

DOI: 10.13288/j.11-2166/r.2018.06.006

关于建立高血压前期人群高血压发病病证结合风险预测模型的思考

王丽颖, 韩学杰*

(中国中医科学院中医临床基础医学研究所, 北京市东城区东直门内南小街 16 号, 100700)

摘要 早期辨识可能发展为高血压病的高血压前期人群, 有针对性地干预, 可降低高血压患病率, 是疾病防治重心前移的战略需求。目前高血压前期人群发生高血压病的预测研究有一定局限性。提出通过收集高血压前期人群的危险因素及四诊信息, 通过数据分析, 筛选出高血压前期发展为高血压病的高危危险因素和中医证候要素, 采用神经网络、支持向量机、决策树等多重方法建立模型, 通过对模型验证、评价, 构建危险因素和中医证候相结合的风险预测模型, 为早期筛查可能发展为高血压病的高血压前期人群提供科学依据。

关键词 高血压前期; 高血压病; 风险预测模型; 病证结合

防治战略前移是当前临床医学的发展趋势。2010 年美国心脏协会 (AHA) 战略规划工作小组把理想的健康行为和健康因素定义为心血管健康^[1]。国内外学者指出, 高血压是心血管疾病的第一战场, 而高血压前期又是高血压病的第一战场^[1-4]。高血压前期由《美国预防、检测、评估与治疗高血压全国联合委员会第七次报告》(JNC7)^[5]首次提出, 是指收缩压为 120 ~ 139 mmHg、舒张压为 80 ~ 89 mmHg 者。国内高血压前期患病率较高, 调查发现, 除个别地区患病率较低外, 其余地区高血压前期患病率为 41.5% ~ 43.9%^[6]。与正常血压人群相比, 高血压前期人群未来发生高血压病的概率和发生心血管疾病的可能性大大增加。北京地区流行病学调查显示, 10 年前正常血压和高血压前期的亚组人群 10 年后发生高血压的比例分别为 22.2% 和 52.6%^[7]。日本一项调查数据显示, 从高血压前期进展到高血压病的受试者发生心血管疾病的风险比保持正常血压或高血压前期的风险高 2.95 倍^[8]。《中国高血压防治指南》^[9]将高血压前期定义为正常高值血压, 其与 JNC7 指南针对该病的建议均为以生活方式改变为主。近年来研究发现, 针对高血压前期人群, 给予膳食添加剂或中医

辨证治疗、穴位埋线等方法, 可以有效地预防或延缓高血压前期向高血压的改变^[10-12]。但是, 高血压前期患病率高、血压跨度大、危险度各异, 无选择性地给予干预将浪费巨大的医疗资源。因此, 识别可能发展为高血压病的高危高血压前期人群, 预测未来发展为高血压病的风险, 有针对性地进行干预显得尤为重要。

1 高血压前期人群早期识别和风险预测的研究现状及存在问题

有学者采用构建数学模型的方法, 寻找高血压前期人群发生高血压病的高危因素及预测因子, 并预测高血压前期人群发展为高血压病的风险。如采用 Cox 比例风险回归分析, 寻找高血压前期人群发生高血压病的独立危险因素^[13-14]。采用多元线性回归探索 C 反应蛋白对高血压前期人群进展至高血压的预测价值^[15]。这种通过建立数学模型, 对未来可能发生的风险状态、影响因素和变化趋势所作的分析和推断的方法, 即为风险预测, 构建的数学模型称为风险预测模型。近年来, 风险预测用于研究致病危险因素与慢性病发病率及病死率之间数量依存关系及其规律性, 将生活方式、理化指标等因素转化为可测量的指标, 用于预测个体在未来一定时间发生疾病或死亡的危险。如刘建平等^[16]采用分类树模型预测缺血性脑卒中发病风险, 正确率

基金项目: 北京中医药科技发展资金 (JJ2015-61); 中国中医科学院基本科研业务费自主选题资金 (Z0549)

* 通讯作者: xuejieh@126.com; (010) 64093238

为 79.30%。文玉娜等^[17]采用多因素 Logistic 回归分析,预测乙型肝炎后肝硬化患者死亡概率,结果显示预测的敏感度为 91.5%,特异度为 92.4%,正确指数为 0.839。杨召等^[18]采用 ARIMA-BPNN 组合模型预测流感发病率,结果显示,ARIMA-BPNN 组合模型预测的平均绝对误差为 0.029、平均误差绝对率为 0.43%。

这些风险预测模型的建立能够较好地筛查疾病高危人群、预测疾病的发病情况,但目前开展的相关研究在预测高血压前期人群发生高血压病的风险时存在一定的局限性。一方面,目前针对高血压前期人群发展为高血压病的预测模型主要采用回归统计方法如 Logistic 回归、Cox 逐步回归等,虽然具有一定的预测能力,但这些模型要求变量满足独立性且不能处理变量之间的共线性问题,因此,用于预测高血压病这样的复杂疾病具有一定的局限性^[19]。另一方面,这些研究都是基于现代医学研究的危险因素建立风险评估工具,缺少高血压前期中医证候学方面的研究内容。此外,也有学者提出采用基于健康管理队列构建高血压病风险预测模型,预测高血压前期人群发展为高血压病的风险^[20]。这种方法也存在一定的缺陷,其危险因素在高血压前期人群与正常人群中的重要程度不同。有研究指出,在高血压前期妇女中,蒙古族、高血压家族史是其发生高血压的独立预测指标,而在正常血压妇女中则不是^[21]。高血压前期人群与正常人群相比更容易发生 2 个以上危险因素聚集的情况,随着危险因素聚集的个数增多,发生高血压前期的危险性越大^[22]。

2 中医证候要素是开展风险预测研究的重要影响因素

中医证候是影响高血压前期人群患高血压病的重要因素。首先,高血压病属于中医学“眩晕”“头痛”范畴,病因有内外两个方面,病机为脏腑阴阳平衡失调,主要为肝肾阴阳平衡失调,病理因素主要为风、火、痰、瘀、虚,病位以肝肾为主,涉及脾、心^[23],因此,高血压前期人群中辨证属于阴虚阳亢或肝火上炎者,更有可能发展为高血压病。其次,有学者采用临床流行病学方法,对比高血压病和高血压前期证候学特征,发现血压由正常高值发展到高血压病的过程是一个机体由实到虚、由肝系病变到肝肾同病的逐渐发展过程,肝火亢盛证、阴阳两虚证可能为高血压前期发展为高血压病

的危险证候,肝火亢盛证可能是高血压前期发展为高血压病的高危证型^[24-26]。因此,我们提出构建高血压前期发展为高血压病的风险预测模型,不仅要考虑高血压前期高危人群危险因素,而且要重视中医证候要素,要将证候要素作为重要的影响因素,建立病证结合的高血压前期发展为高血压的风险预测模型,从而提高模型的预测准确性。

3 队列研究是构建风险预测模型的首选临床研究类型

队列研究是将某一特定人群按是否暴露于某可疑因素或暴露程度进行分组,追踪观察两组或多组成员结局发生的情况,比较各组之间结局发生率的差异,从而判定这些因素与该结局之间有无因果关联及关联程度的一种观察性研究方法^[27],因此,可作为构建风险预测模型的首选临床研究类型。通过文献研究及专家访谈,采集高血压前期发展为高血压病的可能危险因素和四诊信息,在中医理论指导下,围绕高血压前期发展为高血压病的危险因素和中医证候等相关问题,遵循筛选问卷设计的一般原则和方法,通过预调查、信度及效度分析等,形成高血压前期发展为高血压病的危险因素及四诊信息调查问卷。采用队列研究的方法,对自愿接受研究的高血压前期患者,利用高血压前期发展为高血压病的危险因素及四诊信息调查问卷进行调查,追踪随访,随访过程中发生高血压病者全部组成病例组,未发生高血压的高血压前期患者为对照组。对病例组和对照组的数据,采用描述性分析如 t 检验、秩和检验、方差分析、Pearson 相关系数法和 Spearman 相关系数法、Logistic 回归的逐步选择法,探索高血压前期发展为高血压病与各危险因素之间的定量关系。

4 潜在类别模型是客观判断受试者证候类型的可行方法

潜在类别模型是通过间断的潜变量即潜在类别变量来解释外显指标间的关联,使外显指标间的关联通过潜在类别变量来估计,进而维持其局部独立性的统计方法。其基本假设是,外显变量各种反应的概率分布可以由少数互斥的潜在类别变量来解释,每种类别对各外显变量的反应选择都有特定的倾向,因此,本文提出将收集的四诊信息作为显变量,进行潜在类别模型分析。潜在类别模型包括两类参数,一是计算条件概率值。条件概率值反映个体某观察变量属于特定类别的可能性,即通过计算

条件概率值的大小,分析各四诊信息(显变量)属于某一证候要素(潜变量)的概率,根据中医诊断有关内容,判断潜变量各类别所代表的证候要素的内容和性质。二是计算类别概率^[28]。类别概率表示随机抽取的个体属于一个特定类的概率。通过计算潜在类别概率,可以判断每个证素所占的比重和受试者属于各证候要素(潜类别)的概率,然后根据概率的大小将每个受试者进行分类,确定每一位受试者的证候类型,实现将证候纳入到风险预测模型中^[29]。

5 比较不同建模方法结果,挑选合适的风险预测模型

构建疾病的风险预测模型的方法较多,比较常用的有 Logistic 回归^[17]、Cox 回归^[13-14]、支持向量机(SVM)^[6d]、神经网络^[18]、马尔科夫模型^[6d]、决策树^[6i]、分类树^[6d]、贝叶斯网络模型^[6i]、时间序列模型^[18]等。目前普遍认为,不存在某种方法能适合于任何特点的数据。由于数据内部变量间的关系错综复杂,建立的模型之间存在不同的差异是肯定的^[32]。如 Delen 等^[33]比较了运用 Logistic 回归、决策树和神经网络对 20 多万乳腺癌患者生存状况的预测能力,结果显示,预测能力决策树(93.6%) > 神经网络(91.2%) > Logistic 回归(89.2%)。陈建新等^[6i]采用贝叶斯方法、决策树、RBF 神经网络、SVM、Logistic 回归对冠心病患者(共 1069 例)中医诊断分类的准确性进行比较研究,综合各项性能指标,SVM 最好,Logistic 回归次之,决策树、贝叶斯和神经网络依次次之。可见数据类型不同,不同模型的预测能力也不同。然而只要明确了试验设计类型、研究目的、各种建模方法特点等,将不同建模方法的结果进行合理地比较,挑选出较优的模型是完全可能的^[32]。

6 小结

识别发展为高血压病的高危高血压前期人群,预测未来发展为高血压病的风险,具有重要的临床价值。本研究通过分析高血压前期人群早期识别和风险预测的研究现状及存在问题,提出采用队列研究,收集高血压前期发展为高血压病的危险因素及四诊信息,充分考虑中医证候对疾病发生的影响,通过相关的数据分析和数学模型,筛选出对高血压前期发展为高血压病有重要影响的危险因素和中医证候,通过对模型的预测效能进行分析评价,挑选

出较优的包括危险因素和中医证候要素相结合的高血压前期发展为高血压病的风险预测模型,从而为高血压前期发展为高血压病的高危人群筛查提供科学依据。

参考文献

- [1] LLOYD-JONES DM, HONG Y, LABARTHE D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond [J]. Circulation, 2010, 121(4): 586-613.
- [2] GUO F, HE D, ZHANG W, et al. Trends in prevalence, awareness, management, and control of hypertension among United States adults, 1999 to 2010 [J]. J Am Coll Cardiol, 2012, 60(7): 599-606.
- [3] WANG J, XIONG X. Current situation and perspectives of clinical study in integrative medicine in China [J]. Evid Based Complement Alternat Med, 2012: 268542. doi: 10.1155/2012/268542.
- [4] WANG J, XIONG X. Control strategy on hypertension in Chinese medicine [J]. Evid Based Complement Alternat Med, 2012: 284847. doi: 10.1155/2012/284847.
- [5] CHOBANIHAN AV, BAKRIS GL, BLACK HR, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure [J]. JAMA, 2003, 289(19): 2560-2572.
- [6] 张汉明, 冯建章. 高血压前期-流行病学近况 [J]. 心血管病防治知识(学术版), 2012(3): 31-36.
- [7] 孙佳艺, 赵冬, 王薇, 等. 北京地区 2740 人的血压水平 10 年(1990—2002)变化情况 [J]. 高血压杂志, 2005, 13(2): 115-119.
- [8] ISHIKAWA Y, ISHIKAWA J, ISHIKAWA S, et al. Progression from prehypertension to hypertension and risk of cardiovascular disease [J]. J Epidemiol, 2017, 27(1): 8-13.
- [9] 中国高血压防治指南修订委员会. 中国高血压防治指南 2010 [J]. 中华高血压杂志, 2011, 19(8): 701-743.
- [10] PELLICCIA F, PASCERI V, MARAZZI G, et al. Randomised, double-blind, placebo-controlled, assessment of the efficacy and safety of dietary supplements in prehypertension [J]. J Hum Hypertens, 2017, 31(10): 647-653.
- [11] 汪海燕, 李玲, 范娥, 等. 穴位埋线对高血压前期的临床干预疗效观察 [J]. 上海针灸杂志, 2016, 35(10): 1200-1202.
- [12] 汤峰丽, 张晓一, 张晓天. 自拟平肝益肾健脾方联合生活方式干预对高血压前期人群的短期疗效观察 [J]. 中西医结合心脑血管病杂志, 2014, 12(10): 1180-1182.

- [3] BOZORGMANESH M, GHOREISHIAN H, MOHEBI R, et al. Sex-specific predictors of the prehypertension-to-hypertension progression: community-based cohort of a West-Asian population [J]. Eur J Prev Cardiol, 2013, 21(8): 956-963.
- [4] SELASSIE A, WAGNER CS, LAKEN ML, et al. Progression is accelerated from prehypertension to hypertension in blacks [J]. Hypertension, 2011, 58(4): 579-587.
- [5] 吴寿岭, 王娜, 赵海燕. 高敏 C 反应蛋白对高血压前期人群进展至高血压的预测价值 [J]. 中华高血压杂志, 2010, 18(4): 390-394.
- [6] 刘建平, 程锦泉, 张仁利, 等. 应用分类树模型构建缺血性脑卒中发病风险的预测模型 [J]. 中国慢性病预防与控制, 2012, 20(3): 254-258.
- [7] 文玉娜, 陈立章, 薛静, 等. 乙肝后肝硬化患者死亡概率预测模型建立与评价 [J]. 中国公共卫生, 2015, 31(2): 211-214.
- [8] 杨召, 叶中辉, 赵磊, 等. ARIMA-BPNN 组合预测模型在流感发病率预测中的应用 [J]. 中国卫生统计, 2014, 31(1): 16-18.
- [9] DELEN D, WALKER G, KADAM A. Predicting breast cancer survivability: a comparison of three data mining methods [J]. Artif Intell Med, 2005, 34(2): 113-127.
- [10] 于涛, 刘欢乐, 冯新, 等. 基于健康管理队列的高血压风险预测模型 [J]. 山东大学学报(医学版), 2017, 55(6): 61-65.
- [11] SUN Z, ZHENG L, DETRANO R, et al. Risk of progression to hypertension in a rural Chinese women population with prehypertension and normal blood pressure [J]. Am J Hypertens, 2010, 23(6): 627-632.
- [12] 刘冰. 中国农村和城市 35~70 岁人群高血压前期检出率及影响因素分析 [D]. 北京: 中国协和医科大学, 2010: 63.
- [13] 满秋珊, 王爽, 袁敬柏. 高血压前期的中医药预防 [J]. 中国中医药信息杂志, 2010, 17(6): 5-6.
- [14] 李文娟. 高血压与正常高值血压基本证候调查比较研究 [D]. 济南: 山东中医药大学, 2012: 6-18.
- [15] 刘松虎. 基于中医示范社区的高血压人群防治研究(山东平原模式): 社区正常高值血压与高血压人群比较研究 [D]. 济南: 山东中医药大学, 2012: 24-28.
- [16] 戴霞. 高血压与高血压前期中医证候分布规律研究进展 [J]. 中国中医药现代远程教育, 2014, 12(6): 160-161.
- [17] 沈洪兵, 齐秀英. 流行病学 [M]. 8 版. 北京: 人民卫生出版社, 2013: 54.
- [18] 邱皓政. 潜在类别模型的原理与技术 [M]. 北京: 教育科学出版社, 2008: 7-10.
- [19] 许碧云, 陈炳为, 陈后光, 等. 伴有协变量的潜在类别模型在出生缺陷预防知识分析中的应用 [J]. 现代预防医学, 2016, 43(6): 961-964.
- [20] 康殿民, 王洁贞, 薛付忠, 等. Markov 模型在山东省肾综合征出血热流行趋势预测中的应用 [J]. 现代预防医学, 2006, 33(9): 1665-1666.
- [21] 陈建新, 西广成, 王伟. 数据挖掘分类算法在冠心病临床应用的比较 [J]. 北京生物医学工程, 2008, 27(3): 249-252.
- [22] 史忠植. 知识发现 [M]. 北京: 清华大学出版社, 2002: 2-6.

The Idea about Establishing Syndrome-Combination Risk Forecasting Model for Hypertension in Prehypertension Population

WANG Liying, HAN Xuejie

(Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medicine Science, Beijing, 100700)

ABSTRACT Early identification of prehypertension populations that may develop into hypertension and targeted intervention may reduce the prevalence of hypertension, which is the strategic demand of shifting gravity forward of disease prevention and treatment. At present, there are some limitations in prediction study of hypertension in prehypertension population. The authors put forward that by collecting the risk factors and four diagnosis information in prehypertension population, as well as by data analysis, the high risk factors and Chinese medicine syndrome elements for the development of hypertension from prehypertension may be screened out. Model could be established with multiple methods, such as nerve net, support vector machine and decision tree. By validating and evaluating the model, the risk forecasting model combined with risk factors and Chinese medicine syndrome might be established to provide scientific basis for early screening prehypertension population who may develop into hypertension.

Keywords prehypertension; hypertension; risk forecasting model; disease and syndrome combination

(收稿日期: 2017-05-08; 修回日期: 2018-01-06)

编辑: 王利芬