

本文引文格式:刘炎,张玲,李琦,等.基于机器学习的饮食质量指数及其对血压水平及高血压的预测研究[J].右江民族医学院学报,2024,46(3):295-300,306.

【论著与临床报道】

基于机器学习的饮食质量指数及其对血压水平及高血压的预测研究

刘炎,张玲,李琦,杨灿,杜亚楠,袁敏

(安徽医科大学卫生管理学院卫生健康数据科学系,安徽 合肥 230032)

摘要:目的 基于机器学习弹性网变量选择方法提出一种数据自适应的饮食质量指数,并探讨其与血压及高血压患病风险的关联。方法 数据来源于美国营养调查协会(NHANES)2011~2018年4个周期的1 143个亚裔样本数据,其中男性占59.06%,平均年龄为(43.12±15.10)岁;其它变量包括人口统计学变量、29种食物摄入成分以及舒张压和收缩压测量值。在控制相关混杂因素的前提下,利用弹性网变量选择方法筛选出重要食物摄入成分,并以食物成分对应的回归系数为权重计算加权平均值作为数据自适应饮食质量指数(ddDQS)。进一步分析新指数与高血压风险的关联,并与健康饮食指数(HEI-2015)、替代健康饮食指数(AHEI)、得舒饮食指数(DASH)以及地中海饮食指数(MED)4种常用的饮食指数进行比较。结果 高血压患病率为26.15%,患高血压个体的平均年龄[(52.94±14.77)岁]显著高于没有患高血压的个体[(39.64±13.62)岁]。ddDQS包含5个重要的膳食成分(精制谷物、油、酒精、添加糖以及马铃薯),与收缩压(SBP)($\beta = -2.08, 95\% CI = -3.24 \sim -0.92, P < 0.001$)及高血压(HTN)($\beta = -0.482, 95\% CI = -0.72 \sim -0.25, OR = 0.62, P < 0.001$)均相关。摄入马铃薯会增加ddDQS指数,而摄入精制谷物、油、酒精、添加糖则会降低ddDQS指数。亚组分析结果表明40~60岁以及男性人群对于ddDQS更加敏感,这些亚群将更多地从健康膳食模式中获益。HEI-2015与DBP、SBP以及HTN存在一定关联,但P值均大于ddDQS对应P值。本研究未发现DASH、AHEI和MED与DBP、SBP以及HTN的关联。结论 与常用饮食指数相比,ddDQS对高血压有更好的预测。ddDQS越高越有益于血压及高血压风险的降低。针对亚裔人群的饮食指数的构建可指导未来研究进行高血压预防与控制。

关键词:高血压;饮食质量指数;弹性网变量选择;机器学习

中图分类号:R544.1

文献标识码:A

文章编号:1001-5817(2024)03-0295-07

doi:10.3969/j.issn.1001-5817.2024.03.003

Construction of machine learning-based diet quality index and its association with blood pressure levels and hypertension risk

LIU Yan, ZHANG Ling, LI Qi, YANG Can, DU Yanan, YUAN Min

(Department of Health Data Science, School of Health Management, Anhui Medical University, Hefei 230032, Anhui, China)

Abstract: **Objective** This paper aimed to propose a data-adaptive diet quality index based on the elastic net variable selection method in machine learning. Additionally, it explored the association of this index with blood pressure levels and the risk of hypertension. **Methods** The data source was from 1 143 Asian samples in the National Health and Nutrition Examination Survey (NHANES) across four cycles from 2011 to 2018, with 59.06% being male and an average age of (43.12±15.10) years. Other variables included demographic variables, intake levels of 29 food components and diastolic and systolic blood pressure measurements. Controlling for related confounding factors, the elastic net variable selection method was used to screen important food components. The weighted average of food components was defined as the data-driven diet quality score

基金项目:国家自然科学基金项目(82073578);安徽省自然科学面上项目(2008085MA09);人口健康与优生安徽省重点实验室课题(JKYS20233)

第一作者:刘炎,在读硕士研究生,研究方向:医学统计,E-mail:yliu_ahmu@163.com

通讯作者:袁敏,博士,副教授,硕士研究生导师,研究方向:遗传统计、医学统计、生物信息等,E-mail:myuan@ahmu.edu.cn

(ddDQS), using their corresponding regression coefficients as weights. Furthermore, the association between the new index and the risk of hypertension was investigated and compared with four commonly used diet quality indices: the Healthy Eating Index 2015 (HEI-2015), the Alternative Healthy Eating Index (AHEI), the Dietary Approaches to Stop Hypertension (DASH) index, and the Mediterranean Dietary Index (MED). **Results** The prevalence of hypertension was 26.15%, and the average age of individuals with hypertension (52.94 ± 14.77) years was significantly higher than those without hypertension (39.64 ± 13.62) years. The ddDQS included five important dietary components; refined grains, oils, alcohol, added sugars and starchy potatoes. It was significantly associated with systolic blood pressure (SBP) ($\beta = -2.08, 95\% CI = -3.24 \sim -0.92, P < 0.001$) and hypertension (HTN) ($\beta = -0.482, 95\% CI = -0.72 \sim -0.25, OR = 0.62, P < 0.001$). Intake of starchy potatoes increased the ddDQS, while intake of refined grains, oils, alcohol, and added sugars decreased the ddDQS. Subgroup analysis results showed that individuals aged 40~60 years and males were more sensitive to the ddDQS, suggesting that these subgroups would benefit more from adhering to healthy dietary patterns. The HEI-2015 was associated with diastolic blood pressure (DBP), SBP, and HTN to some extent, but the *P*-values were greater than those corresponding to the ddDQS. This study did not find associations between the DASH, AHEI, and MED indices and DBP, SBP, or HTN. **Conclusion** Compared to commonly used diet quality indices, the ddDQS constructed using machine learning methods demonstrates better predictive ability for hypertension. A higher ddDQS is beneficial for reducing blood pressure levels and the risk of hypertension. The development of a diet quality index tailored to the Asian population can guide future research on hypertension prevention and control.

Key words: hypertension; diet quality index; elastic net variable selection; machine learning

高血压(hypertension)是心血管疾病和肾脏疾病等非传染慢性疾病的重要危险因素^[1-4]。由伦敦帝国理工学院和 WHO 牵头的迄今为止全球最大规模的高血压趋势研究发现,近 30 年来,30~79 岁的成年人患有高血压的人数从 6.5 亿增长至 12.8 亿^[5]。高血压是一种复杂的多因素疾病,是多种遗传和环境因素相互作用的结果。广泛研究表明高血压可改变的重要环境风险因素是膳食习惯,膳食方式的改善可降低大约 44%左右的高血压患病风险^[6]。饮食质量指数是综合衡量个体或人群膳食是否健康的重要度量。常用的饮食指数包括饮食健康指数(HEI-2015)^[7]、得舒饮食(DASH)^[8]、替代饮食指数(AHEI)^[9]和地中海饮食(MED)^[10]。研究表明,上述饮食指数越高则饮食越健康,并认为坚持健康的饮食模式有利于降低非传染慢性病的风险^[11-12]。

文献中已有大量基于一般人群或者特定人群的饮食质量指数与高血压风险的研究^[13],然而针对亚裔人群的相关研究相对较少。另外,多数饮食指数都基于营养学家的专业知识,并根据临床报道与高血压显著关联的饮食因素加以更新。考虑到饮食成分之间可能存在的交互作用,基于机器学习方法构造的饮食指数将包含独立于营养学家经验的信息,具有客观性。本文将进一步研究新的饮食指数与高血压患病风险的关联,通过与目前常用饮食指数作比较,评价新指数的优良性。良好的膳食习惯对高血压的三级预防会产生积极影响,从而降低高血压患者的疾病负担^[13-14]。

1 材料与方法

1.1 数据来源 本研究数据来源于 2011~2018 年美国健康与营养调查协会(National Health and Nutrition Examination Survey,简称 NHANES)4 个连续周期。由于 NHANES 在 2011 年开始收集是否为亚裔的种族变量,为尽可能保留样本量,仅将年龄、性别、教育程度、BMI、总胆固醇和甘油三酯作为协变量纳入模型。年龄、性别、种族和受教育程度通过问卷调查收集得到。通过测量身高体重并将体重(kg)除以身高(m)的平方则作为身体质量指数(BMI)。总胆固醇和甘油三酯通过实验室检验获得。根据个体受教育年限将教育程度分为低于高中水平、高中毕业或同等学历、大学及以上学历。NHANES 是一个面向公众的数据库,所有参与者都签署了知情同意书,符合伦理学标准。

1.2 研究变量 ①血压:由专业人员按照标准化程序连续测量 4 次,本研究舍弃第一次测量数据,取后 3 次平均数作为样本的血压值。若只有 1 个测量值可用,则使用唯一测量值;如果有 2 个或 3 个测量值可用,则使用平均值。本研究将收缩压(SBP) ≥ 18.62 kPa 或者舒张压(DBP) ≥ 11.97 kPa 定义为高血压^[15]。②饮食数据及饮食指数:受试者在检查中心完成体检与抽血化验,并进行 24 h 膳食回忆。在 3~10 d 后,通过电话访谈完成第二次 24 h 膳食回忆^[16]。从官方网站下载 FPED 数据集和 NHANES 膳食数据集,并根据 HEI-2015、DASH、AHEI 以及 MED 饮食指数的计分方法计算相应的饮食得分。这些饮食指数包含不同的

食物成分,侧重于饮食质量的不同方面,详细计分方式可参考相关文献^[7-10]。为计算新的饮食指数,本研究通过弹性网变量选择方法从 29 个饮食成分中确定重要饮食成分变量与最终模型。

1.3 统计方法 首先基于弹性网变量选择方法,将 L_1 和 L_2 惩罚同时应用于似然函数,将不重要变量的效应直接估计为 0^[17]。具体而言,记 $X_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4})$ 为年龄、性别、教育程度以及总能量摄入, $Z_i = (Z_{i1}, Z_{i2}, \dots, X_{i29})$ 为食物成分, Y_i 表示血压值或高血压患病状态, $i = 1, 2, \dots, 1143$ 。本文仅对食物成分的变量进行变量选择,通过优化如下目标函数得到相应的参数估计, $\hat{\beta} = \operatorname{argmin} \sum_{i=1}^{1143} l(X_i, Z_i, Y_i | \theta) + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2$, 其中 $l(X_i, Z_i, Y_i | \theta)$ 为第 i 个样本的对数似然函数,参数 $\theta = (\gamma, \beta)'$, γ 为协变量对应的回归系数向量, β 为食物成分对应的回归系数向量, $\|\cdot\|_k, k = 1, 2$ 分别表示 L_1 和 L_2 范数。定义第 i 个样本的数据自适应饮食质量指数为 $ddDQS_i = \sum_{k=1}^{29} Z_{ik} \hat{\beta}_k, i = 1, 2, \dots, 1143$ 。本文使用 R 软件包 `gcdnet` 实现弹性网变量选择和参数估计,将非零系数对应的食物成分确定为重要变量。

确定最终饮食变量和模型后,模型中每个食物成分解释的方差占总方差的比例将被用于衡量新的饮食指数中食物成分的相对重要性。本研究进一步对新指

数和血压及高血压结局进行关联分析,并与 HEI-2015、DASH、AHEI 以及 MED 的关联分析结果进行比较。在亚组分析中,对年龄、性别进行分层,讨论这些重要协变量与饮食质量指数的交互作用。本研究考虑血压值作为结局变量时采用线性回归模型,考虑是否患高血压作为结局变量时采用 Logistic 回归模型。

2 结果

2.1 一般情况 本研究将个体的总能量摄入作为协变量加以控制,在质控和删除缺失数据后,共获得 1 143 个包含完整人口学基本信息、膳食数据以及血压数据的亚裔个体,其中男性 675 例(59.06%),女性 468 例(40.94%)。研究人群的平均年龄为(43.12 ± 15.10)岁,其中患高血压个体的平均年龄[(52.94 ± 14.77)岁]高于没有患高血压的个体[(39.64 ± 13.62)岁],差异有统计学意义($P < 0.001$)。具有大学及以上学历的个体占大多数(62.55%),高中以下学历的人占比较少(6.56%)。5 种饮食指数在没有高血压的人群中较高,表明饮食健康是高血压的一个潜在影响因素。表 1 展示了研究人群总体和按是否高血压交叉分类的人口学特征以及饮食质量指数的分布。患高血压和不患高血压人群在年龄、性别、受教育程度、BMI、总胆固醇、甘油三脂这些变量上比较差异具有统计学意义($P < 0.05$)。

表 1 重要人口学特征和饮食质量指数在高血压及非高血压人群中分布

项目	全人群 ($n = 1\ 143$)	高血压		t/χ^2	P
		否($n = 844$)	是($n = 299$)		
性别				5.688	0.017
男	675 (59.06)	481 (56.99)	194 (64.88)		
女	468 (40.94)	363 (43.01)	105 (35.12)		
年龄/岁	43.12 ± 15.10	39.64 ± 13.62	52.94 ± 14.77	-13.661	<0.001
总能量/KJ	2038.62 ± 817.99	2043.81 ± 818.08	2023.97 ± 818.95	0.360	0.709
受教育程度				7.136	0.028
高中学历以下	75 (6.56)	47 (5.57)	28 (9.36)		
高中文凭	353 (30.88)	254 (30.09)	99 (33.11)		
大学及以上	715 (62.55)	543 (64.34)	172 (57.53)		
BMI/($\text{kg} \cdot \text{m}^{-2}$)	25.10 ± 4.63	24.61 ± 4.39	26.46 ± 5.00	-5.670	<0.001
总胆固醇/($\text{nmol} \cdot \text{L}^{-1}$)	193.16 ± 39.68	189.44 ± 37.64	203.67 ± 43.31	-5.047	<0.001
甘油三脂/($\text{nmol} \cdot \text{L}^{-1}$)	163.40 ± 124.70	152.3 ± 120.92	194.48 ± 130.08	-4.900	<0.001
饮食质量指数					
健康饮食指数(HEI-2015)	56.35 ± 14.28	56.86 ± 14.47	56.16 ± 14.22	0.717	0.473
替代健康膳食数(AHEI)	51.05 ± 14.04	51.55 ± 14.43	50.87 ± 13.90	0.713	0.476
得舒膳食指数(DASH)	27.23 ± 3.63	27.55 ± 3.68	27.12 ± 3.61	1.731	0.084
地中海饮食指数(MED)	6.29 ± 1.21	6.44 ± 1.16	6.23 ± 1.22	2.611	0.009
数据自适应饮食指数(ddDQS)	51.67 ± 11.93	52.06 ± 12.06	50.56 ± 11.50	1.910	0.057

注:表内计数资料数据用[$n(\%)$]表示;计量资料数据以($\bar{x} \pm s$)表示,连续变量检验为 t 检验,属性变量检验用 χ^2 检验。

2.2 新的膳食质量指数构建 本研究所用的 29 种食物成分信息,见表 2。由表 2 可知,多数单个膳食成分在高血压患者和非高血压人群比较差异无统计学意义 ($P > 0.05$)。因此,有必要提出一个综合指数以评估总体膳食摄入的累积效应。通过弹性网变量选择, ddDQS 共筛选出 5 个重要的膳食成分(油、精制谷物、酒精、添加糖、马铃薯)。将每个成分方差的解释比作

为衡量食物成分的相对重要性,并且根据弹性网变量选择方法最终模型中回归系数的正负号定义贡献的不同方向,结果在图 1 中进行展示。马铃薯具有正的效应,表明适当摄入这 2 个食物会增加 ddDQS 指数,而摄入精制谷物、油、酒精、添加糖则会降低 ddDQS 指数。ddDQS 指数越高代表饮食越健康,从而遵循高 ddDQS 指数饮食模式有利于降低高血压风险。

表 2 用于构建 AHEI 等 5 种饮食指数的饮食成分基本信息表

成分	总人群 ($n = 1\ 143$)	高血压		W	P
		否($n = 844$)	是($n = 299$)		
蔬菜(7)					
深绿色蔬菜	0.00(0.00,0.36)	0.00(0.00,0.37)	0.00(0.00,0.35)	130994.000	0.285
西红柿	0.07(0.00,0.34)	0.09(0.00,0.35)	0.03(0.00,0.31)	135083.000	0.061
其它富含胡萝卜素的蔬菜	0.00(0.00,0.17)	0.01(0.00,0.18)	0.00(0.00,0.15)	131652.500	0.232
马铃薯	0.00(0.00,0.35)	0.00(0.00,0.36)	0.00(0.00,0.29)	134356.000	0.066
其它淀粉类	0.00(0.00,0.04)	0.00(0.00,0.03)	0.00(0.00,0.05)	124838.500	0.728
其它蔬菜	0.61(0.21,1.17)	0.61(0.21,1.14)	0.65(0.23,1.32)	119505.000	0.173
除豆类外其它蔬菜	0.00(0.00,0.04)	0.00(0.00,0.04)	0.00(0.00,0.04)	126941.500	0.845
水果(3)					
柑橘、瓜类和浆果类	0.00(0.00,0.31)	0.00(0.00,0.26)	0.00(0.00,0.40)	121253.500	0.239
除上述 3 类外其它整果	0.13(0.00,0.91)	0.09(0.00,0.90)	0.34(0.00,1.04)	116496.500	0.040
果汁	0.00(0.00,0.05)	0.00(0.00,0.04)	0.00(0.00,0.07)	118440.000	0.056
谷物(2)					
全谷物	0.28(0.00,1.72)	0.29(0.00,1.72)	0.28(0.00,1.68)	127197.000	0.827
精制谷物	5.29(3.19,7.67)	5.09(3.16,7.49)	5.84(3.33,8.07)	117313.500	0.071
蛋白质食物(10)					
红肉	0.36(0.00,2.83)	0.30(0.00,2.86)	0.40(0.00,2.80)	125791.500	0.934
腌制肉	0.00(0.00,0.28)	0.00(0.00,0.28)	0.00(0.00,0.18)	127736.500	0.690
内脏	0.00(0.00,0.00)	0.00(0.00,0.00)	0.00(0.00,0.00)	125564.000	0.526
禽类	0.00(0.00,2.36)	0.00(0.00,2.60)	0.00(0.00,1.93)	137072.500	0.015
富含 n-3 脂肪酸的海鲜	0.00(0.00,0.00)	0.00(0.00,0.00)	0.00(0.00,0.00)	126167.500	0.997
缺乏 n-3 脂肪酸的海鲜	0.00(0.00,0.52)	0.00(0.00,0.51)	0.00(0.00,0.62)	126183.000	0.999
鸡蛋	0.13(0.00,0.93)	0.13(0.00,0.93)	0.13(0.00,0.90)	128790.500	0.589
豆制品	0.00(0.00,0.00)	0.00(0.00,0.00)	0.00(0.00,0.00)	126111.500	0.983
花生、坚果和种子	0.00(0.00,0.90)	0.00(0.00,0.75)	0.00(0.00,1.03)	120866.500	0.229
豆类	0.00(0.00,0.18)	0.00(0.00,0.18)	0.00(0.00,0.18)	126929.500	0.847
乳制品(3)					
牛奶	0.18(0.00,0.85)	0.22(0.00,0.87)	0.13(0.00,0.70)	132919.500	0.166
酸奶	0.00(0.00,0.00)	0.00(0.00,0.00)	0.00(0.00,0.00)	133070.000	0.048
奶酪	0.00(0.00,0.56)	0.00(0.00,0.60)	0.00(0.00,0.47)	134549.000	0.055
脂肪(2)					
油	23.76(13.11,39.57)	23.58(13.11,39.36)	24.40(13.10,40.46)	124569.000	0.743
固体脂肪	20.24(10.56,34.71)	21.05(11.14,35.68)	18.34(9.28,32.23)	136001.500	0.045
糖(1)					
添加糖	8.20(3.57,15.18)	7.79(3.00,14.32)	8.30(4.03,15.41)	120513.000	0.248
酒精(1)					
酒精	0.00(0.00,0.34)	0.00(0.00,0.34)	0.00(0.00,0.57)	125129.000	0.783

注:表中数据以 $M(P_{25}, P_{75})$ 表示。

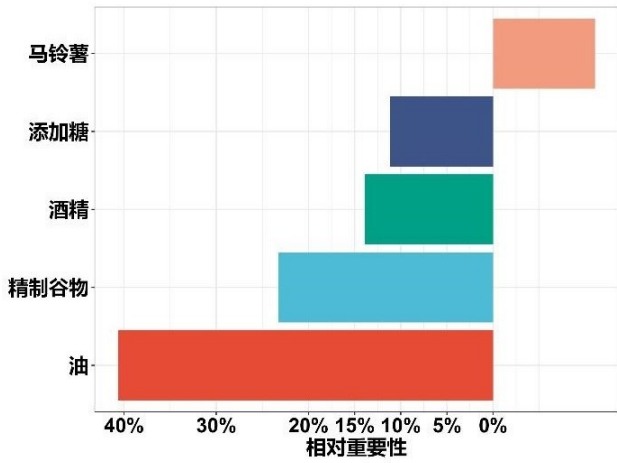


图 1 弹性网变量筛选出的重要食物成分以及它们的相对重要性

2.3 膳食质量指数与血压及高血压的关联 表 3 为 5 种饮食质量指数与 DBP、SBP 和高血压结局 (HTN) 分别进行关联分析的结果。结果显示, ddDQS 与 SBP 及 HTN 均显著相关 ($P < 0.001$)。SBP 和 DBP 对应的回归系数 < 0 以及 HTN 对应的回归系数 < 0 , 即对应 Z 值 < 1 表明 ddDQS 越高越有益于血压及高血压风险的降低。本研究中其它 3 种饮食质量指数均未获得与血压及 HTN 的显著关联。

表 3 5 种饮食质量指数与血压及 HTN 的关联分析比较

饮食指数	<i>b</i>	SE	95% CI	Z	P
DBP					
AHEI	-0.111	0.348	-0.793~0.572	-0.317	0.751
DASH	-0.030	0.348	-0.713~0.652	-0.087	0.931
HEI-2015	-0.572	0.339	-1.237~0.092	-1.690	0.092
MED	-0.081	0.341	-0.749~0.586	-0.239	0.811
ddDQS	-0.622	0.490	-1.582~0.339	-1.270	0.205
SBP					
AHEI	-0.281	0.423	-1.110~0.548	-0.665	0.506
DASH	-0.426	0.423	-1.255~0.403	-1.010	0.314
HEI-2015	-0.684	0.412	-1.491~0.123	-1.660	0.097
MED	-0.489	0.413	-1.299~0.322	-1.180	0.237
ddDQS	-2.080	0.592	-3.238~-0.915	-3.510	< 0.001
HTN					
AHEI	-0.116	0.080	-0.273~0.040	-1.460	0.145
DASH	-0.045	0.080	-0.202~0.111	-0.568	0.570
HEI-2015	-0.120	0.078	-0.273~0.031	-1.550	0.121
MED	0.013	0.079	-0.143~0.169	0.158	0.875
ddDQS	-0.482	0.119	-0.718~-0.253	-4.070	< 0.001

注: HTN 为二值分类变量, $SBP \geq 18.62$ kPa 或 $DBP \geq 11.97$ kPa 时 HTN=1, 否则 HTN=0。

2.4 亚组分析 本研究将人群按照年龄和性别进行分层分析, 结果见图 2。结果表明 40~60 岁的人群比 40 岁以下以及 60 岁以上的人群从遵循健康饮食带来的高血压风险降低收益更大。特别地, 在 40~60 岁人群中 ddDQS 对应的高血压的 OR 为 0.43 (95% CI: 0.28~0.63, $P < 0.001$) 具有显著的统计学意义。另外, 男性似乎比女性从健康饮食模式中获得的高血压风险降低收益更大, ddDQS 的高血压风险 OR 分别为 0.58 和 0.72, P 值分别为 < 0.001 和 0.15。其它的 4 种膳食指标与 ddDQS 具有相同的趋势, 但 ddDQS 相应的 OR 和 P 值更小, 表明基于数据信息挖掘的 ddDQS 与高血压风险的关联更大。

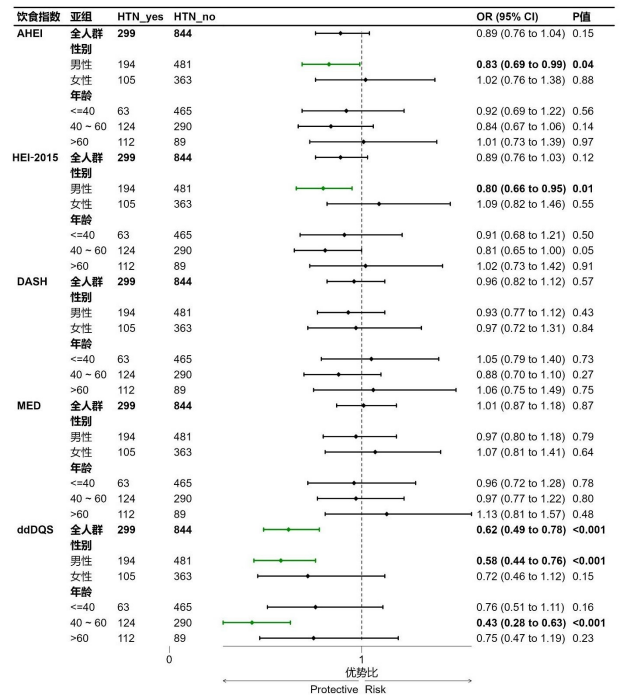


图 2 在不同的年龄和性别亚组内饮食指数与高血压结局的关联分析森林图

2.5 HTN 与饮食质量指数的关联分析 本研究根据 WHO 的标准, 将高血压定义为 $SBP \geq 18.62$ kPa 或 $DBP \geq 11.97$ kPa。在敏感性分析中本研究采用美国心脏病学会和美国心脏协会 2017 年的指南将高血压定义为 $SBP \geq 17.29$ kPa 或 $DBP \geq 10.64$ kPa, 仍然得到一致的关联性结论, 说明基于机器学习的饮食质量指数是稳健的, 结果见表 4。

表 4 HTN 与饮食质量指数的关联分析结果

高血压	饮食质量指数	OR	b	SE	95% CI	Z	P
HTN	AHEI	0.885	-0.122	0.101	-0.319~0.076	-1.210	0.226
	DASH	0.939	-0.063	0.101	-0.263~0.135	-0.623	0.534
	HEI-2015	0.872	-0.137	0.098	-0.330~0.055	-1.400	0.162
	MED	0.920	-0.084	0.100	-0.280~0.113	-0.835	0.404
	ddDQS	0.664	-0.410	0.150	-0.709~-0.120	-2.730	0.006

注:HTN:SBP \geq 17.29 kPa 或 DBP \geq 10.64 kPa。

3 讨论

高血压是影响全球很大一部分人口的重要公共卫生问题。世界卫生组织估计,全世界每年有 940 万人因高血压而死亡,是心血管疾病的主要危险因素。高血压的发生和发展由包括遗传和环境在内的多种因素所决定,在可改变因素中,健康膳食在预防和控制高血压中发挥关键作用。研究表明,亚洲人群富含蔬菜、水果、豆类和全谷物的传统膳食有助于降低血压水平并降低患高血压的风险。然而,随着传统饮食模式向高脂、高盐和高糖以及低纤维的西化膳食的转变,亚洲人群的高血压发病率的逐渐上升^[18-19]。

目前常用的饮食质量指数主要包括世界卫生组织推荐的健康饮食指数(HEI-2015)、哈佛大学公共卫生学院提出的替代健康饮食指数(AHEI),得舒膳食(DASH)以及地中海膳食(MED)。HEI-2015 用于评估个体摄入食物与美国膳食指南的一致性程度。DASH 膳食是一种以植物为中心的膳食,强调全谷物、水果、蔬菜、低脂乳制品、瘦肉和坚果/种子/豆类。MED 是一种植物性膳食,强调全谷物、水果、蔬菜、豆类、坚果、鱼和橄榄油。AHEI 强调水果、蔬菜、全谷物、坚果/种子、豆类、鱼/家禽和健康脂肪。许多研究表明,遵循上述健康膳食模式在一定程度上与较低的高血压风险有关^[20-22]。虽然这些膳食评分有助于从不同方面评估饮食质量,但它们均基于营养学专家的专业知识和临床经验。基于机器学习算法分析数据并识别人眼可能看不到的模式和趋势,能有效提取膳食数据特征,获得更加客观和易操作的饮食质量指数。

本研究提出了一种基于弹性网变量选择方法的数据驱动的饮食指数,研究了其与舒张压、收缩压以及高血压之间的关联性,并且与常用的 4 种饮食质量指数做了比较。结果表明,基于机器学习方法的新指数 ddDQS 与血压和高血压结局之间存在显著的关联性,并且 ddDQS 越高,越有利于降低血压和降低高血压的患病风险。进一步的亚组分析结果表明 40~60 岁以及男性人群对于饮食质量指数更加敏感,这些亚群将更多地从健康的饮食模式中获益,从而降低高血压的患病风险。除 ddDQS 外,其它饮食质量指数均未发现

显著的关联,原因可能在于本研究能获取的亚裔人群样本量,后续需要在更大的亚裔人群队列中进行研究。

总之,目前基于 NHANES 亚裔人群的数据自适应的饮食质量指数并不多见,考虑到亚裔人群的特殊性以及对于中国人群饮食质量指数构造的参考,基于机器学习方法构建针对亚裔人群的饮食质量指数有一定的必要性和意义,新的指数有可能为中国和亚洲其他地区提供更精确和个性化高血压预防控制的建议。

尽管自 2011 年起 NHANES 的数据收集已经开始将亚裔人群纳入分类,但并未对这一群体进行更详细的细分。这种缺乏细致分类的做法限制了对亚裔各子群体在健康结果、行为或状态方面差异的深入分析。这是本研究的一个局限。

参考文献:

- [1] MILLS K T, STEFANESCU A, HE J. The global epidemiology of hypertension [J]. Nat Rev Nephrol, 2020, 16 (4): 223-237.
- [2] NCD RISK FACTOR COLLABORATION(NCD-RISC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants [J]. Lancet, 2017, 389(10064): 37-55.
- [3] SHARP S I, AARSLAND D, DAY S, et al. Hypertension is a potential risk factor for vascular dementia: systematic review [J]. Int J Geriatr Psychiatry, 2011, 26 (7): 661-669.
- [4] EMDIN C A, ANDERSON S G, SALI-MI-KHORSHIDI G, et al. Usual blood pressure, atrial fibrillation and vascular risk: evidence from 4.3 million adults [J]. Int J Epidemiol, 2017, 46(1): 162-172.
- [5] NCD RISK FACTOR COLLABORATION(NCD-RISC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants [J]. Lancet, 2021, 398(10304): 957-980.
- [6] SCHWINGSHACKL L, SCHWEDHELM C, HOFFMANN G, et al. Food groups and risk of hypertension: a systematic review and dose-response meta-analysis of prospective studies [J]. Adv Nutr, 2017, 8(6): 793-803.

(下转第 306 页)

- [17] MOBLEY L R, MAGNUSSEN J. An international comparison of hospital efficiency; does institutional environment matter? [J]. *Applied Economics*, 1998, 30(8): 1089-1100.
- [18] PARKIN D, HOLLINGSWORTH B. Measuring production efficiency of acute hospitals in Scotland, 1991-94; validity issues in data envelopment analysis [J]. *Applied Economics*, 1997, 29(11): 1425-1433.
- [19] EVANS D B, TANDON A, MORRAY C J, et al. Comparative efficiency of national health systems: cross national econometric analysis [J]. *BMJ*, 2001, 323(7308): 307-310.
- [20] CABALLER-TARAZONA M, MOYA-CLEMENTE I, VIVAS-CONSUELO D, et al. A model to measure the efficiency of hospital performance [J]. *Math Comput Model*, 2010, 52(7): 1095-1102.
- [21] 张杰, 唐小芬, 梁秋瑜, 等. 2012—2020 年广西乡镇卫生院卫生资源配置效率评价: 基于 DEA 和 SFA 方法 [J]. *卫生软科学*, 2022, 36(8): 50-55.
- [22] 农圣. 整合型卫生服务体系的内涵和现状 [J]. *卫生经济研究*, 2017(11): 4-8.
- [23] 王波, 杨林. 共享发展理念下医疗卫生资源有效供给: 基于城乡比较 [J]. *东岳论丛*, 2017, 38(9): 158-166.
- [24] 王震, 蔡雪妮, 王清华. 医保报销政策能否引导分级诊疗——基于江苏某市的调查 [J]. *中国医疗保险*, 2017(5): 10-13.
- [25] 张文东, 吴阿元. 镇江市医保支付机制在分级诊疗中的运用效果及思考 [J]. *中国医疗保险*, 2017(3): 45-49.
- [26] 张天天, 戴瑞明, 周奕男, 等. 可分流康复机构的住院病人筛选方法与实证研究 [J]. *中国医院管理*, 2017, 37(11): 3-4, 7.
- [27] 黄汉明, 刘跃华, 卜亚丽. 南京市医保差异化支付与分级诊疗的实证分析 [J]. *中国医疗保险*, 2015(6): 41-44, 48.

收稿日期: 2024-01-26; 修回日期: 2024-03-10

(上接第 300 页)

- [7] KREBS-SMITH S M, PANNUCCI T E, SUBAR A F, et al. Update of the Healthy Eating Index: HEI-2015 [J]. *J Acad Nutr Diet*, 2018, 118(9): 1591-1602.
- [8] MELLEN P B, GAO S K, VITOLINS M Z, et al. Deteriorating dietary habits among adults with hypertension: dash dietary concordance, NHANES 1988-1994 and 1999-2004 [J]. *Arch intern med*, 2008, 168(3): 308-314.
- [9] CHIUVE S E, FUNG T T, RIMM E B, et al. Alternative dietary indices both strongly predict risk of chronic disease [J]. *J Nutr*, 2012, 142(6): 1009-1018.
- [10] TRICHOPOULOU A, COSTACOU T, BAMIA C, et al. Adherence to a Mediterranean diet and survival in a Greek population [J]. *N Engl J Med*, 2003, 348(26): 2599-2608.
- [11] SOTOS-PRIETO M, BHUPATHIRAJU S N, MATTEI J, et al. Changes in diet quality scores and risk of cardiovascular disease among US men and women [J]. *Circulation*, 2015, 132(23): 2212-2219.
- [12] 杨柳青, 田红梅, 石汉平. 三种饮食模式与慢性疾病研究进展 [J]. *首都医科大学学报*, 2022, 43(2): 311-320.
- [13] SCHOENAKER DAJN, SOEDAMAH-MUTHU S S, MISHRA G D. The association between dietary factors and gestational hypertension and pre-eclampsia: a systematic review and meta-analysis of observational studies [J]. *BMC Med*, 2014, 12: 157.
- [14] 王琦. DASH 膳食(得舒饮食)对社区原发性高血压患者生活质量的影响 [J]. *长治学院学报*, 2022, 39(5): 10-14.
- [15] ZHOU B, PEREL P, MENSAH G A, et al. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension [J]. *Nat Rev Cardiol*, 2021, 18(11): 785-802.
- [16] Centers for Disease Control and Prevention (CDC). National health and nutrition examination survey data [EB/OL]. (2014-09) [2024-03-01]. <https://www.cdc.gov/nchs/nhanes/index.htm>.
- [17] ZOU H, ZHANG H H. On the Adaptive Elastic-Net with a diverging number of parameters [J]. *Ann Stat*, 2009, 37(4): 1733-1751.
- [18] CHEUNG B Y, NEIL THOMAS G. The metabolic syndrome and vascular disease in Asia [J]. *Cardiovascular Haematol Disord Drug Targets*, 2007, 7(2): 79-85.
- [19] 苏彦萍, 杨昆, 刘相佟, 等. 北京市成年人饮食行为与慢性病发病风险研究 [J]. *预防医学*, 2021, 33(2): 111-116.
- [20] SANEEI P, SALEHI-ABARGOUEI A, ESMAILL ZADEH A, et al. Influence of Dietary Approaches to Stop Hypertension (DASH) diet on blood pressure: a systematic review and meta-analysis on randomized controlled trials [J]. *Nutr Metab Cardiovasc Dis*, 2014, 24(12): 1253-1261.
- [21] DE PERGOLA G, D'ALESSANDRO A. Influence of mediterranean diet on blood pressure [J]. *Nutrients*, 2018, 10(11): 1700.
- [22] NDANUKO R N, TAPSELL L C, CHARLTON K E, et al. Dietary patterns and blood pressure in adults: a systematic review and meta-analysis of randomized controlled trials [J]. *Adv Nutr*, 2016, 7(1): 76-89.

收稿日期: 2024-02-18; 修回日期: 2024-03-14