

基于自组织映射网络的中老年人慢性病关联及共病模式可视化聚类分析

王福琳^{1,2}, 杨超^{3,4,5*}, 杜建², 孔桂兰², 张路霞^{2,3,4,5}

¹北京大学医学部医学技术研究院, 北京 100191; ²北京大学健康医疗大数据国家研究院, 北京 100191; ³北京大学第一医院肾内科/北京大学肾脏病研究所, 北京 100034; ⁴中国医学科学院免疫介导肾病诊治创新单元, 北京 100034; ⁵浙江省北大信息技术高等研究院, 浙江杭州 311215

[中图分类号] R195.4

[文献标志码] A

[DOI]

10.11855/j.issn.0577-7402.2022.12.1217

[声明]

本文所有作者声明无利益冲突

[引用本文]

王福琳, 杨超, 杜建, 等. 基于自组织映射网络的中老年人慢性病关联及共病模式可视化聚类分析[J]. 解放军医学杂志, 2022, 47(12): 1217-1225.

[收稿日期] 2022-03-01

[录用日期] 2022-06-16

[上线日期] 2022-09-15

[摘要] **目的** 采用可视化聚类分析方法探究我国中老年人中常见慢性病之间的关联关系及共病模式。**方法** 采用中国健康与养老追踪调查2015年数据, 提取年龄 ≥ 45 岁受调查人群的年龄、性别、居住地、14种慢性病患病及治疗信息, 纳入相关信息均完整的研究对象18 542例。采用自组织映射(SOM)神经网络与K-Means结合的两步聚类算法, 对我国中老年人慢性共病模式进行可视化聚类分析。**结果** 患有两种或两种以上慢性病患者8044例, 共病患病率为43.38%, 其中 ≥ 60 岁老年人群的共病患病率达52.28%。在14种慢性病中, 患病率居前三位的分别是关节炎或风湿病(33.02%)、高血压病(31.07%)、胃部疾病或消化系统疾病(23.60%); 其中肝脏疾病的未治疗率达56.99%。所有共病患者被划分为4类人群: 第1类人群慢性肺部疾患的患病率达97.72%, 55.05%罹患关节炎或风湿病; 第2类人群高血压病患病率为98.21%; 第3类人群血脂异常的患病率高达99.49%, 且91.72%患有高血压病; 第4类人群73.39%患关节炎或风湿病, 且68.11%患有胃部疾病或消化系统疾病。女性和城市人群的共病模式略有不同。**结论** 我国中老年人慢性病的负担严重、治疗欠佳, 不同性别、年龄和城乡人口的共病模式有一定差异。基于可视化聚类方法得到的共病模式结果, 对促进多病共防及降低慢性病共病负担有较高的参考价值。

[关键词] 慢性病; 共病模式; 聚类分析; 自组织映射网络; 可视化

Patterns of multimorbidity in the middle-aged and elderly population in China: a cluster analysis based on self-organizing map

Wang Fu-Lin^{1,2}, Yang Chao^{3,4,5*}, Du Jian², Kong Gui-Lan², Zhang Lu-Xia^{2,3,4,5}

¹Institute of Medical Technology, Peking University Health Science Center, Beijing 100191, China

²National Institute of Health Data Science at Peking University, Beijing 100191, China

³Renal Division, Department of Medicine, Peking University First Hospital/Peking University Institute of Nephrology, Beijing 100034, China

⁴Research Units of Diagnosis and Treatment of Immune-Mediated Kidney Diseases, Chinese Academy of Medical Sciences, Beijing 100034, China

⁵Advanced Institute of Information Technology, Peking University, Hangzhou, Zhejiang 311215, China

*Corresponding author, E-mail: yangchao@bjmu.edu.cn

This study was supported by the National Natural Science Foundation of China (82003529, 72125009), and the CAMS Innovation Fund for Medical Sciences (2019-I2M-5-046)

[基金项目] 国家自然科学基金(82003529, 72125009); 中国医学科学院医学与健康科技创新工程项目(2019-I2M-5-046)

[作者简介] 王福琳, 博士研究生, 主要从事健康医疗大数据在重大疾病防治中的应用研究

[通信作者] 杨超, E-mail: yangchao@bjmu.edu.cn

[Abstract] Objective To reveal the pattern of multimorbidity in the middle-aged and elderly population in China by visual cluster analysis. **Methods** Using the 2015 data of China Health and Retirement Longitudinal Study, the age, gender, place of residence, and 14 kinds of non-communicable chronic diseases of the study population were extracted. A total of 18 542 subjects with complete information were included. A two-step clustering algorithm combining self-organizing map and K-Means was used to visually cluster the existence of chronic diseases among the middle-aged and elderly population. **Results** There were 8044 patients with 2 or more chronic diseases. The prevalence of multimorbidity was 43.38% in those patients, and 52.28% in the elderly population aged 60 years or above. Among the 14 chronic diseases, arthritis or rheumatism had the highest prevalence (33.02%), followed by hypertension (31.07%), stomach or other digestive diseases (23.60%). The patterns of multimorbidity included the following four categories: (1) 97.72% of patients in group A had chronic lung diseases, with more than half (55.05%) suffering from arthritis or rheumatism; (2) the prevalence of hypertension in group B was 98.21%; (3) the prevalence of dyslipidemia in group C was as high as 99.49%, and 91.72% of patients had hypertension; (4) 73.39% of patients in group D suffered from arthritis or rheumatism, and 68.11% had stomach or other digestive diseases. The patterns of multimorbidity were slightly different in women and urban populations. **Conclusions** The situation of chronic diseases in the middle-aged and elderly population in China is not optimistic, and the patterns of multimorbidity among different genders, urban and rural populations are different. Those results based on visual cluster analysis are of great significance for co-prevention of multiple conditions and reducing the burden of chronic diseases.

[Key words] non-communicable chronic disease; pattern of multimorbidity; cluster analysis; self-organizing map; visualization

近年来,在全球老龄化程度加深的背景下,随着人口结构的改变和疾病检出率的提升,慢性病的患病率和疾病负担不断增高^[1]。慢性病具有病程长、病因复杂、预后不佳等特点,已成为危害人类健康的重要原因^[2-3],在中老年人群尤其突出^[4]。随着慢性病患病率上升,一人共患多种慢性病的现象越来越严重^[5]。2008年,世界卫生组织将慢性病共病定义为共存于同一患者体内的两种或两种以上慢性病^[6]。慢性病共病会导致患者生活质量下降及多重用药^[7-8],加重医疗负担和医师正确处方的难度^[5]。中老年人慢性病共病及其健康管理已经成为当前重要的公共卫生问题^[9-10]。本研究将一种经典的神经网络——自组织映射(self-organizing map, SOM)神经网络引入共病模式的研究中,通过可视化聚类分析方法,直观揭示我国中老年人中常见慢性病之间的关联关系与共病模式。

1 资料与方法

1.1 数据来源 本研究采用的数据来源于中国健康与养老追踪调查(China Health and Retirement Longitudinal Study, CHARLS)2015年数据集。CHARLS是一项由北京大学国家发展研究院主持、北京大学中国社会科学调查中心与北京大学团委共同执行的大型跨学科调查项目,全国基线调查于2011年开展,至今已在全国28个省(自治区、直辖市)的150个县、450个社区(村)开展调查访问,样本覆盖1.24万户家庭中的近2万名受访者^[11-12]。CHARLS在县(区)和社区(村)抽样阶段均采取分层随机抽样法,因此,利用该数据集对我国中老年人群的慢性病共病

模式进行分析,具有较好的人群代表性。

1.2 研究对象 研究对象为2015年CHARLS调查人口中年龄 ≥ 45 岁的人群,提取的基本人口学因素包括年龄、性别和居住地。在CHARLS追访问卷的“健康状况和功能”模块中,“是否有医生曾经告诉过您有以下这些慢性病?”的问题列出了以下14种慢性病:高血压病、血脂异常、糖尿病或血糖升高、癌症等恶性肿瘤、慢性肺部疾患、肝脏疾病、心脏病、卒中(俗称中风)、肾脏疾病、胃部疾病或消化系统疾病、情感及精神方面问题、与记忆相关的疾病、关节炎或风湿病、哮喘;并且收集被调查者的慢性病药物治疗情况。在2015年CHARLS调查数据中共提取 ≥ 45 岁的受访者19 713例,本研究纳入性别、年龄、居住地、受访类型和患病信息记录均完整的18 542例作为研究对象,其中慢性病共病患者(同时患两种或两种以上慢性病)8044例。在可视化聚类分析中,仅使用慢性病共病患者的数据进行分析。本研究经北京大学生物医学伦理委员会批准(批准文号:IRB00001052-11015),所有研究对象均签署书面知情同意书。

1.3 研究方法 采用SOM神经网络与K-Means结合的两步聚类算法,对我国中老年人群慢性病共病模式进行可视化聚类分析。

1.3.1 SOM神经网络 SOM神经网络是一种竞争学习型的无监督神经网络,可将高维空间中相似的样本点映射到低维空间中的邻近神经元^[13]。当输出层设定为二维空间时,SOM网络结构包含输入层和输出层。输入层负责感知外界的输入信息,将输入模式向输出层传递;输出层则采用竞争机制,被

激活的竞争获胜神经元决定输入向量映射在低维空间中的位置。输出层神经元数目的确定参照如下公式：神经元个数= $5\sqrt{N}$ 。式中N为训练样本数，正方形拓扑边长对神经元个数开根号后再取整。

1.3.2 Davies-Bouldin指数(DBI)作为聚类性能度量指标 DBI综合运用类内距离和类间距离评价聚类效果，基本原则为保证聚类结果得到的各簇簇内相似度高，且簇间相似度低。DBI值越小，提示聚类的效果越好。

1.3.3 SOM二次聚类 本研究采用SOM网络与K-Means结合的二次聚类方法对患病人群进行聚类分析，算法流程如下：(1)训练SOM网络，将患病情况相似的患者映射到相同或相似的神经元中；采用R软件(4.1.1版本)kohonen包进行SOM网络训练，设置迭代次数为10 000，拓扑结构采用二维六边形网络，神经元数目设置为21 × 21，初始学习率为0.05，最终学习率为0.01。训练结束后得到不同疾病分布的可视化展示，以分析相关的共病规律。

(2)采用K-Means聚类方法对SOM网络输出层神经元再次进行聚类，采用DBI确定最佳聚类数目，目的是将相似神经元进一步合并，从而得出最常见的共病模式。

1.4 统计学处理 采用R 4.1.1软件进行统计分析。计数资料以例(%)表示，组间比较采用 χ^2 检验。 $P<0.05$ 为差异有统计学意义。

2 结 果

2.1 基本情况 研究对象年龄(60.2 ± 10.1)岁，其中48.94%为男性，78.49%居住在农村地区。仅患一种慢性病者5052例(27.25%)，患有两种或两种以上慢性病的共病患者8044例(43.38%)。其中， ≥ 60 岁的老年人群慢性病共病患病率高于45~59岁人群(52.28% vs. 34.42%， $P<0.05$)，女性人群慢性病共病患病率高于男性人群(46.06% vs. 40.59%， $P<0.05$)，城市人口慢性病共病患病率高于农村人口(45.73% vs. 42.74%， $P<0.05$ ，表1)。

表1 研究对象慢性病基本情况

Tab.1 General information of chronic diseases in study subjects

项目	例数(%)	慢性病患者情况[例(%)]			前三位慢性病(%)		
		不患病	患一种病	共病(≥ 2)	第1位	第2位	第3位
性别							
男	9074(48.94)	2872(31.65)	2519(27.76)	3683(40.59)	高血压病(30.05)	关节炎或风湿病(27.73)	胃部疾病或消化系统疾病(20.53)
女	9468(51.06)	2574(27.19)	2533(26.75)	4361(46.06)	关节炎或风湿病(38.10)	高血压病(32.04)	胃部疾病或消化系统疾病(26.54)
年龄(岁)							
45~60	9239(49.83)	3479(37.66)	2580(27.93)	3180(34.42)	关节炎或风湿病(27.89)	胃部疾病或消化系统疾病(22.64)	高血压病(22.60)
≥ 60	9303(50.17)	1967(21.14)	2472(26.57)	4864(52.28)	高血压病(39.48)	关节炎或风湿病(38.12)	胃部疾病或消化系统疾病(24.55)
居住地							
城市	3989(21.51)	1179(29.56)	986(24.72)	1824(45.73)	高血压病(34.22)	关节炎或风湿病(25.70)	血脂异常(21.94)
农村	14 553(78.49)	4267(29.32)	4066(27.94)	6220(42.74)	关节炎或风湿病(35.03)	高血压病(30.21)	胃部疾病或消化系统疾病(24.74)
合计	18 542(100)	5446(29.37)	5052(27.25)	8044(43.38)	关节炎或风湿病(33.02)	高血压病(31.07)	胃部疾病或消化系统疾病(23.60)

2.2 慢性病患病及共病情况 在研究涉及的14种慢性病中，患病率居前三位的是关节炎或风湿病(33.02%)、高血压病(31.07%)、胃部疾病或消化系统疾病(23.60%)；不同人群的前三位慢性病患者情况略有不同，其中城市人口中血脂异常的患病率(21.94%)较高(表1)。关节炎或风湿病的共病患病率(25.00%)最高。14种慢性病的疾病共病患病率均 $>75\%$ ，其中哮喘的疾病共病患病率(95.05%)最高。此外，在心脏病、卒中、与记忆相关的疾病和肾脏疾病患者中，共病的比例也较高，显示患有上述慢性病的中老年人群更容易共患其他疾病(表2)。

2.3 慢性病治疗情况 在上述14种慢性病中，CHARLS针对血脂异常、慢性肺部疾患、肝脏疾病、心脏病、肾脏疾病、胃部疾病或消化系统疾病、与记忆相关的疾病、关节炎或风湿病共8种疾病的用药情况进行了调查。结果显示，上述8种慢性病的治疗率均 $>40\%$ ，用药类型以西药为主；其中肝脏疾病的未治疗率最高(56.99%，表3)。

2.4 慢性病共病模式

2.4.1 总体共病聚类结果 SOM网络的训练结果显示，不同节点内扇形的比例可直观展现映射到对应神经元内的患者各疾病的患病情况(图1)。网络

表2 研究对象中14种慢性病的患病及共病情况

Tab.2 Statistics of 14 chronic diseases of study subjects

疾病	例数	患病率(%)	同时患有该疾病与其他疾病的例数	共病患病率(%)	疾病共病患病率(%)
关节炎或风湿病	6123	33.02	4636	25.00	75.71
高血压病	5761	31.07	4479	24.16	77.75
胃部疾病或消化系统疾病	4376	23.60	3487	18.81	79.68
心脏病	2664	14.37	2463	13.28	92.45
血脂异常	2576	13.89	2305	12.43	89.48
慢性肺部疾患	2286	12.33	1993	10.75	87.18
糖尿病或血糖升高	1576	8.50	1380	7.44	87.56
肾脏疾病	1229	6.63	1113	6.00	90.56
哮喘	829	4.47	788	4.25	95.05
肝脏疾病	765	4.13	675	3.64	88.24
卒中	501	2.70	459	2.48	91.62
情感及精神方面问题	425	2.29	368	1.98	86.59
与记忆相关的疾病	407	2.20	370	2.00	90.91
癌症等恶性肿瘤	231	1.25	182	0.98	78.79

共病患病率=(同时患有该疾病与其他疾病的例数/受访者总数)×100%; 疾病共病患病率=(同时患有该疾病与其他疾病的例数/患有该疾病的例数)×100%

表3 研究对象中8种慢性病的治疗情况[例(%)]

Tab.3 Treatment of 8 chronic diseases of study subjects [n(%)]

疾病	例数	服用中药	服用西药	服药以外的其他治疗方法	以上均无
血脂异常	2576	332(12.89)	1167(45.30)	236(9.16)	1121(43.52)
慢性肺部疾患	2286	398(17.41)	1079(47.20)	204(8.92)	950(41.56)
肝脏疾病	765	145(18.95)	238(31.11)	50(6.54)	436(56.99)
心脏病	2664	624(23.42)	1503(56.42)	201(7.55)	823(30.89)
肾脏疾病	1229	275(22.38)	437(35.56)	111(9.03)	586(47.68)
胃部疾病或消化系统疾病	4376	881(20.13)	2410(55.07)	263(6.01)	1554(35.51)
与记忆相关的疾病	407	73(17.94)	172(42.26)	24(5.90)	199(48.89)
关节炎或风湿病	6123	1209(19.75)	2353(38.43)	1725(28.17)	2306(37.66)

训练结束后患病情况相似的慢性病共病患者已初步映射到相邻的神经元。SOM网络训练后的U-Matrix矩阵和各疾病的分布情况见图2。各疾病聚类分布图中,红色表示相应位置神经元中该疾病患病率较高,蓝色则反之。

进一步对SOM网络训练得到的神经元进行K-Means聚类,借助DBI确定最佳聚类数目为4。因此,最终结果将所有慢性病共病患者划分为4类人群,每种颜色对应一个聚类簇(图3);并绘制出每簇人群中各慢性病的患病率热图(图4)。第1类人群慢性肺部疾患的患病率达97.72%,关节炎或风湿病患者占55.05%;第2类人群高血压病患病率达98.21%,关节炎或风湿病患病率(56.57%)也较高;第3类人群血脂异常患病率为99.49%,此外91.72%患高血压病,糖尿病或血糖升高(40.05%)和心脏病(39.80%)患病率也较高;第4类人群73.39%患关节炎或风湿病,68.11%患胃部疾病或消化系统疾病。

2.4.2 不同人群慢性病共病模式的差异 不同慢性病共病模式人群的性别、年龄以及居住地特征见图5。4类人群之间比较,类别1和类别2中≥60岁及农村人口比例较高,其中类别1的男性比例偏高;类别3中城市人口比例相对较高;类别4中女性比例较高,且45~60岁的比例相对较高。

此外,按照性别、年龄、居住地对慢性病共病患者分别进行聚类分析,不同亚组的聚类结果及各聚类簇内的患病情况见图6。其中,女性人群呈现较为集中的患病模式:类别1中慢性肺部疾患的患病率达99.15%,同时97.16%患心脏病;类别2中心脏病患病率达99.47%,同时61.88%患有高血压病。城市人口中心脏病和高血压病关联性较高。

3 讨论

慢性病共病的研究涉及临床医学、流行病学以及社会学等领域。有研究显示,年龄较大人群、女

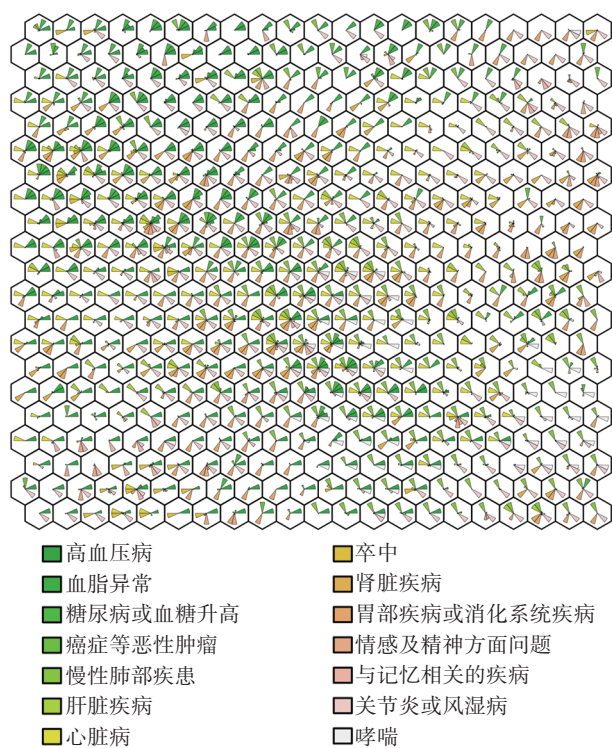


图1 采用SOM神经网络训练后得到的共病聚类结果

Fig.1 Cluster results on patterns of multimorbidity by training of self-organizing map

SOM神经网络. 自组织映射神经网络; 图中每个神经元中, 按照圆心角平分为14个扇形, 分别对应14种慢性病, 某种疾病对应的扇形半径越大, 则映射到该神经元的输入样本中患该病的例数越多

性人群、社会经济水平及教育水平较高的人群患慢性病共病的可能性更大, 慢性病共病可能引发一系列健康不平等、灾难性卫生支出等问题^[11]。既往针

对慢性病共病模式的研究主要以描述性分析为主, 研究对象更多关注老年人群, 所采用的聚类方法集中于因子分析、关联规则算法等传统方法, 研究结果的可视化效果欠佳, 对于患病规律的表达能力不强, 未能对我国中老年人群的慢性病负担和共病情况进行深入挖掘。

本研究采用SOM神经网络对我国中老年人群的慢性病患者情况和共病模式进行可视化聚类分析, 结果显示该人群中共病患病率达43.38%, 其中 ≥ 60 岁人群的慢性病共病情况更为突出, 但是慢性病的治疗率相对欠佳。本研究将共病患者进一步聚类为4类人群, 其主要特点分别为: 慢性肺部疾患/关节炎或风湿病; 高血压病/关节炎或风湿病; 血脂异常/高血压病(同时可能伴有心脏病、糖尿病或血糖升高); 关节炎或风湿病/胃部疾病或消化系统疾病。不同类别人群的共病模式挖掘结果显示, 女性人群中心脏病可能与高血压病、慢性肺部疾患有更强的关联, 城市人口中心脏病和高血压病之间存在较强的关联, 提示需要在中老年人慢性病共病模式的基础上, 重点关注特定人群共病模式的差异, 有所侧重地开展慢性病共病的筛查与防控, 采取精准化的治疗和管理措施。

近年来, 我国慢性病患者率显著增高, 尤其是中老年人的慢性病共病问题逐渐凸显^[14-15]。郭小榕^[9]以CHARLS 2015年数据集中 ≥ 60 岁的人群作为研究对象, 利用web图发现了胃部疾病/关节炎或风湿病、高血压/关节炎或风湿病两种关联度较强的共病模式, 使用关联规则Apriori算法挖掘到以下3种较为常见的慢性病共病模式: 哮喘/慢性肺部疾

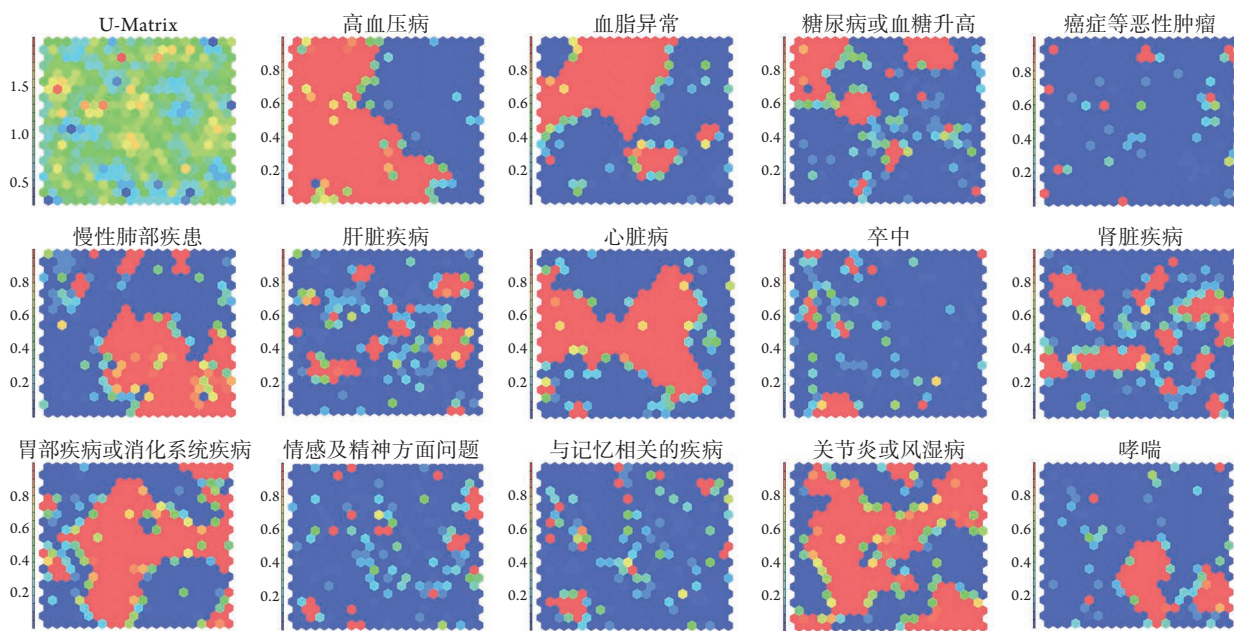


图2 U-Matrix及各种慢性病聚类分布图

Fig.2 U-Matrix and distributions of each chronic disease

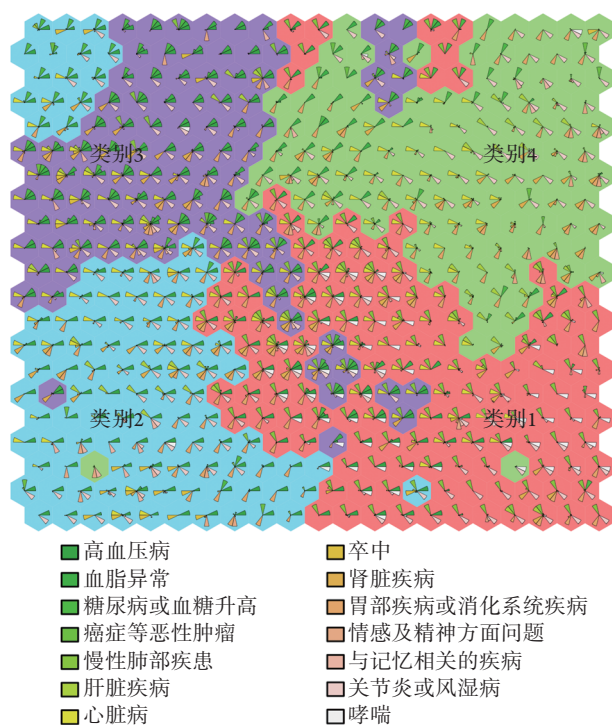


图3 全体中老年慢性病共病人群的二次聚类结果

Fig.3 Secondary clustering results of all comorbid populations

患, 关节炎或风湿病、哮喘/慢性肺部疾患, 血脂异常、高血压、关节炎或风湿病/心脏病; 上述模式与本研究报告中报告的共病模式基本一致。程杨杨等^[16]选取CHARLS 2015年数据集中 ≥ 45 岁的人群, 使用关联规则算法挖掘到以下3种常见的慢性病共病模式: 血脂异常/心脏病、糖尿病/血脂异常、哮喘/慢性肺部疾病, 其中血脂异常与心脏病、糖尿病之间的关联在本研究结果中也有所体现。既往也有研究者关注到相关疾病之间的关联关系。例如, Boeckxstaens等^[17]指出高血压病和关节炎之间存在较为紧密的关联; Bieber等^[18]指出, 类风湿关节炎患者更容易罹患慢性阻塞性肺疾病。既往使用CHARLS数据进行的共病研究中, 大多数只挖掘到数据集中频繁出现(患病率较高)的疾病的关联信息; 而本研究使用SOM网络, 通过较好的可视化能力对疾病的分布情况进行直观展示, 显示出一些人群中患病率较低的疾病也可能具有相应的聚集趋势。例如, 卒中、情感及精神方面问题、与记忆相关的疾病等三类疾病分布规律存在一定相似性, 肝脏疾病和肾脏疾病的分布也有相似之处, 高血压病

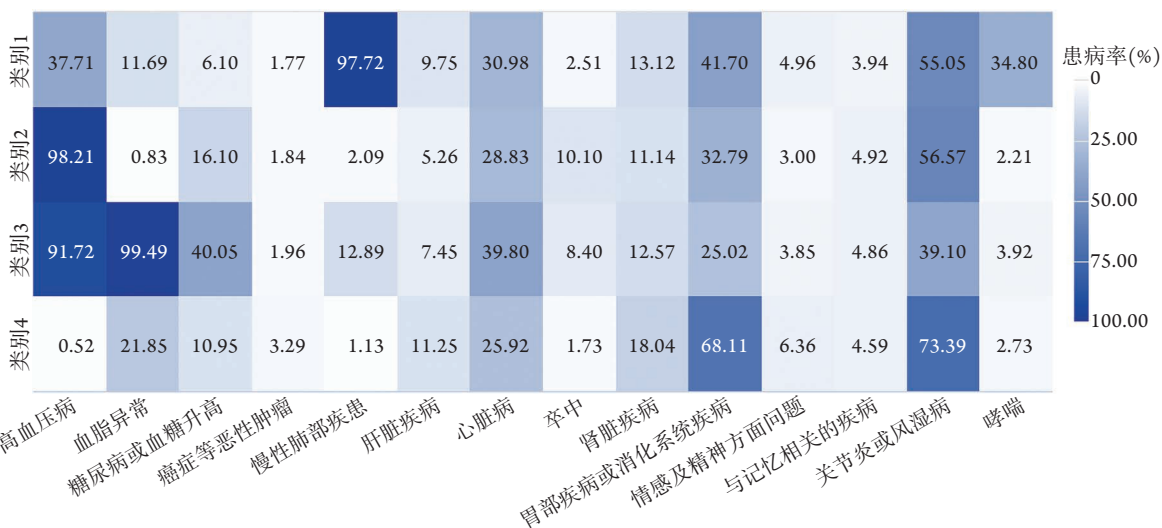


图4 全体中老年慢性病共病人群的各项聚类簇患病率统计热图

Fig.4 Heat map of prevalence of each disease in all comorbid populations

与关节炎或风湿病、心脏病、血脂异常等的聚类分布具有较大面积的重叠等, 提示这些疾病之间可能存在关联关系, 同样需要引起重视, 建议加强对这些疾病的多病共管、同防。

此外, 本研究较为详细地分析了不同共病模式人群人口学特征的差异, 结果显示: “慢性肺部疾患/关节炎或风湿病”类别中男性比例较高, 这可能与男性群体中吸烟比例较高有关; “血脂异常/高血压病”类别中城市人口比例较高, 而其他3类包含关节炎或风湿病的共病模式中农村人口比

例较高, 这可能与城乡人口的生活方式、体力劳动和饮食等存在差异有关。同时, 本研究根据性别、年龄、居住地对共病患者分别进行了聚类分析, 结果显示女性和城市人口有独特的共病模式, 这可能与女性围绝经期后雌激素水平下降以及不同地区的经济水平和人群生活方式、行为习惯等存在差异有关^[19], 提示应针对不同性别、年龄和城乡人口采取有针对性的干预策略, 根据人群特征“量体裁衣”, 制定差异化的慢性病管理方案。

以往的共病模式研究较多采用的方法有因子

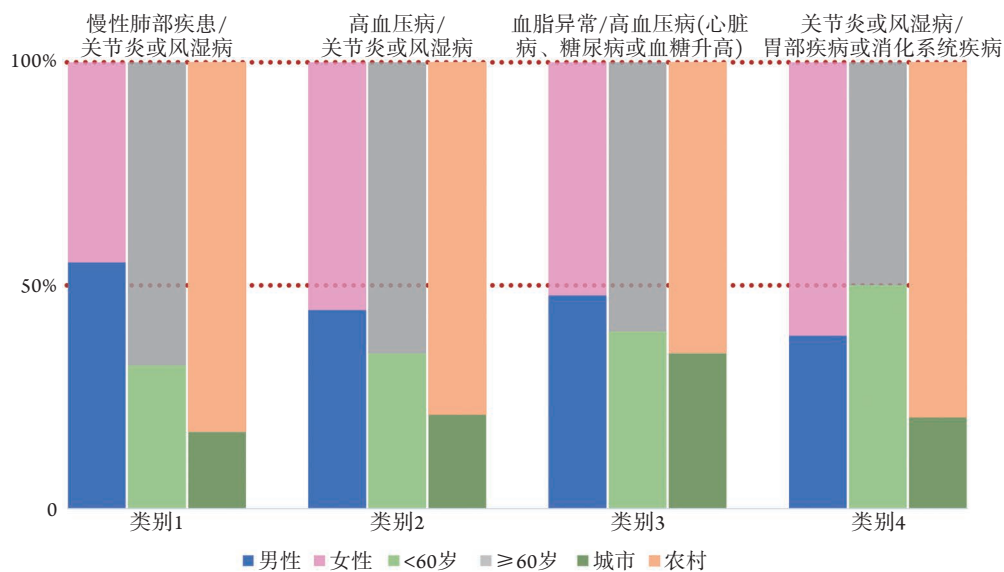


图5 不同慢性病共病模式中老年人的人口学特征

Fig.5 Demographic characteristics for each pattern of multimorbidity

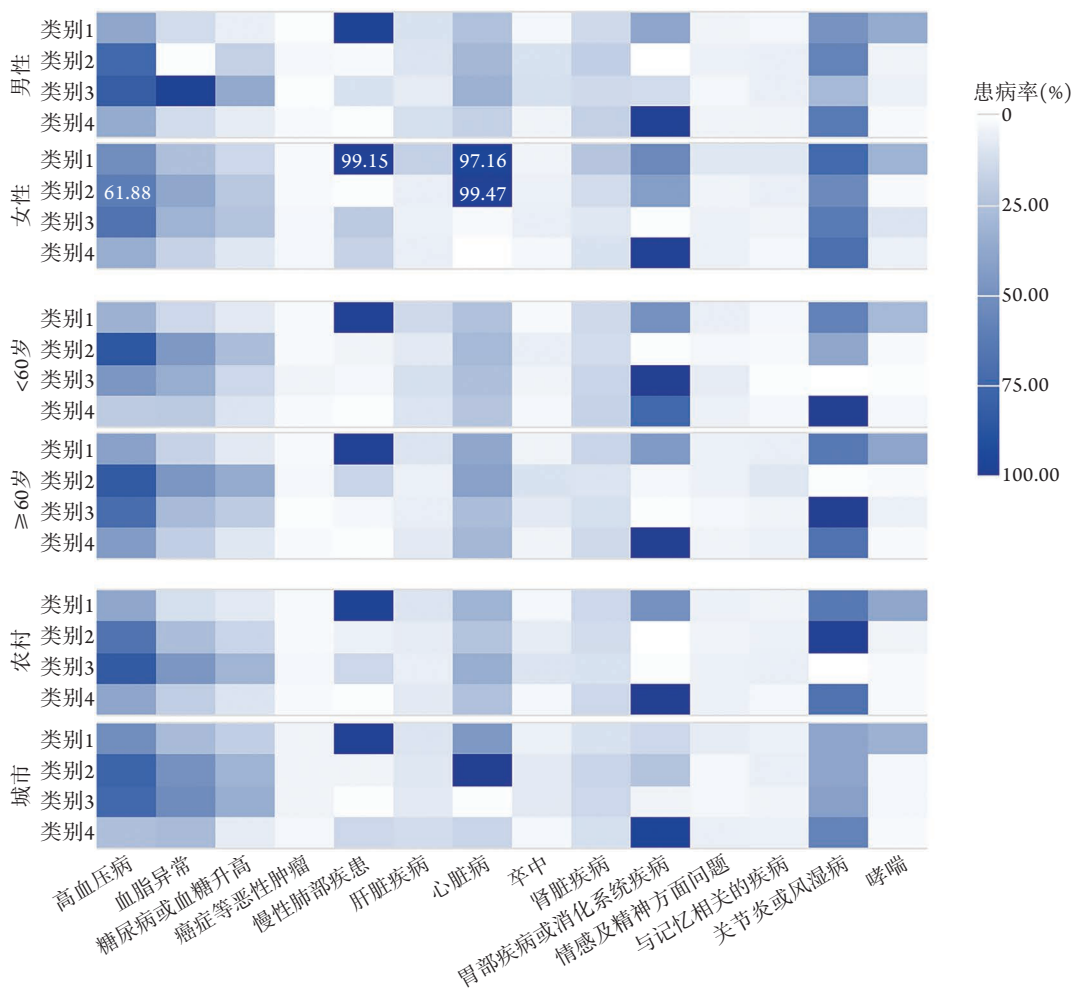


图6 各亚组慢性病共病中老年人中各聚类簇患病率统计热图

Fig.6 Heat map of prevalence of each disease in each subgroup of multimorbidity

按照性别、年龄、居住地分别对慢性病共病患者进行二次聚类分析得到的4个类别及其患病情况

分析、关联规则挖掘、层次聚类分析^[16,20-21]等。其中，因子分析是将多个疾病指标转化为少数综合指

标，得到的类别可解释性较高；关联规则挖掘的应用场景主要是挖掘数据集中频繁出现的疾病组

合,方法简单,是当前较为常用的共病模式分析方法之一;聚类分析方法可以发现数据的聚集趋势,适合探索新的共病假设(表4)。目前共病分析中常用的聚类分析方法有针对疾病指标进行聚类的层次聚类分析和针对患病人群进行聚类的K-Means系列方法,但这两种方法可能存在计算复杂度较高、易受噪声点影响、对疾病关联的展示能力较弱等局限性。本研究采用的SOM二次聚类算法,使用场景较为灵活,可以选择与层次聚类或K-Means聚类搭配使用,分别实现对疾病指标或患病人群的聚类分析^[13]。对于人群较大、疾病数量较多的高维数

据,SOM二次聚类方法可以在有效降低计算量的同时,将数据映射到低维空间,直观地展现患者患病模式的聚集趋势,有助于理解疾病之间的关联关系及共病模式。本研究结果显示,将SOM网络结合K-Means的两步聚类方法应用于共病模式的研究,不仅方法合理可行,而且结论较为可靠。但本研究仍存在一定的局限性,包括:可能存在研究对象的回忆和报告偏倚;仅基于横断面调查数据对14种慢性病的患病情况进行调查,患病率计算可能会受到人群选择偏倚的影响;未对慢性病共病模式的发生机制及其相关危险因素作深入挖掘等。

表4 常用共病模式研究方法的比较

Tab.4 Comparison of common methods for multimorbidity analysis

研究方法	应用场景	优点	缺点
因子分析	将多个疾病指标转化为少数综合指标	结果从医学角度可解释性较高、容易理解	相比于聚类分析,更倾向于描述在患病原理上相似的疾病群,而非探索新的共病关联 ^[20]
关联规则挖掘(Apriori算法)	挖掘数据集中频繁出现的疾病组合	方法简单明了	有最小支持度限制,患病率较低的疾病会被剔除;计算复杂度较高,数据量较大时不适用
聚类分析	发现疾病或患病人群的聚集趋势	适合探索新的共病关联	不同聚类算法得到的聚集趋势可能存在差异
层次聚类分析	对疾病指标进行聚类	可以发现共病类别间的层次关系;能够将患病率较低的疾病纳入分析结果	计算复杂度高,疾病数过多时不适用
K-Means系列	对患病人群进行聚类	算法原理简单、使用方便	容易受噪声点影响;对疾病关联的层次展示、细致捕捉能力较弱;需要依次尝试以确定类别数目,消耗计算资源
SOM二次聚类	对疾病指标或患病人群进行聚类	可视化效果较好;不会排除患病率较低的疾病;降低噪声点的影响;SOM网络可以搭配层次聚类和K-Means算法使用,兼具两者优点;降低部分计算量	可解释性及方法细节优化等问题有待进一步研究探讨

综上所述,本研究采用一种新的共病模式可视化方法,对我国中老年人慢性病患者情况和共病模式进行聚类分析,结果显示,我国中老年人慢性病的慢性负担严重、治疗欠佳,不同性别、年龄和城乡人群的共病模式存在一定差异,需引起医疗从业者和相关卫生政策制定者的重视,并建议将慢性病共病管理纳入基层医疗卫生机构的工作中。基于可视化聚类方法得到的慢性病共病模式,能够对多病共防和用药方案制定发挥指导作用,对于降低我国慢性病共病负担,构建“以人为本”的整合型医疗卫生服务体系也有一定的参考价值。

【参考文献】

- Zemedikun DT, Gray LJ, Khunti K, et al. Patterns of multimorbidity in middle-aged and older adults: an analysis of the UK Biobank data[J]. *Mayo Clin Proc*, 2018, 93(7): 857-866.
- Yin P, Qi JL, Liu YN, et al. Burden of disease in the Chinese population from 2005 to 2017[J]. *Chin Circ J*, 2019, 34(12): 1145-1154. [殷鹏, 齐金蕾, 刘韞宁, 等. 2005—2017年中国疾病负担研究报告[J]. *中国循环杂志*, 2019, 34(12): 1145-1154.]
- Qian HS, Ma AX. Multimorbidity status and its risk factors to multimorbidity in middle-aged people in China[J]. *Chin J Drug Eval*, 2017, 34(6): 476-480. [钱焯森, 马爱霞. 我国中老年人共病现状及其影响因素分析[J]. *中国药物评价*, 2017, 34(6): 476-480.]
- Wang M, Tan P, Xie XQ. A retrospective analysis of the comprehensive evaluation results of health development in Beijing from 2012 to 2016[J]. *Chin J Health Inform Manag*, 2020, 17(4): 461-466. [王梅, 谭鹏, 谢学勤. 2012—2016年北京市卫生发展综合评价结果回顾性分析[J]. *中国卫生信息管理杂志*, 2020, 17(4): 461-466.]
- Yan W, Lu Y, Zhang R, et al. Multimorbidity status of the elderly in China—research based on CHARLS data[J]. *Chin J Dis Control Prev*, 2019, 23(4): 426-430. [闫伟, 路云, 张冉, 等. 基于CHARLS数据分析的我国老年人共病现状研究[J]. *中华疾病控制杂志*, 2019, 23(4): 426-430.]
- World Health Organization. The World Health Report 2008-Primary Health Care (Now more than ever)[EB/OL]. [2020-11-07]. <http://www.who.int/whr/2008/en/>.
- Hunger M, Thorand B, Schunk M, et al. Multimorbidity and health-related quality of life in the older population: results from the German KORA-Age study[J]. *Health Qual Life Outcomes*, 2011, 9: 53.
- Gu J, Chao JQ, Chen WJ, et al. Multimorbidity and health-related quality of life among the community-dwelling elderly: a

- longitudinal study[J]. Arch Gerontol Geriatr, 2018, 74: 133-140.
- [9] Guo XR. Study on multimorbidity and its main influence factors among elderly in China[D]. Xiamen: Xiamen University, 2019. [郭小榕. 中国老年人慢性病共病现状及主要影响因素的研究[D]. 厦门: 厦门大学, 2019.]
- [10] Liu XH, Cao H, Zhang H, *et al.* Effects of long-term exposure to atmospheric ozone on depressive, anxiety and stress symptoms of natural population in the community[J]. New Medicine, 2021, 31(1): 5-13. [刘晓慧, 曹寒, 张晗, 等. 大气臭氧长期暴露对社区自然人群抑郁、焦虑和压力状况的影响[J]. 医学新知, 2021, 31(1): 5-13.]
- [11] Zhao Y, Atun R, Oldenburg B, *et al.* Physical multimorbidity, health service use, and catastrophic health expenditure by socioeconomic groups in China: an analysis of population-based panel data[J]. Lancet Glob Health, 2020, 8(6): e840-e849.
- [12] Zhao YH, Hu YS, Smith JP, *et al.* Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS)[J]. Int J Epidemiol, 2014, 43(1): 61-68.
- [13] Vesanto J, Alhoniemi E. Clustering of the self-organizing map[J]. IEEE Trans Neural Netw, 2000, 11(3): 586-600.
- [14] Li LZ, Fang FS, Tian H, *et al.* Increase in incidence and influential factor of abnormal glycometabolism in a group of middle-old age cadres during 3 years[J]. Med J Chin PLA, 2011, 36(5): 505-507. [李立柱, 方福生, 田慧, 等. 中老年干部人群3年体检中糖代谢异常率增长情况及影响因素分析[J]. 解放军医学杂志, 2011, 36(5): 505-507.]
- [15] He Y, Zeng Q, Zhao XL. Association of body mass index and waist circumference with risk of hypertension and diabetes in Chinese adults[J]. Med J Chin PLA, 2015, 40(10): 803-808. [贺媛, 曾强, 赵小兰. 中国成人肥胖、中心性肥胖与高血压和糖尿病的相关性研究[J]. 解放军医学杂志, 2015, 40(10): 803-808.]
- [16] Cheng YY, Cao Z, Hou J, *et al.* Investigation and association analysis of multimorbidity in middle-aged and elderly population in China[J]. Chin J Dis Control Prev, 2019, 23(6): 625-629. [程杨杨, 曹志, 侯洁, 等. 中国中老年人慢性病现状调查与共病关联分析[J]. 中华疾病控制杂志, 2019, 23(6): 625-629.]
- [17] Boeckxstaens P, Peersman W, Goubin G, *et al.* A practice-based analysis of combinations of diseases in patients aged 65 or older in primary care[J]. BMC Fam Pract, 2014, 15: 159.
- [18] Bieber V, Cohen AD, Freud T, *et al.* Autoimmune smoke and fire--coexisting rheumatoid arthritis and chronic obstructive pulmonary disease: a cross-sectional analysis[J]. Immunol Res, 2013, 56(2-3): 261-266.
- [19] Li YN, Wang YQ. Prevalence and patterns of multimorbidity among Chinese elderly people[J]. Chin Gen Pract, 2021, 24(31): 3955-3962, 3978. [黎艳娜, 王艺桥. 我国老年人慢性病共病现状及模式研究[J]. 中国全科医学, 2021, 24(31): 3955-3962, 3978.]
- [20] Roso-Llorach A, Violán C, Foguet-Boreu Q, *et al.* Comparative analysis of methods for identifying multimorbidity patterns: a study of 'real-world' data[J]. BMJ Open, 2018, 8(3): e018986.
- [21] Yao SS, Cao GY, Han L, *et al.* Prevalence and patterns of multimorbidity in a nationally representative sample of older Chinese: results from the China Health and Retirement Longitudinal Study[J]. J Gerontol A Biol Sci Med Sci, 2020, 75(10): 1974-1980.

(责任编辑: 蒋铭敏)