

复杂网络社区发现视角下慢性阻塞性肺疾病患者的 共病特征分析*

杨青青¹ 何航帆¹ 李莉芳² 王莹¹ 李少凡¹ 赵卉^{2△} 张岩波^{1,3△}

【摘要】目的 探讨不同性别慢性阻塞性肺疾病患者的共病模式,为慢性阻塞性肺疾病患者分层管理提供理论依据。**方法** 选取山西医科大学第二医院 2011 年 12 月至 2020 年 5 月慢性阻塞性肺疾病患者作为研究对象,横断面收集患者诊断信息建立共病网络,采用 Louvain 法对共病网络进行社区划分,并分析比较不同性别患者共病类型、组合及患病率。**结果** 研究表明不同性别患者单个慢性病和共病对的患病率存在差异;各患者群体对应的共病网络和社区结构也不同。**结论** 男性患者共病网络较女性患者共病网络更聚集和复杂,男性患者的共病模式比女性患者更复杂多样,疾病间联系更为紧密,在慢性阻塞性肺疾病患者中更应多关注男性患者,以期减少共病的发生。

【关键词】 慢性阻塞性肺疾病 共病 复杂网络 社区发现

【中图分类号】 R563 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.02.004

Analysis of Comorbidity Characteristics Found in Community under Complex Network of Patients with Chronic Obstructive Pulmonary Disease

Yang Qingqing, He Hangzhi, Li Lifang, et al (Department of Epidemiology and Biostatistics, School of Public Health Shanxi Medical University(030001), Taiyuan)

【Abstract】 Objective To explore the comorbidity patterns of different gender patients with chronic obstructive pulmonary disease (COPD) and provide theoretical basis for the hierarchical management of COPD patients. **Methods** The patients with chronic obstructive pulmonary disease from December 2011 to May 2020 in the Second Hospital of Shanxi Medical University were selected as the study subjects. The cross-sectional diagnostic information of patients was collected to establish a comorbidity network. The community of the comorbidity network was divided by Louvain method, and the comorbidity types, combinations and prevalence rates of patients with different genders were analyzed and compared. **Results** Our study manifested that there were differences in the prevalence of single chronic disease and comorbidity in patients with different genders. The corresponding comorbidity network and community structure of each patient group were also different. **Conclusion**

The comorbidity network of male patients is more concentrated and complex than that of female patients. The comorbidity pattern of male patients is more complex and diverse than that of female patients, and the relationship between diseases is closer. It verifies that the comorbidity pattern of different gender patients is different. In patients with chronic obstructive pulmonary disease, more attention should be paid to male patients in order to reduce the incidence of comorbidity.

【Key words】 COPD; Comorbidity; Complex network; Community discovery

慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)以气道和/或肺泡异常导致持续存在的呼吸道症状和进行性加重的气流受限为特征^[1],是一种可以预防和治疗的常见病。尽管 COPD 主要影响肺部,但也被认为是一种复杂的、多成分疾病,临床上将 COPD 患者广泛存在的共患慢性病如心脑血管疾病、消化系统疾病及精神疾病等多种疾病称为 COPD 共病^[2-3]。COPD 共病增加了患者的复杂性和治疗难度,给其家庭带来沉重经济负担^[4],探索患者共病模式有望改善干预策略,对患者的预后风险控制非常重要。

在以往对 COPD 的研究中,大多数研究仅考虑到两三种疾病组合的情况,并不能全面准确地探究 COPD 及其共病间的关系,考虑到 COPD 患者在性别上存在差别^[5],所以本研究对不同性别 COPD 患者的电子病历数据进行挖掘,利用复杂网络模型可视化 COPD 患者共病情况,探究不同性别 COPD 患者的共病模式差异,以期指导临床干预。

对象与方法

1. 研究对象

选取某院 2011 年 12 月至 2020 年 5 月诊断为 COPD 的患者构建对应共病网络,知情同意并严格按照纳入标准和排除标准选择研究对象。纳入标准:①根据 2021 年版 GOLD 诊断指南确诊为 COPD 的患者^[6];②年龄 ≥ 18 岁;③接受 COPD 药物及其他治疗措施。排除标准:①单病种患者;②病历诊断数据不全。为确保构建疾病网络的有效性和代表性,剔除:①重复的医疗记录;②患病率小于 1% 的患者数据^[7](本

* 基金项目:国家自然科学基金(82173631)中央引导地方科技发展专项资金项目(YDZX20191400004736)

1. 山西医科大学公共卫生学院流行病学与卫生统计学教研室(030001)

2. 山西医科大学第二医院呼吸与危重症医学科

3. 重大疾病风险评估山西省重点实验室(201805D111006)

△通信作者:张岩波, E-mail: sxmuzyb@126.com; 赵卉, E-mail: hui_zhao@sxmu.edu.cn

研究暂不考虑罕见疾病)。在完成数据处理步骤后,其他共病诊断按诊断指南进行确诊^[8]。

2.研究方法

(1)社区发现

网络系统由“节点(node)”和“边(edge)”组成,将某一系统的特征和信息以网络的形式呈现。社区发现(community detection)是指在网络图中根据一定规则将网络划分成若干个社区的过程,本质是对复杂网络图中节点的聚类,其方法包括图划分方法、基于模块度的方法、基于层次聚类的方法等^[9]。

本文采用基于模块度的 Louvain 算法对疾病网络进行社区发现。Louvain 算法基本思想是网络中节点尝试遍历所有邻居的社团,并选择最大化模块度增量的社区标签,在最大化模块度后,每个社区看成一个新的节点,重复直到模块度不再增大。Louvain 算法步骤如下:

步骤 1:模块度优化阶段。初始时将每个节点视为一个社团,然后对于任意一个节点*i*,考虑其邻居*j*,通过从节点*i*所属社团移除节点*i*,然后将其加入属于节点*j*的社团,再评估模块度收益。将孤立节点*i*移动到社团*m*模块度的增量Δ*Q*表达式为:

$$\Delta Q = \left[\frac{\sum in + S_i, in}{2e} - \left(\frac{\sum tot}{2e} \right)^2 \right] - \left[\frac{\sum in}{2e} - \left(\frac{\sum tot}{2e} \right)^2 - \left(\frac{S_i}{2e} \right)^2 \right] \tag{1}$$

式(1)中∑*in*为社团*m*内边权重的总和,∑*tot*为连接到社团*m*内节点的边的权重总和,*S_i*为连接到节点*i*的边的权重和,*S_i,in*为从节点*i*连接到社团*m*内节点的边权重和,*e*为网络中边权重的总和。

步骤 2:网络凝聚阶段。构建一个新网络,其节点属于第一阶段发现的社团,新节点之间的连边权重由相应的两个社团的节点连边权重和给定,再用第一阶段进行迭代循环,直到模块度不再增大。对于无权图所有边权重设为 1,用节点的度进行计算。

(2)复杂网络的描述统计

本文使用严格局部计算法对疾病网络基础特性进行度量,指标包括密度、节点数;采用全局计算法对网络聚集结构进行度量,指标包括聚集系数、边数量、社区数、边介数。*V*表示节点数量,*E*表示连边数量,*ε*表示节点集。

①密度:网络的密度主要用来衡量网络中各节点间的连接强度,*D*取值区间为[0,1]。公式表示为:

$$D = \frac{E}{\binom{V}{2}} = \frac{2E}{V(V-1)} \tag{2}$$

式(2)中 $\binom{V}{2}$ 表示在网络中节点所有可能连接的方案数,它是从网络中所包含的*V*个节点中任意选择 2 个节点进行组合。当密度 *D*=0 时,此时网络是一个零

图;当 *D*=1 时,此时网络是一个完全图。

②聚集系数:用于度量网络中节点聚集程度的系数,*CC*∈[0,1]。表达式为:

$$CC_i = \frac{2|e_i|}{k_i(k_i-1)}; CC = \frac{1}{V} \sum_{i \in \epsilon} CC_i \tag{3}$$

式(3)中 *CC_i* 表示节点 *i* 的聚类系数,|*e_i*|表示节点 *i* 的邻域内节点间的连接边数(即由节点 *i* 以及其领域内的两个节点形成的三角形数量),*k_i*表示节点 *i* 的度数。如果 *CC*=1,说明网络中所有点都是相连的,而如果 *CC*趋近于 0,说明网络的连接较为松散。

③边介数:用于度量网络中每对节点位于最短路径上的程度,边介数 *B_ω* 值越大表示这对节点在网络信息传递过程中可能有更大的影响力。表达式为:

$$B_\omega = \sum_{i,j \in V} \frac{\eta_{ij}^\omega}{\eta_{ij}} \tag{4}$$

式(4)中 η_{ij}^ω 表示从节点 *i* 到节点 *j* 的最短路径中经过节点 ω 的最短路径数, η_{ij} 为节点 *i* 到节点 *j* 的所有可能的最短路径数。

3.统计分析 with 网络构建

在完成数据清理步骤后,采用 R 4.1.1 进行统计描述:计量资料采用 $\bar{x} \pm s$ 表示,计数资料采用 *n*(%) 表示,使用 χ^2 检验差异,检验水准 $\alpha = 0.05$ 。将纳入的疾病整理为点数据,疾病与疾病间共患情况整理为边数据,通过 R 中“igraph”包构建基于点数据和边数据的无权无向网络及社区发现;使用 Cytoscape 3.8.2 软件进行网络结构的可视化和参数呈现。

结 果

1.基本信息

本研究共纳入 3497 例患者病历数据,其中男性患者 2683 人,占比 76.72%,女性患者 814 人,占比 23.28%;平均年龄 73.72±11.07 岁,共病个数为 2~7 的患者最多,占比 73.83%(如表 1,其中共病数不包括 COPD)。

表 1 全部研究对象一般人口学分布(%)

基本信息		研究对象(<i>n</i> = 3497)
性别	男	2683 (76.72)
	女	814 (23.28)
年龄(岁)	18~39	12 (0.34)
	40~64	690 (19.73)
	65~	2795 (79.93)
吸烟史	有	3431 (98.11)
	无	66 (1.89)
共病数	2~7	2582 (73.83)
	8~13	853 (24.39)
	14~19	62 (1.78)

2.不同性别共病情况

在纳入的 37 种慢性病中有 12 种单个慢性病患病率在性别上具有统计学差异,其中男性后天性肾囊肿、肝囊

肿、高同型半胱氨酸血症的患病率高于女性(表 2)。

表 2 不同性别研究对象单个慢性病患病率比较(%)

疾病	男 (n=2683)	女 (n=814)	χ^2	P
高血压	843(31.42)	343(42.14)	32.007	<0.001
心力衰竭	459(17.11)	166(20.39)	4.593	0.032
哮喘	383(14.28)	163(20.02)	15.669	<0.001
后天性肾囊肿	304(11.33)	55(6.76)	14.183	<0.001
2 型糖尿病	213(7.94)	121(14.86)	34.678	<0.001
肝囊肿	278(10.36)	48(5.90)	14.727	<0.001
胆囊结石	157(5.85)	66(8.11)	5.326	0.021
继发性肥胖	97(3.62)	46(5.65)	6.599	0.010
高脂血症	58(2.16)	28(3.44)	4.253	0.039
甲状腺功能异常	46(1.71)	23(2.83)	3.986	0.046
高同型半胱氨酸血症	37(1.38)	4(0.49)	4.247	0.039

在所有研究对象中,除 COPD 外,共病对患病率前 10 位多为高血压和肺部感染的组合(表 3)。其中共病对高血压 & 肺部感染患病率最高,为 8.15%;高血压 & 肺部感染、高血压 & 心力衰竭、高血压 & 冠心病、高血压 & 2 型糖尿病、高血压 & 哮喘、高血压 & 脂肪肝、高血压 & 电解质紊乱与肺部感染 & 胸腔积液的性别患病率差异有统计学意义。

3. 共病网络分析

(1) 网络基础特性比较

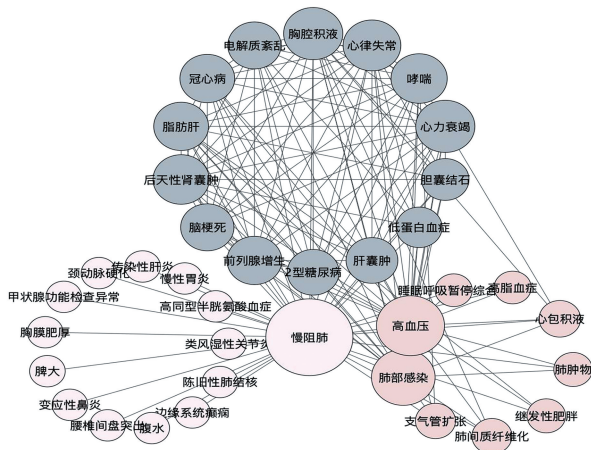


图 1 全部研究对象共病网络

表 3 不同性别前 10 位共病对患病率差异性检验(%)

共病模式	男 (n=2683)	女 (n=814)	χ^2	P
高血压 & 肺部感染	194(7.23)	91(11.18)	13.008	<0.001
高血压 & 心力衰竭	149(5.55)	85(10.44)	23.906	<0.001
高血压 & 冠心病	159(5.93)	65(7.99)	4.417	0.036
高血压 & 心律失常	156(5.81)	59(7.25)	2.225	0.136
高血压 & 脑梗死	160(5.97)	47(5.77)	0.040	0.841
高血压 & 2 型糖尿病	122(4.55)	85(10.44)	38.973	<0.001
肺部感染 & 胸腔积液	155(5.78)	46(5.65)	0.018	0.892
高血压 & 哮喘	134(4.99)	61(7.49)	7.410	0.006
高血压 & 脂肪肝	121(4.51)	51(6.27)	4.115	0.042
高血压 & 电解质紊乱	104(3.88)	50(6.14)	7.619	0.006

在构建的三个疾病网络中(如图 1、2、3,每种颜色代表一个社区),“节点”代表慢性病,其大小与患病率高低成正比;“边”表示慢性病的共患情况。全部研究对象共病网络被划分为 3 个社区,最大的社区包含 14 种疾病,最小的社区包含 9 种疾病,呼吸系统疾病易被划分到一个社区,心血管疾病易被划分到一个社区,高血压所在社区的慢性病种类最少。男性 COPD 患者共病网络与女性 COPD 患者共病网络均划分为了 4 个社区,女性中最大的社区包含的慢性病种类多于男性中最大的社区,男性共病网络明显比女性共病网络更稠密。

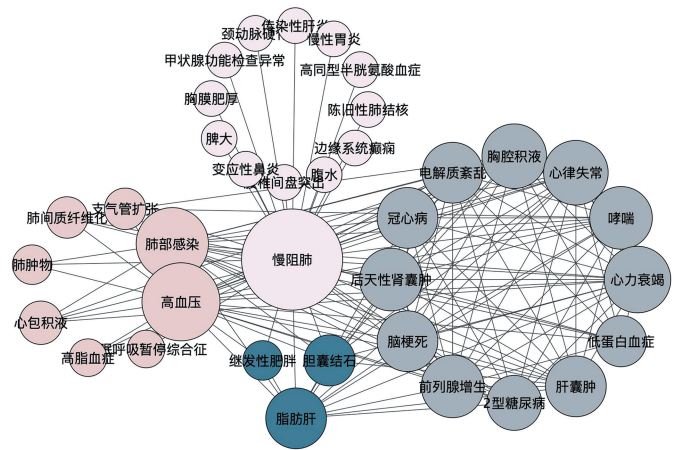


图 2 男性 COPD 患者共病网络

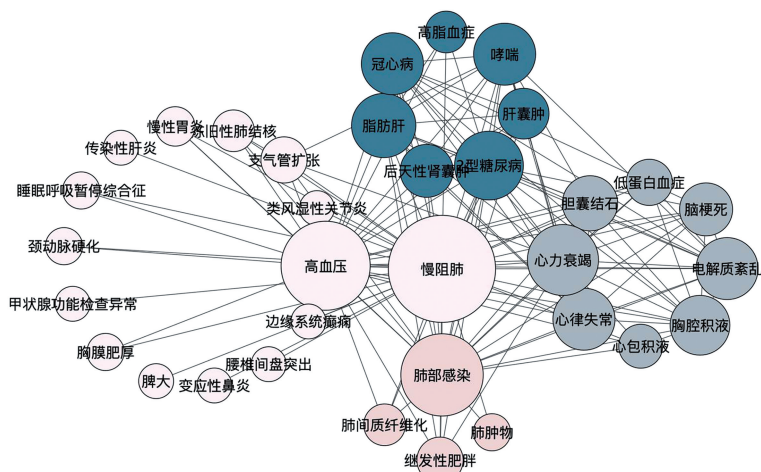


图 3 女性 COPD 患者共病网络

对所建立的三个共病网络进行基本结构特征比较,结果显示,男性的共病网络聚集系数和网络密度最大,分别为 0.864 和 0.243(如表 4)。

表 4 共病网络的特征参数统计

网络	类型	V	E	D	CC	社区数
全部研究对象共病网络	无向	37	152	0.228	0.860	3
男性 COPD 患者共病网络	无向	36	153	0.243	0.864	4
女性 COPD 患者共病网络	无向	34	140	0.210	0.845	4

(2) 不同性别的共病网络分析

使用边介数值来发现共病网络中重点疾病对,截取前 5 位共病对,结果显示男性 COPD 患者更倾向出现心力衰竭 & 前列腺增生的问题,女性 COPD 患者更倾向于出现 2 型糖尿病 & 冠心病的问题(表 5)。

表 5 不同性别共病网络边介数

性别	共病模式	边介数
男性	高血压 & 呼吸睡眠暂停综合征	23.00
	高血压 & 高脂血症	23.00
	高血压 & 肺间质纤维化	13.67
	高血压 & 心包积液	11.83
	心力衰竭 & 前列腺增生	6.96
女性	高血压 & 呼吸睡眠暂停综合征	25.00
	高血压 & 高脂血症	17.17
	高血压 & 肺间质纤维化	15.17
	高血压 & 心包积液	13.97
	2 型糖尿病 & 冠心病	8.50

讨 论

及时识别 COPD 共病以及准确判断其对患者的影响,制定出 COPD 和共病个体化干预方案,对患者的风险防控十分重要^[10]。目前大量 COPD 共病研究只是关注到某种特定的疾病与 COPD 的共病情况,而不是着眼于人群中存在的复杂共病网络。与其他共病研究相比,本文所构建的共病网络可以直观地反映出各疾病间错综复杂的联系,从整体将疾病进行分区获得共病模式,缩小慢性病的行为干预范围,从而使疾病预防策略更精准有效。

本研究结果显示,不同性别的 COPD 患者在单个慢性病和共病对患病率上都存在差异,男性 COPD 患者共病种类多于女性 COPD 患者。众多研究也证明女性 COPD 患者的患病危险因素、症状严重程度及肺功能的临床特点有异于男性^[11-12],造成不同性别 COPD 患者共病种类和患病率的差异。与女性相较而言,男性吸烟率和过量饮酒率较高是其患慢性肺部疾病、肝脏疾病、肾脏疾病患病率高于女性的重要原因。女性高血压、心脏病、血脂异常、糖尿病或血糖升高及肥胖的患病率高于男性,可能与女性群体特有的风险因素有关,如绝经、妊娠增加了其罹患上述慢性病的可

能性^[13]。这些相关因素导致了 COPD 男女患者在临床表现、疾病进展、对治疗的反应性等方面均存在明显的异质性^[14],学界也将性别作为临床表型分型的一个重要影响因素。

本研究还显示男性患者共病网络较女性患者共病网络更聚集和复杂,疾病间联系更为紧密。网络分析显示男性 COPD 患者重点共病对为心力衰竭 & 前列腺增生,有流行病学和临床研究表明,前列腺疾病和心血管疾病危险因素之间并存^[15];女性 COPD 患者重点共病对为 2 型糖尿病 & 冠心病,先前的研究表明心力衰竭大多发生于老年及女性患者,常合并糖尿病、冠心病、心律失常等^[16]。处于同一社区的慢性病可能由于其存在潜在的共同的患病机制或存在共同的危险因素,也就是说,出现某社区内慢性病后与该社区内其他慢性病共病的可能性更大。

由于本文为横断面研究,无法明确疾病间的因果关系,因此后续研究中,应考虑时序因素建立动态网络,在真实世界复杂网络中的动态社区结构检测对于测量网络结构和性质具有更为重要的临床意义。

共病网络之间的差异证实了共病模式的异质性。为了解疾病关系、降低共病复杂性、简化临床管理、制定共病指南等提供参考,目前有许多学者试图探索人群共病患病率、影响因素和共病模式^[17]。本研究通过共病网络能够较好地归纳 COPD 患者共病模式,并且进一步分析不同性别共病网络中出现的频次和顺位以发现共病中的重点共病对,可为分层管理 COPD 患者提供新的决策视角,以期有效提高人群中 COPD 共病的预防和治疗质量。

参 考 文 献

[1] Laura F, Aaron D. Chronic obstructive pulmonary disease(COPD). The Lancet Respiratory Medicine, 2017, 5(1):18-19.

[2] Negewo N, Gibson G, McDonald V. COPD and its comorbidities: Impact, measurement and mechanisms. Respirology, 2015, 20(8): 1160-1171.

[3] Ben N, Sue J, Andrew M, et al. Comorbidities and Medication Burden in Patients With Chronic Obstructive Pulmonary Disease Attending Pulmonary Rehabilitation. Journal of Cardiopulmonary Rehabilitation and Pre, 2014, 34(1):75-79.

[4] Ho TW, Tsai YJ, Huang CT, et al. Impact of tobacco-related chronic obstructive pulmonary disease on developmental trajectories of comorbidities in the Taiwan population. Scientific Reports, 2020, 10(1):21025.

[5] 刘伟,贾爱芹,崔欣宏,等.2011-2015 年某三甲医院出院病人前十位病种分析.中国卫生统计,2017,34(5):812-813.

[6] 中华医学会呼吸病学分会慢性阻塞性肺疾病学组,中国医师协会呼吸医师分会慢性阻塞性肺疾病工作委员会.慢性阻塞性肺疾病诊治指南(2021 年修订版).中华结核和呼吸杂志,2021,44(3): 170-205.