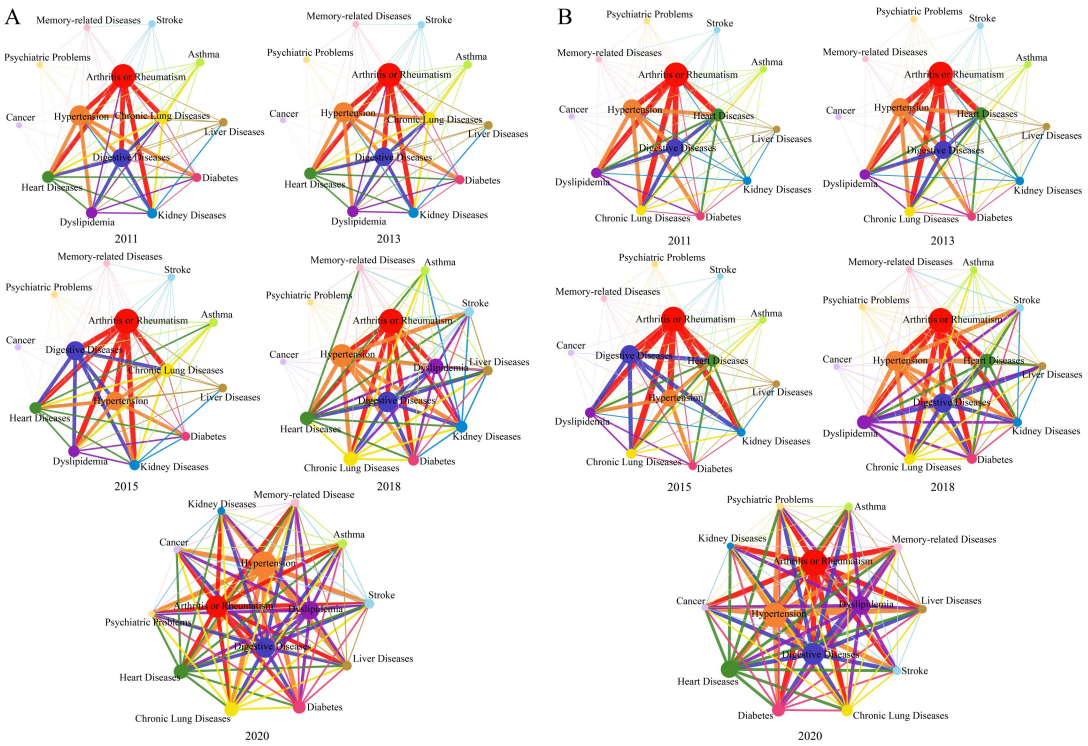
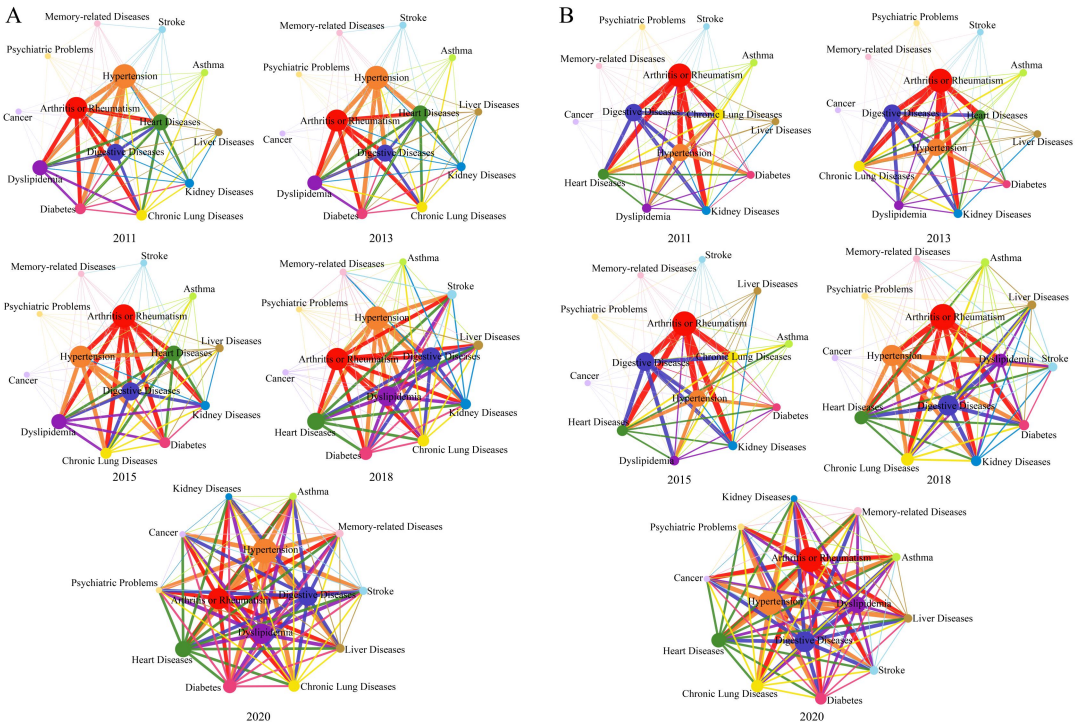


Figure 1 The co-disease network of multimorbidity patterns from 2011 to 2020



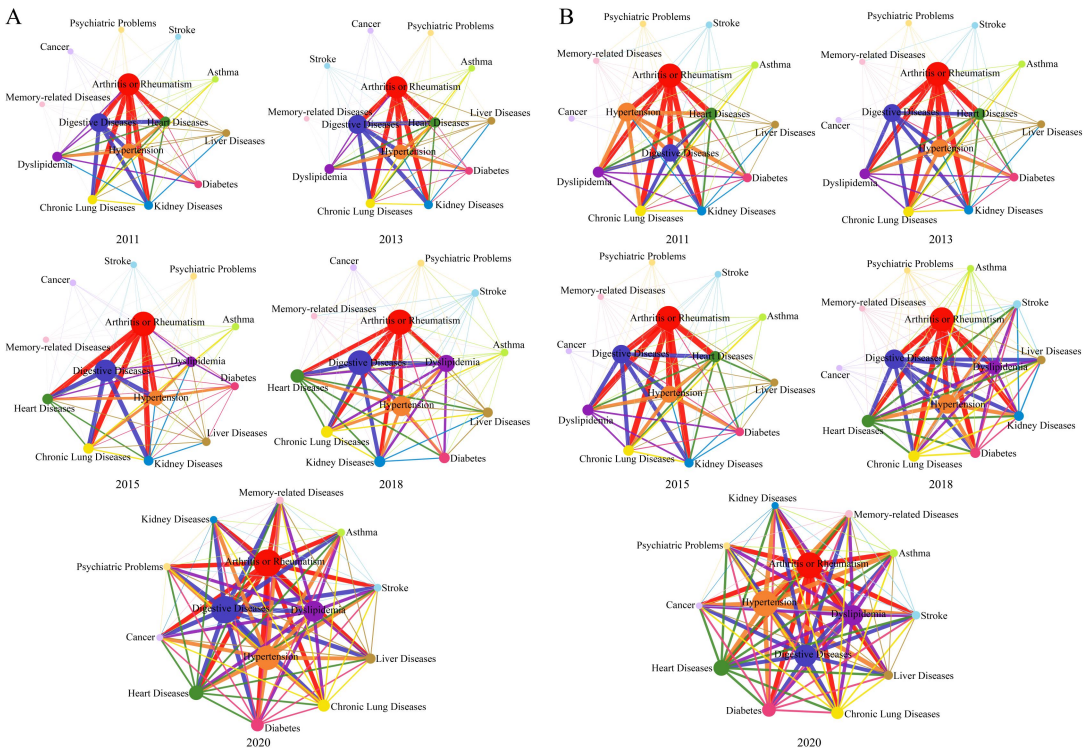
A: Male; B: Female

Figure 2 The co-disease network of multimorbidity patterns by gender from 2011 to 2020



A: Urban; B: Rural

Figure 3 The co-disease network of multimorbidity patterns by residence from 2011 to 2020



A: People between 45 to 54 years old; B: People between 55 to 64 years old

Figure 4 The co-disease network of multimorbidity patterns by age group from 2011 to 2020

Figure 1 shows that from 2011 to 2020, the nodes of the multimorbidity network changed significantly, with a pronounced rise in the multimorbidity pattern of dyslipidemia, an overall tendency for the probability of multimorbidity to increase, and an overall tendency for the density to increase. Figure 2 shows the network of multimorbidity patterns for males (Figure 2A) and females (Figure 2B). It can be seen that males and females have different center network nodes and different multimorbidity patterns, males have higher support for chronic lung diseases multimorbidity and females have higher support for heart diseases multimorbidity. Figure 3 shows the network of urban (Figure 3A) and rural (Figure 3B) multimorbidity patterns. Both urban and rural networks present varied from different years, yet, it can be seen that the density of the urban and rural multimorbidity pattern network has been increasing from 2011 to 2020, suggesting that the tie strength among disease-pairs are increased over time. In addition to the same multimorbidity pattern, support for the multimorbidity pattern for heart diseases was higher in urban areas. Figure 4 shows the multimorbidity pattern network for the age groups of 45-54 (Figure 4A) and 55-64 (Figure 4B) that we focused on exposing from 2011 to 2020. The 55- 64 years old group has an increasing density of multimorbidity pattern network, and the change in multimorbidity pattern network graph for the 45-54 years old group has been relatively stable. The multimorbidity pattern of dyslipidemia in the younger age group showed an increasing trend. The details of the network diagrams are shown in Table 4.

Table 4 The characteristics of multimorbidity patterns network graph from 2011 to 2020

Year	Populations	Average Degree	Average Weighted Degree	Graph Density	Average Clustering Coefficient	Average Path Length
2011	All	10.571	130.009	0.813	0.887	1.187
	Male	10.286	125.403	0.791	0.877	1.209
	Female	10.857	129.868	0.835	0.894	1.165

	Rural	10.286	124.112	0.791	0.878	1.209
	Urban	10.857	131.427	0.835	0.890	1.165
	45-54 years old group	9.571	106.476	0.736	0.836	1.264
	55-64 years old group	10.286	136.110	0.791	0.881	1.209
2013	All	10.429	129.202	0.802	0.884	1.198
	Male	10.429	125.317	0.802	0.884	1.198
	Female	10.286	128.201	0.791	0.883	1.209
	Rural	10.286	126.809	0.791	0.878	1.209
	Urban	10.857	128.333	0.835	0.890	1.165
	45-54 years old group	9.857	106.074	0.758	0.823	1.242
	55-64 years old group	10.000	131.694	0.769	0.870	1.231
2015	All	10.714	138.897	0.824	0.868	1.176
	Male	10.857	146.26	0.835	0.904	1.165
	Female	11.000	150.640	0.846	0.900	1.154
	Rural	11.000	140.772	0.846	0.900	1.154
	Urban	11.000	155.391	0.846	0.894	1.154
	45-54 years old group	10.571	115.476	0.813	0.866	1.187
	55-64 years old group	10.714	144.122	0.824	0.891	1.176
2018	All	12.000	193.657	0.923	0.935	1.077
	Male	12.286	201.062	0.945	0.957	1.055
	Female	12.143	211.274	0.934	0.953	1.066
	Rural	12.143	200.562	0.934	0.948	1.066
	Urban	12.000	210.739	0.923	0.937	1.077
	45-54 years old group	11.714	137.359	0.901	0.924	1.099
	55-64 years old group	12.143	195.610	0.934	0.953	1.066
2020	All	12.143	335.848	0.934	0.948	1.066
	Male	13.000	358.889	1.000	1.000	1.000
	Female	13.000	395.733	1.000	1.000	1.000
	Rural	12.714	362.294	0.978	0.980	1.022
	Urban	13.000	381.423	1.000	1.000	1.000
	45-54 years old group	13.000	326.076	1.000	1.000	1.000
	55-64 years old group	13.000	368.851	1.000	1.000	1.000

4 Discussion and conclusions

In this study, we applied nationally representative middle-aged and elderly health-related data from CHARLS 2011-2020 to analyze the multimorbidity prevalence of chronic diseases and their evolutionary trends in a longitudinal insight, identify multimorbidity patterns among Chinese people ≥45 years old. Instead of the traditional way in the area of management for chronic diseases, we employed a network approach to figure out the co-occurrence network for diseases-pairs of multimorbidity and their evolutionary characteristics over time, providing a new insight of visualizations in the research of health metrics.

Descriptive statistics of the included data can help us to quickly understand the basic situation

of the population with different characteristics and the prevalence of the 14 chronic diseases, and can help us to discover the general pattern and trend of the data from 2011 to 2020. The chi-square test is applicable to the comparison of two or more categorical data, and the reliability of the test results is high. By conducting the chi-square test on different characteristic populations and multimorbidity, we can quickly find out whether there is a significant correlation or independence between different characteristic populations and the prevalence of multiple diseases.

We found that the prevalence of multimorbidity ranged from 30.62% to 57.47%, which is similar to the results of other researches (Chen et al., 2018; Yao et al., 2020; Zhang et al., 2022). The prevalence of multimorbidity decreased and then increased over time, with an overall increasing trend, which may be partly related to the sampling response rate at the time of the survey, and partly due to the prevalence and detection measures of chronic diseases in the survey population. Second, although multimorbidity increased significantly from 2011-2020 in both males and females, the prevalence of multimorbidity in females is higher than that of males, which may be related to longer average life expectancy and poorer health status in females compared with males (Zhong et al., 2022). Third, among the five waves of data, in three of the waves the prevalence of multimorbidity in rural residents is higher than that of urban residents, which may be related to the rise in economic level and the enhancement of health literacy as well as lifestyle of both rural and urban residents. Moreover, the results about characteristics of multimorbidity networks for the lower older age groups revealed that the condition of multimorbidity in the 45-54 years old group was as complex and severe as in the 55-64 years old group, suggesting that multimorbidity in the lower age groups deserves more attention in the context of the trend towards an aging population.

Under the same association rule condition, the 14 common binary chronic disease multimorbidity patterns developed from 74 in 2011 to 85 in 2020. Among them, arthritis or rheumatism was the disease with the highest multimorbidity rate, followed by hypertension and digestive diseases, which may be associated with the high prevalence of these three diseases. It can be seen from the top 10 strong association rules that the multimorbidity combinations associated with arthritis or rheumatism, hypertension, and digestive diseases are changing from 2011 to 2020. As shown in Figures 2-4, the binary multimorbidity patterns in different populations have been changing and increasing from 2011 to 2020. There are also many differences in the prevalence of multimorbidity and the main multimorbidity patterns among different genders, age groups, and residences, which provide evidence for policy-making. On the one hand, the emphasis on diseases with high morbidity such as arthritis or rheumatism, hypertension, and digestive diseases and their multimorbidity combinations can be strengthened in the formulation and practice of public health policies. In the management of common chronic diseases multimorbidity, primary healthcare institutions need to formulate personalized management plans according to different age groups, different genders, different places of residence and other different characteristics of the population. On the other hand, general practitioners need to strengthen the thinking of chronic diseases multimorbidity management, use common chronic disease multimorbidity patterns as a reference to adjust chronic diseases multimorbidity prevention and management strategies in a timely manner, and to be alerted to the change of chronic disease multimorbidity patterns.

In general, the issue of chronic diseases multimorbidity among the elderly in China is prominent and serious, and how to effectively prevent and manage chronic diseases multimorbidity among the elderly, and improve the physical functioning and health status of the elderly with multimorbidity is an important challenge for our health system.

There are some limitations in this study: the CHARLS database we used has limited coverage and timeliness, and may not fully reflect the latest health status of the middle-aged and elderly populations; only 14 common chronic diseases in the CHARLS program were included in this study, and the results of the prevalence of multimorbidity in this study may have some deviation from the objective reality; the CARMA algorithm may face the problem of dimensional catastrophe when dealing with high dimensional data, which leads to a decrease in the efficiency of the algorithm. It is sensitive to noise and missing values in the data, which may affect the accuracy and

reliability of the association rules mined; this study only analyzes the chronic diseases multimorbidity status and binary multimorbidity patterns of middle-aged and elderly people of different genders, residences, and lower elderly age groups, health-related factors such as their economic status, physical factors, level of social support were not included; and this study only reveals the common pattern of the binary multimorbidity but fails to explore the longitudinal correlation and the reasons for that in-depth.

In summary, this study found an overall increasing trend in multimorbidity prevalence from 2011 to 2020; the prevalence of multimorbidity and the pattern of multimorbidity also differed by gender, place of residence, age group, etc. The finding in this study provide evidence for the effective prevention and management of multimorbidity.

References

Agrawal, R., Imielinski, T., & Swami, A.N. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record, 22*, 207 - 216.

Asogwa, O. A., Boateng, D., Marzà-Florensa, A., Peters, S., Levitt, N., van Olmen, J., & Klipstein-Grobusch, K. (2022). Multimorbidity of non-communicable diseases in low-income and middle-income countries: a systematic review and meta-analysis. *BMJ open, 12*(1), e049133. https://doi.org/10.1136/bmjopen-2021-049133

Chen, H., Cheng, M., Zhuang, Y., & Broad, J. B. (2018). Multimorbidity among middle-aged and older persons in urban China: Prevalence, characteristics and health service utilization. *Geriatrics & gerontology international, 18*(10), 1447–1452. https://doi.org/10.1111/ggi.13510

Fan, J., Sun, Z., Yu, C., Guo, Y., Pei, P., Yang, L., Chen, Y., Du, H., Sun, D., Pang, Y., Zhang, J., Gilbert, S., Avery, D., Chen, J., Chen, Z., Lyu, J., Li, L., & China Kadoorie Biobank Collaborative Group (2022). Multimorbidity patterns and association with mortality in 0.5 million Chinese adults. *Chinese Medical Journal, 135*(6), 648–657. https://doi.org/10.1097/CM9.0000000000001985

Hidber, C. (1999). Online association rule mining. *Sigmod Record, 28*(2), 145–156. https://doi.org/10.1145/304181.304195

Kernick, D., Chew-Graham, C., & O'Flynn, N. (2017). Clinical assessment and management of multimorbidity: NICE guideline. *British Journal of General Practice, 67*(658), 235–236. https://doi.org/10.3399/bjgp17x690857

Kudesia, P., Salimarouny, B., Stanley, M., Fortin, M., Stewart, M., Terry, A., & Ryan, B. L. (2021). The incidence of multimorbidity and patterns in accumulation of chronic conditions: A systematic review. *Journal of Multimorbidity and Comorbidity, 11*, 26335565211032880. https://doi.org/10.1177/26335565211032880

Li, X. L., Huang, H., Lu, Y., Stafford, R. S., Lima, S. M., Mota, C., & Shi, X. (2023). Prediction of multimorbidity in Brazil: Latest fifth of a century population study. *JMIR Public Health and Surveillance, 9*, e44647. https://doi.org/10.2196/44647

Liu, X., Zhang, J., Zhang, S., Peng, S., Pei, M., Dai, C., Wang, T., & Zhang, P. (2023). Quality of life and associated factors among community-dwelling adults with multimorbidity in Shanghai, *China: A cross-sectional study. Nursing Open, 10*(8), 5328–5337. https://doi.org/10.1002/nop2.1770

Mohamud, M. A., Campbell, D. J. T., Wick, J., Leung, A. A., Fabreau, G. E., Tonelli, M., & Ronksley, P. E. (2023). 20-year trends in multimorbidity by race/ethnicity among hospitalized patient populations in the United States. *International Journal For Equity in Health, 22*(1), 137. https://doi.org/10.1186/s12939-023-01950-2

Ni, W., Yuan, X., Zhang, Y., Zhang, H., Zheng, Y., & Xu, J. (2023). Sociodemographic and lifestyle determinants of multimorbidity among community-dwelling older adults: findings from 346,760 SHARE participants. *BMC Geriatrics, 23*(1), 419. https://doi.org/10.1186/s12877-023-04128-1

Pati, S., Swain, S., Hussain, M. A., van den Akker, M., Metsemakers, J., Knottnerus, J. A., & Salisbury, C. (2015). Prevalence and outcomes of multimorbidity in South Asia: a systematic review. *BMJ Open, 5*(10), e007235. https://doi.org/10.1136/bmjopen-2014-007235

Rajoo, S. S., Wee, Z. J., Lee, P. S. S., Wong, F. Y., & Lee, E. S. (2021). A Systematic Review of the Patterns of Associative Multimorbidity in Asia. *BioMed Research International, 2021*, 6621785.

https://doi.org/10.1155/2021/6621785

Rivera-Almaraz, A., Manrique-Espinoza, B., Ávila-Funes, J. A., Chatterji, S., Naidoo, N., Kowal, P., & Salinas-Rodríguez, A. (2018). Disability, quality of life and all-cause mortality in older Mexican adults: association with multimorbidity and frailty. *BMC Geriatrics, 18*(1), 236. https://doi.org/10.1186/s12877-018-0928-7

Sasseville, M., Smith, S. M., Freyne, L., McDowell, R., Boland, F., Fortin, M., & Wallace, E. (2019). Predicting poorer health outcomes in older community-dwelling patients with multimorbidity: prospective cohort study assessing the accuracy of different multimorbidity definitions. *BMJ Open, 9*(1), e023919. https://doi.org/10.1136/bmjopen-2018-023919

Schäfer, I., von Leitner, E. C., Schön, G., Koller, D., Hansen, H., Kolonko, T., Kaduszkiewicz, H., Wegscheider, K., Glaeske, G., & van den Bussche, H. (2010). Multimorbidity patterns in the elderly: a new approach of disease clustering identifies complex interrelations between chronic conditions. *PloS One, 5*(12), e15941. https://doi.org/10.1371/journal.pone.0015941

Silva, D. S. M. D., Silva, M. F., Assumpção, D., Francisco, P. M. S. B., Neri, A. L., Yassuda, M. S., & Borim, F. S. A. (2023). Influência de padrões de multimorbidade nas atividades de vida diária da pessoa idosa: seguimento de nove anos do Estudo Fibra [Influence of multimorbidity patterns on the activities in the day-to-day lives of the elderly: nine-year follow-up of the Fibra Study]. *Ciencia & Saude Coletiva, 28*(7), 2003–2014. https://doi.org/10.1590/1413-81232023287.14842022

Yao, S. S., Cao, G. Y., Han, L., Chen, Z. S., Huang, Z. T., Gong, P., Hu, Y., & Xu, B. (2020). Prevalence and Patterns of Multimorbidity in a Nationally Representative Sample of Older Chinese: Results From the China Health and Retirement Longitudinal Study. The journals of gerontology. Series A, *Biological Sciences and Medical Sciences, 75*(10), 1974–1980. https://doi.org/10.1093/gerona/glz185

Zhang, X., Padhi, A., Wei, T., Xiong, S., Yu, J., Ye, P., Tian, W., Sun, H., Peiris, D., Praveen, D., & Tian, M. (2022). Community prevalence and dyad disease pattern of multimorbidity in China and India: a systematic review. *BMJ Global Health, 7*(9), e008880. https://doi.org/10.1136/bmjgh-2022-008880

Zhao, Y., Wang, J., Zhu, X., Zhang, X., Zhang, Y., Zhang, W., & Dong, Y. (2023). Multimorbidity and polypharmacy in hospitalized older patients: a cross-sectional study. *BMC Geriatrics*, *23*(1), 423. https://doi.org/10.1186/s12877-023-04109-4

Zhong, Y., Xi, H., Guo, X., Wang, T., Wang, Y., & Wang, J. (2022). Gender and Socioeconomic Differences in the Prevalence and Patterns of Multimorbidity among Middle-Aged and Older Adults in China. *International Journal of Environmental Research and Public Health, 19*(24), 16956. https://doi.org/10.3390/ijerph192416956