

· 临床研究 ·

小细胞肺癌中医智能辨证探索性研究

臧晓彤 张培彤[▲]

摘要 **目的:**提取小细胞肺癌中医证素,初步探索建立小细胞肺癌中医智能辨证模型。**方法:**采用SPSS 23.0进行因子分析,采用SPSS Modeler 18.0建立MLP神经网络智能辨证模型。**结果:**①共收集到43个症状条目,KMO检验统计量为0.648,Bartlett的球形度检验近似卡方值为2332.572($P < 0.05$)。因子分析提取到7个公因子,结合临床经验总结出相应的证素,总结为7个核心单证,分别为肺血瘀、肺阴虚、心阴虚、胃阴虚、肺气虚、脾气虚、肾阳虚。②建立的小细胞肺癌智能辨证模型训练集正确率为92.63%,测试集正确率为62.79%。**结论:**因子分析可较为准确地提取小细胞肺癌中医证素,在此基础上通过神经网络算法建立的智能辨证模型能够提升小细胞肺癌中医辨证效率和精准度。

关键词 因子分析;小细胞肺癌;证素;单证;智能辨证;机器学习;神经网络

小细胞肺癌(small cell lung cancer, SCLC)是一种高度侵袭性疾病,占有肺癌的15%,其中位生存期小于2年,是最致命的癌症之一^[1-2]。在过去的几十年中,小细胞肺癌的治疗方法进展缓慢,现有的疗法未能延长晚期小细胞肺癌患者的生存期^[3]。临床研究^[4-7]显示,中医药干预可以有效减轻化疗给小细胞肺癌患者带来的不良反应,改善临床症状,提高生存质量及治疗效果。

辨证论治是中医临床实践的核心,信息化背景下借助新算法实现智能辨证对提升SCLC中医诊疗效率及精度具有重要意义。目前,中医智能辨证研究^[8-12]对症状、证候的规范化不足,辨证结果在术语表达上没有完全统一,研究结果间的可比性差。笔者研究团队前期探索建立了基于证素理论的证候分层诊断体系^[13-14],为本研究奠定了理论基础。证素是构成中医证候的基本要素,找准核心证素能帮助临床医生实现精准辨证论治,并更好地评估疗效。因此,如何精准的提取证素是中医智能辨证的关键。中医智能辨证能敏感地捕捉患者病程中证候的变化,精准评价治疗前后单证的改善与否,有助于提高SCLC的中医诊疗效率。

1 资料与方法

1.1 一般资料 纳入病例为2018年6月1日—2020年6月1日于中国中医科学院广安门医院肿瘤科就诊的138例经病理确诊为SCLC的患者。患者平均年龄(59.50 ± 9.74)岁;男性占76.09%,女性占23.91%;局限期81例,广泛期34例,另有23例病例就诊资料不完整,无法确定分期。

1.2 诊断标准

1.2.1 西医诊断标准 采用《原发性肺癌诊疗规范(2018年版)》^[15]中SCLC的诊断标准。

1.2.2 中医证型诊断标准 参考中医证候分层诊断标准^[13-14],将八纲辨证、脏腑辨证、气血津液辨证相结合,借助证素理念,采用单证的研究方法,将证候分层至不能再进一步划分。

1.3 纳入标准 ①符合上述SCLC诊断标准,有明确病理诊断的SCLC患者;②就诊时无其他影响症状表现的疾病者。

1.4 排除标准 ①有其他恶性疾病者;②既往有其他恶性肿瘤病史者;③认知障碍或依从性差,不能配合量表评价者。

1.5 调查方法 采用横断面调查法,由中医肿瘤专业的硕士研究生对就诊的研究对象进行现场调查,收集刻下症信息,并经肿瘤科硕士研究生导师确认补充。

[▲]通信作者 张培彤,男,医学博士,主任医师。研究方向:肺癌的证候研究。E-mail: drzhangpeitong@hotmail.com

• 作者单位 中国中医科学院广安门医院(北京 100053)

1.6 症状术语规范方法 结合既往研究成果及专家意见,初步确定由病位证素和病性证素组合成单证名称。采取《中医临床常见症状术语规范修订》^[16]的症状描述,采用课题组时美伶研究的相似症状归类法^[17]进行整理,共计129个标准化后症状纳入症状术语库。

1.7 统计学方法 应用SPSS 23.0软件进行统计学处理和分析,采用因子分析法提取证候要素。

1.8 建模方法 将提取的公因子按证素整理合并,形成138例患者的核心单证诊断结果,采用SPSS modeler 18.0的MLP神经网络算法,按照7:3的比例自动拆分训练集和测试集,构建智能辨证模型。

2 结果

2.1 信度 对数据进行检验变量间偏相关性的KMO (Kaiser-Meyer-Olkin)检验,统计量为0.648,变量相关性较强,适合做因子分析。Bartlett的球形度检验近似卡方值为2332.572 ($P < 0.05$)。拒绝变量间独立性的假设,适合做因子分析。

2.2 频率分析 对138例患者的1553条症状进行整理和术语规范,得到标准化后的中医症状129个,再对其进行频率分析。为了集中反映SCLC患者的证候特征,选取频率>5%的症状进行分析。其中脉沉、咳嗽的频数分别为114和104,为SCLC患者的共性表现,对证素区分意义不大,故剔除,共保留典型症状43个。见表1。

2.3 因子分析 对上述43个症状进行因子分析。运用因子分析中的主成分分析法提取公因子,其中特征根>2的公因子有7个(用f表示)。用方差最大旋转法求出因子成分矩阵。见表2。

2.4 证候要素提取 在上述结果取载荷>0.5为遴选阈值,筛选出SCLC患者的常见症状,结合临床经验,总结相关证素。见表3。

2.5 智能辨证模型 对依据43个关键症状得出的7个核心单证构建神经网络模型,隐藏层有14个神经元节点,见图1。模型训练集正确率为92.63%,测试集正确率为62.79%,用于模型评估的受试者操作特征曲线,见图2。

3 讨论

证候是疾病发展过程中某一阶段的病理概括,反映了疾病的本质,是中医临床的核心。证素^[18]是构成

表1 SCLC患者典型症状分布表

症状	频数	频率(%)	症状	频数	频率(%)
苔黄	84	60.87	咽干	25	18.12
咳痰	60	43.48	口干	24	17.39
背痛	47	34.06	鼻干	23	16.67
脉弦	44	31.88	恶寒	23	16.67
舌暗紫	43	31.16	便干	22	15.94
胁痛	40	28.99	声嘶	22	15.94
胸痛	38	27.54	心烦	22	15.94
乏力	37	26.81	白痰	21	15.22
苔白	36	26.09	舌尖红	20	14.49
胸闷	36	26.09	畏寒	20	14.49
脉无力	34	24.64	盗汗	18	13.04
舌有瘀斑	33	23.91	舌少津	17	12.32
脉数	32	23.19	齿痕舌	15	10.87
纳差	32	23.19	自汗	14	10.14
气短	31	22.46	黄痰	12	8.70
痰中带血	31	22.46	面色少泽	11	7.97
干咳	30	21.74	舌红	9	6.52
舌暗红	29	21.01	夜尿多	8	5.80
腰痛	29	21.01	便秘	7	5.07
便溏	26	18.84	恶心	7	5.07
脉细	26	18.84	手足麻木	7	5.07
脉滑	25	18.12			

中医证候的基本要素,包含病位、病性证素组成的单证及由二个或多个单证组合成的复证。本研究通过对SCLC患者的中医症状进行因子分析提取公因子,结合专家经验分析总结核心单证并构建智能辨证模型,为单证的量化及复合证候的研究工作奠定基础。

因子分析法^[19]是指从研究指标相关矩阵内部的依赖关系出发,把一些信息重叠、关系复杂的变量归结为少数几个不相关的综合因子的一种多元统计分析方法,能够在保留尽可能多的原始信息的基础上,减小数据噪音的影响。在医学领域常用于通过多个可观测的指标之间的相关性,寻求有限的不可观测的潜在变量,例如通过对可以观测的症状进行分析以提取不可直接观测的证素。中医认为SCLC的病位主要在肺,可累及脾、肾^[20],其病机主要在“虚”和“瘀”^[21]两方面。笔者提取出SCLC的7个单证,分别为肺血瘀、肺阴虚、心阴虚、胃阴虚,肺气虚、脾气虚、肾阳虚,符合中医临床对肺癌病位、病机的认识,对中医临床诊断具有实际意义。

人工神经网络^[22]是在模拟人脑信息处理方式的基础上建立起来的一种具有快速数据处理能力的非线性动力系统,能够通过样本的学习,自动提取数

表2 SCLC 患者典型症状最大方差旋转公因子成分矩阵表

症状	公因子						
	f1	f2	f3	f4	f5	f6	f7
便溏	-0.12	-0.09	0.19	-0.17	-0.12	-0.72	0.13
手足麻木	0.05	0.02	0.05	0.30	-0.12	-0.21	-0.02
舌红	0.13	0.10	0.02	0.01	-0.10	-0.12	-0.15
脉细	-0.06	0.07	-0.79	-0.12	-0.06	0.10	0.01
脉弦	0.66	-0.06	0.06	0.08	-0.12	-0.01	0.12
脉数	-0.09	-0.09	0.05	0.72	-0.03	0.04	-0.03
乏力	-0.18	-0.11	0.06	-0.18	0.59	-0.04	0.01
气短	-0.11	0.02	0.05	-0.19	0.64	0.02	0.07
恶寒	0.03	-0.17	0.21	0.14	0.18	0.17	-0.40
胁痛	0.58	-0.09	0.12	-0.28	-0.12	0.17	-0.20
腰痛	-0.19	-0.03	0.17	0.01	-0.05	-0.08	-0.65
苔白	-0.23	-0.27	0.22	-0.25	-0.45	0.41	-0.00
脉无力	-0.05	-0.09	-0.06	-0.12	0.04	0.13	-0.67
白痰	0.08	-0.30	0.15	-0.01	0.25	0.00	0.27
胸闷	-0.11	-0.18	0.27	-0.13	0.06	-0.06	0.40
脉滑	-0.29	-0.08	0.35	-0.13	-0.07	0.18	0.39
纳差	-0.23	-0.04	0.08	-0.21	-0.04	-0.70	0.14
舌暗红	-0.02	0.06	0.21	0.01	-0.14	0.13	-0.33
齿痕舌	-0.18	-0.13	0.15	-0.18	-0.09	-0.74	0.08
痰中带血	0.70	-0.25	-0.07	-0.14	-0.15	0.08	0.01
胸痛	0.70	-0.01	0.19	0.03	-0.06	0.04	0.08
背痛	0.59	-0.17	0.22	-0.23	-0.18	-0.04	-0.10
舌暗紫	0.51	-0.03	0.01	-0.06	0.26	0.02	0.28
舌有瘀斑	0.79	-0.03	0.04	-0.01	-0.01	0.03	0.03
盗汗	-0.11	-0.09	-0.82	-0.13	-0.05	0.09	-0.01
心烦	-0.13	-0.03	-0.81	0.01	-0.05	0.14	0.08
舌尖红	-0.20	-0.07	-0.83	-0.13	-0.12	0.10	0.05
口干	-0.04	0.07	0.05	0.78	-0.10	0.12	-0.03
舌少津	-0.06	-0.03	0.12	0.71	-0.12	0.09	0.01
干咳	-0.14	0.84	-0.02	0.02	-0.06	0.00	0.04
自汗	-0.16	-0.21	0.19	-0.10	0.75	0.22	0.05
畏寒	-0.10	-0.21	0.07	-0.04	-0.07	0.01	-0.59
声嘶	0.01	0.80	-0.01	-0.11	0.02	0.02	0.14
咽干	-0.13	0.86	0.06	-0.04	-0.04	0.12	0.05
鼻干	-0.06	0.82	0.06	-0.03	-0.06	0.07	0.06
便秘	-0.08	0.07	0.21	-0.27	-0.04	0.33	0.18
夜尿多	0.25	-0.00	-0.02	0.18	0.01	0.11	0.02
便干	-0.08	0.01	-0.02	0.77	-0.14	0.17	-0.03
黄痰	-0.03	-0.25	0.21	-0.18	-0.07	-0.01	0.33
恶心	-0.02	0.16	-0.29	0.05	-0.09	-0.15	0.06
面色少泽	-0.06	-0.05	0.09	-0.10	0.69	0.14	0.02
咳痰	0.29	-0.53	0.12	-0.16	0.07	0.09	0.30
苔黄	0.26	0.29	-0.08	0.18	0.51	-0.44	-0.06

表3 SCLC 患者证素提取表

公因子	包含症状	病位	病性
f1	脉弦(0.66), 胁痛(0.58), 痰中带血(0.70), 胸痛(0.70), 背痛(0.59), 舌暗紫(0.51), 舌有瘀斑(0.79)	肺	血瘀
f2	干咳(0.84), 声嘶(0.80), 咽干(0.86), 鼻干(0.82)	肺	阴虚
f3	脉细(0.79), 盗汗(0.82), 心烦(0.81), 舌尖红(0.83)	心	阴虚
f4	脉数(0.72), 口干(0.78), 舌少津(0.71), 便干(0.77)	胃	阴虚
f5	乏力(0.59), 气短(0.64), 自汗(0.75)	肺	气虚
f6	便溏(0.72), 纳差(0.70), 齿痕舌(0.74)	脾	气虚
f7	腰痛(0.65), 脉无力(0.67), 畏寒(0.59)	肾	阳虚

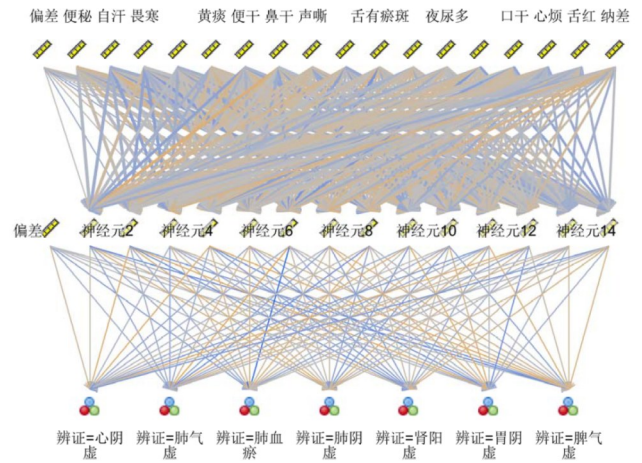
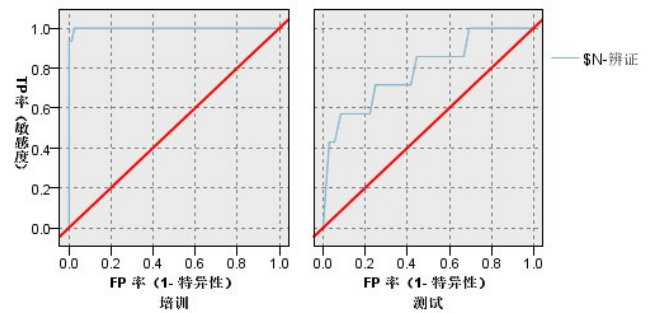


图1 SCLC 患者核心单证神经网络模型图



训练集曲线下面积AUC=0.998, P<0.05; 测试集曲线下面积AUC=0.778, P<0.05

图2 受试者操作特征曲线图

据中的隐藏规律,并根据这些规律自动对新样本进行归类。中医临床辨证是在症状、体征的基础上形成对证候的判断,因此可以认为,临床经验的积累与在人大脑中建立症状、证候之间的映射关系在一定程度上是相同的。因此,采用人工神经网络建模能够模拟复杂

的中医辨证逻辑。此次建模的受试者操作特征曲线显示模型拟合度欠佳,可能的原因是在对因子分析进行降维的过程中造成了信息的失真。此外,患者通常会同时具备多个单证,且各单证之间存在一定的关联,而各目标变量之间的关联性也会影响模型准确度。本研究条件有限,样本量较小,且目前只能提取出患者最为主要的单证,还无法对各单证在复合证候中所占的比重进行分析,后续研究中笔者将不断地探索改进。

综上所述,因子分析可较为准确地提取 SCLC 中医核心证素,结合既往研究成果,能够基本实现中医临床辨证的客观化,在此基础上通过神经网络算法建立的智能辨证模型能够提升 SCLC 中医辨证效率和精度。今后可以通过大样本、多中心的研究,纳入足够的样本,引入更优的算法,增加模型可靠性,为智能辨证模型的临床应用奠定基础。此外,若能在本研究结果中核心单证的基础上加入各兼证并分别量化赋值,通过症状及证候的量化研究,可使中医临床辨证进一步实现精准化和客观化,如此,SCLC 的精准中医诊疗未来可期。

参考文献

[1] DENEKA AY, BOUMBER Y, BECK T, et al. Tumor-targeted drug conjugates as an emerging novel therapeutic approach in small cell lung cancer (SCLC)[J]. *Cancers (Basel)*, 2019, 11(9): 1297.

[2] 邓宇,郝博,耿庆.小细胞肺癌治疗现状及展望[J].*临床外科杂志*, 2020, 28(7): 696-699.

[3] RECK M, HEIGENER D, REINMUTH N. Immunotherapy for small-cell lung cancer: emerging evidence [J]. *Future Oncol*, 2016, 12(7): 931-943.

[4] DENG C, LOU Y, GAO Y, et al. Efficacy and safety of Shengjiang Xiexin decoction in prophylaxis of chemotherapy-related diarrhea in small cell lung cancer patients: study protocol for a multicenter randomized controlled trial[J]. *Trials*, 2020, 21(1): 370.

[5] 呼永华,张慧芳.参苓白术散对小细胞肺癌患者行EP方案化疗胃

肠道反应的疗效观察[J].*甘肃医药*, 2013, 32(11): 817-818.

[6] 苏春雨,宋卓,徐竟男,等.健脾补肾法防治小细胞肺癌转移[J].*中医学报*, 2016, 31(5): 621-624.

[7] 徐筱青,王达洋,张培彤.小细胞肺癌中西医结合治疗研究概述[J].*天津中医药*, 2017, 34(11): 789-792.

[8] 赵汉青.基于数据驱动的中医精准诊疗模型研究[D].北京:中国中医科学院, 2019.

[9] 周晟.基于案例推理的糖尿病肾病中医临床决策辅助系统的设计与构建[D].北京:北京中医药大学, 2018.

[10] 韦昌法,罗丽琴,晏峻峰.中医数字辨证配套医案智能采集与分析系统构建研究[J].*湖南中医药大学学报*, 2020, 40(1): 70-74.

[11] 鲍平波.运用Python优化中医辨证诊疗系统的思考[J].*中医药管理杂志*, 2020, 28(4): 61-62.

[12] 杨佳澄.基于群体智能的中医辨证诊断研究[D].兰州:兰州交通大学, 2019.

[13] 王笑丹.原发性支气管肺癌虚证分级分层量化诊断标准的探索性研究[D].北京:北京中医药大学, 2016.

[14] 王笑丹,张培彤.中医证候分层诊断模式及标准初探[J].*中华中医药杂志*, 2017, 32(3): 1209-1213.

[15] 中国政府网.关于印发原发性肺癌等18个肿瘤诊疗规范的通知:国卫办医函[2018]1125号[EB/OL].(2018-12-21)[2018-12-21].<http://www.nhc.gov.cn/ewebeditor/uploadfile/2018/12/20181225162102206.docx>.

[16] 黎敬波,马力.中医临床常见症状术语规范·修订[M].北京:中国医药科技出版社, 2015: 1-328.

[17] 时美伶.恶性肿瘤相关中医症状术语规范化的探索性研究[D].北京:中国中医科学院, 2018.

[18] 朱文锋.证素辨证学[M].北京:人民卫生出版社, 2008: 1-318.

[19] 孙振球,徐勇勇.医学统计学[M].北京:人民卫生出版社, 2014: 333-341.

[20] 陈红,杨永,孙旭,等.抗癌解毒、化痰祛瘀治疗小细胞肺癌[J].*中医杂志*, 2018, 59(23): 2061-2063.

[21] 岳斌,李占林,何晓华.从扶正化瘀探讨小细胞肺癌的临床治疗[J].*环球中医药*, 2017, 10(8): 901-903.

[22] 赵铁牛,于春泉,王惠君,等.人工智能在中医证候学中的应用初探[J].*中华中医药杂志*, 2014, 29(3): 831-833.

(收稿日期:2021-12-19)

(本文编辑:金冠羽)