

挥发性有机物暴露与代谢功能障碍相关脂肪性肝病患病风险存在正向关联

李凯, 曾文倩, 张艳孜, 朱秀玲, 郭冰

四川大学华西公共卫生学院//四川大学华西第四医院流行病学与卫生统计学系, 四川 成都 610041

摘要:目的 评估一般成年人群尿液挥发性有机化合物代谢物(mVOCs)暴露水平与代谢功能障碍相关脂肪性肝病(MASLD)患病风险之间的关联。方法 基于2011~2018年美国国家健康与营养调查的四轮横断面调查,采用广义线性模型评估单一mVOC与MASLD患病风险之间的关联,进一步构建最小绝对收缩和选择算子-加权分位数(LASSO-WQS)两阶段回归模型探究mVOCs混合暴露与MASLD患病风险之间的联系,并量化各化合物的相对贡献。结果 单一暴露分析结果表明,在调整混杂因素后,2-氨基噻唑啉-4-羧酸(ATCA)、N-乙酰-S-2-羧乙基-L-半胱氨酸(CEMA)和N-乙酰-S-3,4-二羟基丁基-L-半胱氨酸(DHBMA)与MASLD患病风险之间存在显著正向关联。混合暴露两阶段分析显示,第一阶段LASSO回归筛选出6种与MASLD患病风险更为相关的mVOCs,第二阶段WQS回归提示mVOCs混合暴露与MASLD患病风险之间存在显著的正向关联关系(OR=1.306, 95% CI: 1.132~1.507, P<0.001),其中CEMA贡献权重最大(36%)。结论 VOCs对肝脏可能具有潜在健康风险,其中CEMA是VOCs混合暴露中贡献最大的独立风险因子,建议开展机制研究验证其肝毒性通路,并考虑将其纳入优先管控清单。

关键词:代谢功能障碍相关脂肪性肝病;挥发性有机物;混合暴露;加权分位数和回归

Exposures to volatile organic compounds are positively correlated with risks of metabolic dysfunction-associated steatotic liver disease

LI Kai, ZENG Wenqian, ZHANG Yanzi, ZHU Xiuling, GUO Bing

Department of Epidemiology and Health Statistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu 610041, China

Abstract: Objective To assess the association between urinary levels of volatile organic compound metabolites (mVOCs) and risks of metabolic dysfunction-associated steatotic liver disease (MASLD) in the general adult population. **Methods** Based on 4 cycles of cross-sectional surveys from the 2011-2018 National Health and Nutrition Examination Survey (NHANES), generalized linear models were employed to evaluate the associations between individual mVOC exposures and the risk of MASLD. A two-stage Least Absolute Shrinkage and Selection Operator-Weighted Quantile Sum (LASSO-WQS) regression model was constructed to investigate the relationship between mixed mVOCs exposures and MASLD risk, and the relative contributions of the individual compounds were quantitatively analyzed. **Results** The single-exposure analysis revealed significant positive associations of 2-aminothiazoline-4-carboxylic acid (ATCA), N-acetyl-S-(2-carboxyethyl)-L-cysteine (CEMA), and N-acetyl-S-(3,4-dihydroxybutyl)-L-cysteine (DHBMA) with MASLD risk after adjusting for confounders. In the two-stage mixed-exposure analysis, the first-stage LASSO regression identified 6 mVOCs with stronger association with MASLD risk. The second-stage WQS regression demonstrated a statistically significant positive association between mixed mVOCs exposures and MASLD risk (OR=1.306, 95% CI: 1.132-1.507; P<0.001), with CEMA contributing the highest weight (36%). **Conclusion** The study reveals a significant positive association between urinary levels of mVOCs mixtures and MASLD risk, suggesting potential hepatotoxic effects of VOC (especially CEMA) exposures, which urges future mechanistic studies of VOC mixture-related health impacts and listing of CEMA for health risk control.

Keywords: metabolic dysfunction-associated steatotic liver disease; volatile organic compounds; mixed exposure; weighted quantile sum regression

肝脏是机体新陈代谢及维持稳态的重要器官,也是抵御内源性和外源性毒素的关键屏障^[1]。代谢功能障碍相关脂肪性肝病(MASLD)是最普遍的慢性肝病,可进展为代谢功能障碍相关脂肪性肝炎、纤维化等更为严重的病理状态,显著增加肝硬化与肝癌风险^[2],累及

全球约30%人口^[3]。环境挥发性有机物(VOCs)是易挥发的低分子含碳有机化合物^[4],其具有强自由基特性和高挥发性^[5],易通过呼吸途径进入人体而直接诱导氧化应激反应危害健康^[6],或在光化学反应作用下转化为其他环境污染物间接造成健康损害^[7]。实验研究已证实VOCs对啮齿动物的肝脏功能具有毒理学效应^[8,9];流行病学研究也发现mVOCs暴露与非酒精性脂肪肝病(NAFLD)的患病风险增加及肝功能指标异常存在关联^[10,11]。MASLD相较于NAFLD将心脏代谢风险因素纳入诊断标准^[12],其可以更全面反映个体代谢状态的肝

收稿日期:2025-07-07

基金项目:国家自然科学基金(82103943)

Supported by National Natural Science Foundation of China (82103943).

作者简介:李凯,在读硕士研究生,E-mail: likai05132024@163.com

通信作者:郭冰,副研究员,硕士生导师,E-mail: guobing0111@scu.

edu.cn

脏表现^[13]。目前,普通人群中 mVOCs 混合暴露与 MASLD 患病风险之间的关联尚未系统阐明。此外,人体实际暴露场景中常同时接触高度相关的多源 VOCs 混合物,传统回归模型在处理高度相关的多种污染物面临多重共线性问题,要么分析单一 mVOC^[10],要么仅将相同或相似的 mVOCs 进行简单加和处理^[11],难以稳定地识别出真正驱动健康结局相关的特定污染物组合^[14],导致结果缺乏明确的生物学或毒理学解释^[14],无法全面准确地认识其对肝脏的危害。

因此,本研究旨在探究尿液中 mVOCs 混合暴露与 MASLD 患病风险之间的关联,基于最新 MASLD 诊断标准,应用 LASSO 回归筛选 MASLD 患病风险显著相关的 mVOCs 高风险暴露簇,与混合效应模型 WQS 进行结果互补,量化高风险暴露簇 mVOCs 对 MASLD 患病的相对贡献。通过 LASSO 筛选关键污染物组合,为我国在复杂新污染背景下精准识别环境健康高危人群和制定差异化协同治理政策,提供了科学参考。本研究有望为制定基于 MASLD 危害的 VOCs 管控策略提供直接证据。

1 资料和方法

1.1 研究对象

本研究以 NHANES 2011~2018 年 4 个周期的调查人群为研究对象。NHANES 是由美国国家卫生统计中心开展的一项全国性、分层多阶段的概率抽样调查,旨

在评估一般人群的健康状况^[15],采集信息包括基础人口统计信息、膳食信息、实验室检查数据、体格检查数据和问卷数据。根据研究目的,本研究纳入 18 岁及以上、进行了尿液 mVOCs 检测、且具有结局 MASLD 相关判断信息的研究对象 9085 名。进一步排除对研究结果会产生干扰的参与者:因协变量缺失会降低统计检验的能力,排除协变量数据缺失者($n=5278$);因妊娠期生理状态发生显著变化,排除孕妇($n=30$);鉴于过量饮酒可能导致代谢性与酒精性脂肪肝的病难以区分,本研究基于问卷采集的酒精摄入频率数据排除中重度饮酒者($n=1420$),定义为:过去 12 个月内饮酒且女性饮酒量 ≥ 2 杯/d、男性 ≥ 3 杯/d^[16];病毒性肝炎对肝脏病理改变具有潜在影响,其引发的肝脂肪变性与代谢功能障碍之间的相关性较低,因此排除血清学检测乙型或丙型肝炎病毒阳性的参与者($n=235$)。最终纳入 2122 名研究对象。

1.2 暴露和结局的定义

暴露变量:根据 NHANES 介绍文档^[17],VOCs 浓度是经超高效液相色谱-电喷雾串联质谱对尿液样品进行实验室分析获得,所有检测流程均遵循标准化操作程序并实施严格的质量控制措施^[17]。本研究筛选出检出率 $>80\%$ 的 mVOCs 作为暴露变量,共计 15 种(表 1);进一步对浓度低于检出限(LOD)的尿液 mVOCs 数据,采用 $LOD/\sqrt{2}$ 进行赋值以降低统计偏差。

表 1 检出率高于 80% 的 15 种 mVOCs

Tab.1 Fifteen types of volatile organic compound metabolites (mVOCs) with detection rates exceeding 80% in the urine samples

Biomarker code	Abbreviation	Detection rate	Chemical name
URX2MH	2-MHA	92.6%	2-Methylhippuric Acid
URX34M	3,4-MHA	99.5%	3,4-Dimethylhippuric Acid
URXAMC	AMCC	99.7%	N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine
URXATC	ATCA	94.0%	2-Aminothiazoline-4-carboxylic Acid
URXBMA	BMA	99.4%	N-Acetyl-S-(benzyl)-L-cysteine
URXCEM	CEMA	99.1%	N-Acetyl-S-(2-carboxyethyl)-L-cysteine
URXCYM	CYMA	87.2%	N-Acetyl-S-(2-cyanoethyl)-L-cysteine
URXDHB	DHBMA	99.9%	N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine
URXHP2	2HPMA	95.2%	N-Acetyl-S-(2-hydroxypropyl)-L-cysteine
URXHPM	3HPMA	99.8%	N-Acetyl-S-(3-hydroxypropyl)-L-cysteine
URXMAD	MA	98.8%	Benzoic Acid
URXPHG	PGA	99.4%	Phenylglyoxylic Acid
URXAAM	AAMA	99.9%	N-Acetyl-S-(2-carbamoylethyl)-L-cysteine
URXMB3	MHBMA3	97.2%	N-Acetyl-S-(4-hydroxy-2-buten-1-yl)-L-cysteine
URXPMM	HMPMA	99.9%	N-Acetyl-S-(1-methyl-3-hydroxypropyl)-L-cysteine

结局变量:MASLD 的诊断标准为肝脏存在脂肪堆积,且同时伴有至少一项心脏代谢风险因素异常^[13]。首

先,利用脂肪肝指数(FLI)评估是否存在脂肪堆积^[18]。FLI 是经过验证的可反映肝脏脂肪堆积的临床生化指

标^[19],计算所需数据包括体质指数(BMI)、腰围(WC)、甘油三酯(TG)和 γ -谷氨酰转氨酶(GGT)水平,FLI \geq 60诊断为存在肝脏脂肪堆积,计算公式如下:

$$FLI = \left(\frac{e^{0.953 \times \log_e(TG) + 0.139 \times BMI + 0.718 \times \log_e(GGT) + 0.053 \times WC - 15.745}}{1 + e^{0.953 \times \log_e(TG) + 0.139 \times BMI + 0.718 \times \log_e(GGT) + 0.053 \times WC - 15.745}} \right) \times 100$$

进一步,结合以下心脏代谢风险因素条件之一判定MASLD:非西班牙裔白种人BMI \geq 25 kg/m²、非西班牙裔亚洲人BMI \geq 23 kg/m²,或男性腰围 $>$ 94 cm、女性腰围 $>$ 80 cm;空腹血糖 \geq 5.6 mmol/L(或糖化血红蛋白 \geq 5.7%)或已确诊2型糖尿病并接受治疗;血压 \geq 130/85 mmHg或正在服用降压药;血浆甘油三酯 \geq 1.70 mmol/L(150 mg/dL)或正在接受降脂治疗;男性高密度脂蛋白胆固醇 \leq 1.0 mmol/L(40 mg/dL)、女性 \leq 1.3 mmol/L(50 mg/dL)或正在接受降脂治疗。

1.3 协变量的定义

本研究选取的协变量基于既往研究中发现的可能影响尿液mVOCs浓度与MASLD相关的因素,涵盖人口统计学特征(年龄、性别、种族、家庭收入水平、婚姻状况)、健康状况(糖尿病、高血压患病情况)、实验室检测指标(总胆固醇水平)、问卷访谈内容(饮食模式、吸烟状况)。其中,饮食模式属于二分类变量,其判定依据是从问卷数据中提取的受试者是否具有特殊饮食模式。所有协变量数据均可以从NHANES的人口统计信息、实验室检查数据和问卷数据中提取,用于综合调整潜在混杂因素。

1.4 统计学分析

描述性分析部分,采用均数 \pm 标准差描述连续变量,并对于满足正态性的数据通过 t 检验或对于满足非正态数据通过Mann-Whitney U 检验比较MASLD组与非MASLD组的差异;采用频数(百分比)形式描述分类变量,组间差异通过卡方检验评估。鉴于尿液浓度会对检测结果产生干扰,故采用尿肌酐校正法对mVOCs浓度进行标准化处理。尿液中mVOCs浓度呈右偏分布,所以进行自然对数转换以符合正态分布假设,以便于后续分析。

关联性分析部分,首先探索mVOC单一暴露与MASLD患病风险的关联(单暴露分析),进一步估计mVOCs混合暴露与MASLD患病风险的综合关联(混合暴露分析)。单暴露分析中,采用广义线性回归模型估计关联大小,并进一步计算比值比(OR)及95%置信区间(CI)^[20]。混合暴露分析中,构建LASSO-WQS两阶段回归模型^[21]:第一阶段,采用LASSO回归模型来克服暴露变量之间可能的多重共线性,应用10倍交叉验证来计算LASSO中的最小均方误差(MSE)和最佳超参数 λ ,确定与导致MASLD风险相关的mVOCs;第二阶段,将选定的代谢物纳入WQS回归模型,以评估混合

mVOCs与MASLD的相关性,并估计每种代谢物的贡献。在WQS分析中,考虑到VOCs暴露对肝脏的不利影响,假设方向为正,将40%的数据作为训练集,60%为验证集,通过5000次bootstrap抽样过程确定各化学物质的权重。

在上述关联性分析模型中,均采用两种协变量调整策略:模型1未调整任何协变量,模型2调整年龄、性别、种族、婚姻状况、收入水平、饮食类型、高血压、糖尿病、总胆固醇水平等潜在混杂因素。为比较不同亚组之间的效应差异,本研究根据性别和血清中可替宁的浓度(吸烟暴露生物标志物)将人群分为不同亚组,比较不同亚组间mVOCs混合暴露与MASLD患病风险的关联的异质性;为验证结果的稳定性,本研究进一步纳入2005~2006年调查周期的数据以及交换WQS分析集的权重来开展敏感性分析。所有统计分析均基于R4.4.3完成,设定双侧 $P < 0.05$ 为差异具有统计学意义。

2 结果

2.1 研究人群的一般特征

本研究共纳入2122名研究对象,平均年龄为45.30岁,其中55%为男性,MASLD患病率为45.85%(表2)。MASLD组患者呈现出年龄更大,体质量指数(BMI)、总胆固醇及尿肌酐水平更高的特征。统计分析表明,两组在年龄、性别、收入水平、婚姻状况,以及糖尿病和高血压患病情况、总胆固醇水平、尿肌酐水平方面的差异均具有统计学意义($P < 0.05$),受教育程度这一协变量在两组间的差异不具有统计学意义($P > 0.05$)。

2.2 单一mVOC的暴露水平与MASLD的关联

为了评估暴露于单一mVOCs成分的健康风险,进行了广义线性回归(表3)。经多变量调整模型校正潜在混杂因素后,发现纳入研究的mVOCs中3种成分(ATCA、CEMA和DHBMA)与患MASLD的风险增加呈显著正相关。在单一mVOC与MASLD关联之间未观察到有意义的负向关联。

2.3 混合mVOCs与MASLD的关联

Spearman相关分析显示(图1),15种尿液mVOCs之间存在显著的正相关关系,表明这些mVOCs之间存在高度相关性和共线性,因此执行LASSO回归模型以筛选一组与MASLD更具关联的mVOCs暴露簇。在最佳拟合LASSO模型中选择的具有非零系数的mVOCs共有6种(图2A),分别是ATCA、BMA、CEMA、DHBMA、MA、HMPMA,被认为与MASLD更相关。图2B表示LASSO回归中MSE变异与对数转换 λ 之间的关联,使用MASLD为响应变量,选择0.00485作为MSE最小的最优 λ 值。

基于LASSO分析的结果,使用正向WQS模型来

表2 研究对象的一般特征

Tab.2 General characteristics of the study population

Variable	Overall (n=2122)	Non-MASLD (n=1149)	MASLD (n=973)	P
Age (year, Mean±SD)	45.30 (14.17)	43.43 (14.63)	47.52 (13.27)	<0.001
Gender [n (%)]				<0.001
Male	1168 (55%)	602 (52%)	566 (58%)	
Female	954 (45%)	547 (48%)	407 (42%)	
Education level [n (%)]				0.13
Less than high school	354 (17%)	181 (16%)	173 (18%)	
High school or equivalent	413 (19%)	212 (18%)	201 (21%)	
College or above	1355 (64%)	756 (66%)	599 (62%)	
Income level [n (%)]				0.03
Low income	519 (24%)	283 (25%)	236 (24%)	
Middle income	804 (38%)	408 (36%)	396 (41%)	
High income	799 (38%)	458 (40%)	341 (35%)	
Marital status [n (%)]				<0.001
Married	1400 (66%)	734 (64%)	666 (68%)	
Divorced	329 (16%)	166 (14%)	163 (17%)	
Single	393 (19%)	249 (22%)	144 (15%)	
BMI (kg/m ² , Mean±SD)	29.19 (6.83)	24.88 (3.48)	34.27 (6.31)	<0.001
Diabetes [n (%)]	216 (10%)	63 (5.5%)	153 (16%)	<0.001
Hypertension [n (%)]	620 (29%)	219 (19%)	401 (41%)	<0.001
Total cholesterol level (mmol/L, Mean±SD)	4.94 (1.06)	4.80 (1.00)	5.11 (1.10)	<0.001
Urinary creatinine level (mg/100 mL, Mean±SD)	122.22 (79.13)	113.38 (76.69)	132.66 (80.72)	<0.001

表3 单一mVOC与MASLD之间的关联

Tab.3 Correlation between single mVOC and metabolic dysfunction-associated steatotic liver disease (MASLD)

VOC	Model 1		Model 2	
	OR (95% CI)	P	OR (95% CI)	P
2-MHA	1.02 (0.77-1.36)	0.871	1.05 (0.93-1.18)	0.427
3,4-MHA	0.95 (0.78-1.15)	0.582	1.03 (0.94-1.13)	0.532
AMCC	0.85 (0.48-1.48)	0.558	0.98 (0.90-1.08)	0.743
ATCA	1.06 (0.95-1.18)	0.279	1.13 (1.03-1.25)	0.008
BMA	1.02 (0.94-1.12)	0.585	1.07 (0.94-1.22)	0.297
CEMA	1.98 (1.53-2.56)	<0.001	1.28 (1.11-1.46)	<0.001
CYMA	0.76 (0.55-1.04)	0.086	0.98 (0.89-1.08)	0.682
DHBMA	1.33 (1.14-1.56)	<0.001	1.29 (1.17-1.43)	<0.001
2HPMA	1.00 (0.91-1.09)	0.975	1.03 (0.94-1.12)	0.539
3HPMA	0.59 (0.46-0.76)	<0.001	1.02 (0.93-1.12)	0.697
MA	1.12 (0.60-2.09)	0.713	1.17 (0.94-1.46)	0.164
PGA	0.95 (0.54-1.66)	0.849	1.15 (0.94-1.42)	0.181
AAMA	0.81 (0.67-0.99)	0.036	1.00 (0.91-1.10)	0.963
MHBMA3	0.87 (0.64-1.18)	0.357	1.05 (0.95-1.15)	0.359
HMPMA	1.57 (1.11-2.21)	0.011	1.09 (0.99-1.21)	0.083

Model 1: No covariate adjustment performed; Model 2: Adjusted for age, gender, ethnicity, dietary pattern, income level, marital status, occupation, diabetes, hypertension, smoking status, and total cholesterol level.

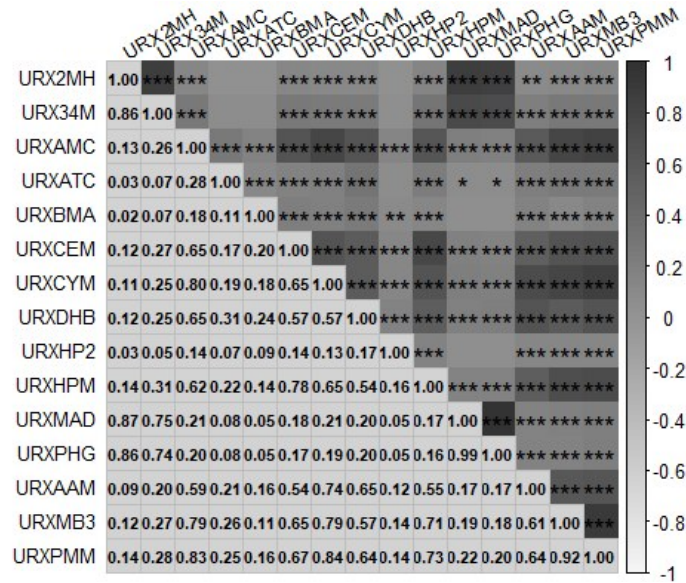


图1 纳入研究的15种mVOCs之间的相关性

Fig.1 Correlation between the 15 mVOCs included in the study. The numbers in the figure represent the correlation coefficients between two VOCs. Darkest shades indicate a strong positive correlation (close to 1), while the light shades indicate a strong negative correlation (close to -1). The asterisks (*) within the squares denote the statistical significance of the correlation coefficients. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

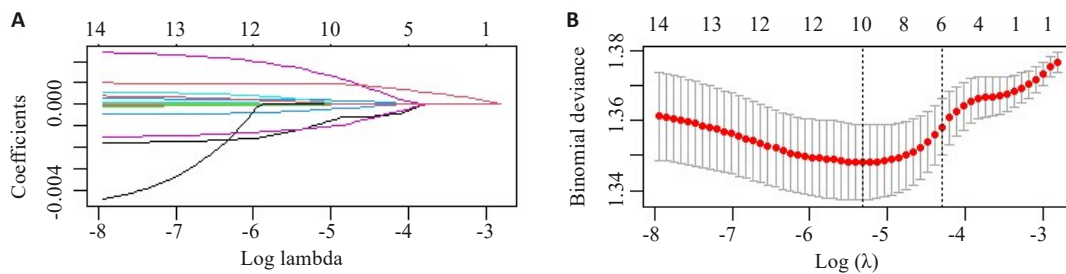


图2 纳入研究的15种mVOCs的LASSO回归分析

Fig.2 LASSO regression analysis of the 15 types of mVOCs included in the study. **A**: Screening pathway for mVOCs associated with MASLD risk. **B**: Correlation between log-transformed λ and MSE. The red dashed line and its error bars in **B** represent the mean MSE values and their corresponding 95% confidence intervals (95% CI). The left black dashed line indicates the optimal λ value yielding the minimum MSE, while the right black dashed line shows the λ value from the simplest model obtained at one standard error below the minimum MSE.

探索多种尿液中mVOCs混合暴露的联合效应并量化各成分的相对贡献权重(图3)。在调整混杂因素前后WQS分析的结果显示:尿液中mVOCs混合暴露与MASLD患病风险之间存在统计学上显著的正向关联(未调整模型:OR=1.344, 95% CI: 1.193~1.515, $P < 0.001$, 调整混杂因素后的模型:OR=1.306, 95% CI: 1.132~1.507, $P < 0.001$)。在这些mVOCs中,贡献最大的3个分别是:CEMA、DHBMA和BMA,其估计权重分别为36.0%、31.1%和14.7%(图3)。

2.4 亚组分析与敏感性分析

基于WQS模型,根据性别(男性/女性)以及基于血清可替宁浓度划分的吸烟状态进行分层分析,其中吸烟状态定义为:当前吸烟者(>3.08 ng/mL)、被动吸烟者

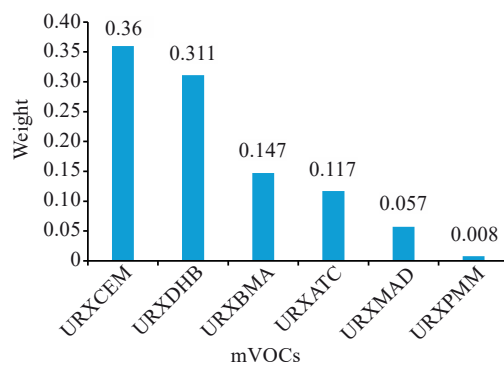


图3 WQS模型正向回归mVOCs混合物与MASLD的权重图

Fig. 3 WQS model forward regression of mVOCs mixtures with MASLD weight plots. The chart above were adjusted for age, gender, ethnicity, dietary patterns, income level, marital status, occupation, diabetes, hypertension, smoking status, and total cholesterol levels.

(二手烟暴露者, >0.015 ng/mL 且 ≤3.08 ng/mL) 和未吸烟者 (≤0.015 ng/mL)。经对混杂因素校正后的亚组进行WQS分析(表4), 男性和女性尿液中mVOCs水平与MASLD之间的关联性差异无统计学意义($P>0.05$)。吸烟状况的分组中, 不吸烟组尿液中mVOCs水平与MASLD

患病风险之间的关联不具有显著差异, 被动吸烟组和吸烟组尿液中mVOCs水平与MASLD患病风险具有显著的正向关联, 且吸烟组的患病风险较被动吸烟组更大。亚组分析权重图(图4), 吸烟组所占权重最大的mVOC为CEM, 被动吸烟组所占权重最大的mVOC为MA。

表4 不同分组mVOCs与MASLD的关联

Tab.4 Associations between different groupings of mVOCs and MASLD

Subgroup	OR (95% CI)	P
Gender		
Male	1.21 (0.99, 1.47)	0.056
Female	1.06 (0.86, 1.30)	0.579
Smoking status		
Non-smoker	1.10 (0.83, 1.46)	0.501
Secondhand smoker	1.27 (1.02, 1.57)	0.031
Current smoker	1.49 (1.19, 1.87)	<0.001

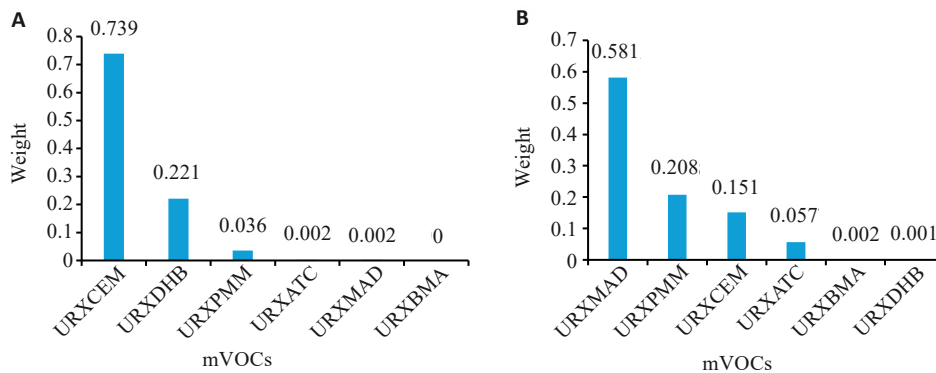


图4 吸烟状况分组中WQS模型正向回归mVOCs混合物与MASLD的权重图

Fig. 4 WQS model forward regression of mVOCs mixtures with MASLD weight plots in groups with different smoking statuses. A: Weight map of mVOCs in the smoking group after adjusting for confounding factors. B: Weight map of mVOCs in the secondhand smoking exposure group after adjusting for confounding factors.

为验证研究结果的稳健性, 采用两种不同策略进行敏感性分析。策略一, 扩充2005~2006年调查周期的数据至主分析数据集; 策略二, 在WQS模型框架下, 通过交换训练集与验证集的权重分配, 对主分析数据集进行交叉验证。两种验证结果一致表明, 尿液中VOCs混合暴露与MASLD患病风险之间的正向关联仍保持不变, 关联结果依然具有统计学意义。

3 讨论

VOCs作为一类具有多重毒理学效应的化学物质, 对人体健康的危害可表现肝脏代谢功能障碍、肾脏排泄能力受损及神经系统传导异常等多个靶器官损害^[22], 部分成分已被我国《新污染物治理行动方案》列为重点管控物质^[23]。与发达国家相比, 我国在VOCs高效降解技术和环境友好型替代品研发方面仍存在明显差距^[24], 这使得

我国由新型污染物引发的健康风险防控面临严峻挑战。

本研究基于NHANES中2011-2018四个周期的数据, 探究从单一暴露到联合暴露不同暴露模式下尿液中mVOCs和MASLD患病风险之间的关联。单一暴露模型结果显示, ATCA、CEMA和DHBMA与MASLD患病风险呈正向关联, 这之前单一尿液中mVOC与NAFLD关联的研究结果相似^[10], 以往研究显示成人尿液中ATCA、CEMA和MUCA与NAFLD具有统计学意义的正向关联。本研究混合暴露模型结果显示, 经LASSO筛选的尿液mVOCs高风险暴露簇与MASLD患病风险呈显著正向关联(OR=1.294, 95% CI: 1.124~1.490, $P<0.001$), 其中CEMA和DHBMA在总体效应中起重要的作用(CEMA权重为36.0%, DHBMA权重为31.1%), 而MA和HMPMA在混合物中的作用较小(MA所占权重为5.7%, HMPMA所占权重为0.8%), 这

与以往研究结果存在部分差异^[11]。以往研究通过WQS模型和BKMR模型观察到mVOC混合物与部分肝脏生物标志物和NAFLD患病风险之间具有显著正相关关系,但在总体效应中起重要的作用是2,3,4-MHA和HMPMA(2,3,4-MHA所占权重为42%,HMPMA所占权重为32%)。这可能是由于mVOCs之间具有高度的相关性和共线性^[21],以往研究并未考虑mVOCs之间的高度相关性,混合暴露效应因为所纳入的暴露簇不同而表现出不同的结果,且研究设计、样本量和种族遗传背景也可能导致结果的差异。

本研究在混合暴露的亚组分析中发现性别差异无统计学意义,可能的原因是由于VOCs代谢通路和毒理机制本身缺乏强烈的性别特异性,VOC与其它健康结局之间关联的性别差异也不具有统计学意义^[25]。另亚组分析显示尿液中mVOCs混合物与MASLD患病风险的关联在不同的血清可替宁浓度分组中存在差异,在吸烟组和被动吸烟组的尿液中mVOCs混合物与MASLD之间存在显著正向关联,而在非吸烟组中尿液mVOCs与MASLD之间不存在统计学关联,类似吸烟状态之间的差异也在mVOCs与炎症的关联中表现出来^[26],烟草烟雾可能是VOCs暴露的主要来源^[27]。未来的研究可以在不同特征的人群中进一步分析VOCs混合物与MASLD之间的关系。

VOCs与MASLD相关的生物学机制目前尚未完全阐明,根据现有研究表明VOCs混合物对MASLD的影响机制可能涉及多个生物学通路和系统。肝脏中VOCs的代谢会消耗谷胱甘肽并产生氧自由基,导致氧化应激增加,氧化损伤和炎症是VOCs诱导肝损伤的常见机制^[28]。本研究表明贡献最大尿液中的VOC为CEMA,在总体效应中所占的权重为36%,CEMA为丙烯醛的代谢产物,其导致肝脏损伤的基本机制涉及多种应激反应^[29,30],包括诱导氧化应激、损害线粒体功能以及激活内质网应激反应^[31],丙烯醛可以通过引发脂质过氧化产物来传播氧化应激^[32],从而通过该途径参与NAFLD的发病过程^[33]。丙烯醛暴露也会增加内质网应激标志物的表达,无法激活保护性的未折叠蛋白,可能通过改变脂蛋白的功能和结构特性^[34],进而促进肝脏脂肪的堆积^[35],增加MASLD的风险。部分芳香族VOCs(如DHBMA、BMA等)作为微塑料的组成成分,可通过肝脏蓄积作用通过物理性损伤直接降低肝细胞活力,从而促进肝内脂质蓄积^[36]。另外,多种VOCs在人体内代谢过程中存在显著的共线性特征,各组分可能存在复杂的交互作用^[37],部分VOCs可能通过竞争性结合同一靶标或通过调节相同的生物途径,增强或削弱彼此的效应。目前关于VOCs与MASLD之间的确切作用机制,仍需通过更多的实验研究结合人群队列研究进一步阐明。

本研究通过LASSO回归解决了mVOCs高度相关性问题的,首次识别出驱动MASLD的核心mVOCs暴露簇,为mVOCs混合暴露与MASLD患病风险之间的关联提供了新的证据。同时,本研究也存在一些局限性。首先,本研究为横断面设计,无法确定尿液中mVOCs与MASLD之间的因果关系;其次,考虑到协变量缺失过多会影响统计效能,故本研究分析使用数据集将协变量缺失者均进行排除,经对分析数据集和目标数据集基线特征比较发现,分析数据集可能存在“健康者偏倚”^[38],这可能会低估VOCs与MASLD之间的关联。尽管本研究考虑多个混杂因素后明确尿液中mVOCs暴露与MASLD患病增加具有统计学意义上的关联,但未测量的其他混杂因素(如他汀类药物使用等)仍然会导致结果的偏差。另脂肪肝的判定依赖于脂肪肝指数(FLI \geq 60),而非影像学或组织学等金标准。尽管最初验证研究提示,FLI的受试者工作特征曲线下面积(AUROC)为0.84^[18],但仍可能引入误分类偏倚(假阳性或假阴性)。FLI常用于大规模流行病学研究中替代超声检查,以平衡可行性与成本,本研究旨在揭示人群层面的潜在关联,未来仍需通过影像学等更精确的测量方法在不同队列中进行验证。本研究使用LASSO回归进行变量筛选,其目的是降维后筛选重要变量,故在建立模型时未考虑交互或非线性关系。考虑到后续混合暴露所用的WQS模型在研究中为正向,所以LASSO回归进行筛选时并未做负向和双向检验,未来研究可采用贝叶斯核机器回归等更复杂模型,同步评估混合暴露的整体效应、非线性关系及成分间交互作用。另外,受问卷信息所限,未能全面收集电子烟及其它烟草制品的使用情况,而血清可替宁作为尼古丁暴露的共同生物标志物,无法区分其具体来源,部分不产生烟雾的烟草使用者可能被误归入暴露组,但是此类错误分类为非差异性错分,可能会低估暴露与结局之间真实关联。此外,尽管与血液中VOCs相比,尿液中VOCs代谢产物的半衰期更长^[39],但在NHANES中,尿液中mVOCs仅作为单一评估进行了一次测量,mVOCs在人体中存在昼夜变化,仅包含一次mVOCs浓度测量的流行病学调查可能会导致测量误差。另在研究结果中发现部分单一暴露(如ATCA、CEMA)的关联强度OR在模型调整协变量后发生了变化显著,可能是有效的混杂控制所致,也可能是优先考虑了控制混杂以获取无偏估计,而未考虑协变量之间可能的共线性而使得OR在校正混杂因素前后出现波动。因此,为了更好地了解VOCs与健康结果之间的关联,需要进一步的前瞻性研究,并且在研究设计和统计分析中考虑更多的因素,以更全面、准确地评估环境VOCs暴露的长期影响。

本研究结果为制定针对挥发性有机物(VOCs)的

公共卫生防控策略提供了直接依据。首先,研究通过统计模型识别出丙烯腈代谢物(CEMA)、1,3-丁二烯代谢物(DHBA)和甲苯代谢物(BMA)等在VOCs混合暴露与MASLD患病风险关联中权重较高的关键成分,提示丙烯腈、甲苯等工业与环境常见污染物应作为肝脏健康风险防控的重点对象,为修订新污染物治理清单和确立干预优先级提供了科学依据。此外,亚组分析表明,吸烟及二手烟暴露会加剧VOCs混合暴露对MASLD患病风险的影响,提示应将吸烟控制与VOCs暴露防护结合推进,加强公众健康教育,倡导减少烟草使用并避免VOCs双重暴露场景。

本研究首次发现mVOCs混合暴露与MASLD患病风险的正向关联,其中CEMA、DHBA和BMA是权重最高的关键化合物。本研究建议将CEMA等优先纳入环境监测,并且在修订职业暴露限值时考虑其代谢毒性。后续研究应着重阐明VOCs暴露与MASLD的时序关系,以及CEMA诱导肝脏脂肪变的具体通路。

Declaration of interests: The authors declare no competing interests.

参考文献:

- [1] Trefts E, Gannon M, Wasserman DH. The liver[J]. *Curr Biol*, 2017, 27(21): R1147-51.
- [2] Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases[J]. *Hepatology*, 2018, 67(1): 328-57.
- [3] Younossi ZM, Golabi P, Paik JM, et al. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): a systematic review[J]. *Hepatology*, 2023, 77(4): 1335-47.
- [4] Inamdar AA, Morath S, Bennett JW. Fungal volatile organic compounds: more than just a funky smell? [J]. *Annu Rev Microbiol*, 2020, 74: 101-16.
- [5] Zhou XH, Zhou X, Wang CM, et al. Environmental and human health impacts of volatile organic compounds: a perspective review [J]. *Chemosphere*, 2023, 313: 137489.
- [6] Lei T, Qian H, Yang JX, et al. The association analysis between exposure to volatile organic chemicals and obesity in the general USA population: a cross-sectional study from NHANES program [J]. *Chemosphere*, 2023, 315: 137738.
- [7] Hanna GB, Boshier PR, Markar SR, et al. Accuracy and methodologic challenges of volatile organic compound-based exhaled breath tests for cancer diagnosis: a systematic review and meta-analysis[J]. *JAMA Oncol*, 2019, 5(1): e182815.
- [8] Wahlang B, Hardesty JE, Head KZ, et al. Hepatic injury caused by the environmental toxicant vinyl chloride is sex-dependent in mice [J]. *Toxicol Sci*, 2020, 174(1): 79-91.
- [9] Wang F, Liu F, Liu H, et al. Effects of immunological and hematological parameter in mice exposed to mixture of volatile organic compounds[J]. *Inhal Toxicol*, 2016, 28(4): 164-9.
- [10] Duan XX, Chen ZH, Liao J, et al. The association analysis between exposure to volatile organic compounds and fatty liver disease in US Adults[J]. *Dig Liver Dis*, 2025, 57(2): 535-41.
- [11] Liu W, Cao ST, Shi D, et al. Single-chemical and mixture effects of multiple volatile organic compounds exposure on liver injury and risk of non-alcoholic fatty liver disease in a representative general adult population[J]. *Chemosphere*, 2023, 339: 139753.
- [12] Targher G, Byrne CD, Tilg H. MASLD: a systemic metabolic disorder with cardiovascular and malignant complications[J]. *Gut*, 2024, 73(4): 691-702.
- [13] Lazarus JV, Newsome PN, Francque SM, et al. Reply: a multi-society Delphi consensus statement on new fatty liver disease nomenclature[J]. *Hepatology*, 2024, 79(3): E93-4.
- [14] Billionnet C, Sherrill D, Annesi-Maesano I. Estimating the health effects of exposure to multi-pollutant mixture[J]. *Ann Epidemiol*, 2012, 22(2): 126-41.
- [15] Johnson CL, Paulose-Ram R, Ogden CL, et al. National health and nutrition examination survey: analytic guidelines, 1999-2010[J]. *Vital Health Stat 2*, 2013(161): 1-24.
- [16] Marti-Aguado D, Calleja JL, Vilar-Gomez E, et al. Low-to-moderate alcohol consumption is associated with increased fibrosis in individuals with metabolic dysfunction-associated steatotic liver disease[J]. *J Hepatol*, 2024, 81(6): 930-40.
- [17] Alwis KU, Blount BC, Britt AS, et al. Simultaneous analysis of 28 urinary VOC metabolites using ultra high performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (UPLC-ESI/MSMS) [J]. *Anal Chim Acta*, 2012, 750: 152-60.
- [18] Bedogni G, Bellentani S, Miglioli L, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population[J]. *BMC Gastroenterol*, 2006, 6: 33.
- [19] Park J, Kim G, Kim BS, et al. The associations of hepatic steatosis and fibrosis using fatty liver index and BARD score with cardiovascular outcomes and mortality in patients with new-onset type 2 diabetes: a nationwide cohort study[J]. *Cardiovasc Diabetol*, 2022, 21(1): 53.
- [20] Niu ZP, Duan ZZ, He WX, et al. Kidney function decline mediates the adverse effects of per- and poly-fluoroalkyl substances (PFAS) on uric acid levels and hyperuricemia risk[J]. *J Hazard Mater*, 2024, 471: 134312.
- [21] Dong R, Chang D, Shen C, et al. Association of volatile organic compound exposure with metabolic syndrome and its components: a nationwide cross-sectional study[J]. *BMC Public Health*, 2024, 24(1): 671.
- [22] Oyerinde AS, Selvaraju V, Babu JR, et al. Potential role of oxidative stress in the production of volatile organic compounds in obesity[J]. *Antioxidants: Basel*, 2023, 12(1): 129.
- [23] 国务院办公厅关于印发新污染物治理行动方案的通知. 中华人民共和国国务院公报(2022.16号).
- [24] 张伟霞, 陈柄旭. 国内外VOCs控制技术面面观[J]. *中国经济周刊*, 2017, (17): 72-3.
- [25] Zhang WP, Lin AQ, Chen WQ. The effect of dietary oxidation balance scores and volatile organic compounds exposures on inflammation[J]. *Ecotoxicol Environ Saf*, 2024, 286: 117163.
- [26] Hu YZ, Niu ZP, Cao CS, et al. Volatile organic compounds (VOC)

- metabolites in urine are associated with increased systemic inflammation levels, and smokers are identified as a vulnerable population[J]. *Ecotoxicol Environ Saf*, 2024, 288: 117398.
- [27] McGraw KE, Domingo-Reloso A, Riggs DW, et al. Exposure to volatile organic compounds and blood pressure in NHANES 2011 to 2018 [J]. *Hypertension*, 2025, 82(1): 136-48.
- [28] Kuang HX, Feng JL, Li ZL, et al. Volatile organic compounds from second-hand smoke may increase susceptibility of children through oxidative stress damage[J]. *Environ Res*, 2022, 207: 112227.
- [29] Mohammad MK, Avila D, Zhang JW, et al. Acrolein cytotoxicity in hepatocytes involves endoplasmic reticulum stress, mitochondrial dysfunction and oxidative stress[J]. *Toxicol Appl Pharmacol*, 2012, 265(1): 73-82.
- [30] Moghe A, Ghare S, Lamoreau B, et al. Molecular mechanisms of acrolein toxicity: relevance to human disease[J]. *Toxicol Sci*, 2015, 143(2): 242-55.
- [31] Wang B, Wang X, Yu LL, et al. Acrylamide exposure increases cardiovascular risk of general adult population probably by inducing oxidative stress, inflammation, and TGF- β 1: a prospective cohort study[J]. *Environ Int*, 2022, 164: 107261.
- [32] Wang B, Liu W, Yu L, et al. Acrolein exposure impaired glucose homeostasis and increased risk of type 2 diabetes: an urban adult population-based cohort study with repeated measures[J]. *Environ Sci Technol*, 2023, 57(18): 7162-73.
- [33] Lang AL, Beier JI. Interaction of volatile organic compounds and underlying liver disease: a new paradigm for risk[J]. *Biol Chem*, 2018, 399(11): 1237-48.
- [34] Tan X, Zhao T, Wang Z, et al. Acrylamide defects the expression pattern of the circadian clock and mitochondrial dynamics in C57BL/6J mice liver and HepG2 cells[J]. *J Agric Food Chem*, 2018, 66(39): 10252-66.
- [35] Chen WY, Wang M, Zhang JW, et al. Acrolein disrupts tight junction proteins and causes endoplasmic reticulum stress-mediated epithelial cell death leading to intestinal barrier dysfunction and permeability[J]. *Am J Pathol*, 2017, 187(12): 2686-97.
- [36] Lu Y, Zhang Y, Deng Y, et al. Uptake and accumulation of polystyrene microplastics in zebrafish (*Danio rerio*) and toxic effects in liver[J]. *Environ Sci Technol*, 2016, 50(7): 4054-60.
- [37] Wang Y, Yu Y, Zhang X, et al. Combined association of urinary volatile organic compounds with chronic bronchitis and emphysema among adults in NHANES 2011-2014: The mediating role of inflammation[J]. *Chemosphere*, 2024, 361: 141485.
- [38] Costello S, Picciotto S, Rehkopf DH, et al. Social disparities in heart disease risk and survivor bias among autoworkers: an examination based on survival models and g-estimation[J]. *Occup Environ Med*, 2015, 72(2): 138-44.
- [39] Konkle SL, Zierold KM, Taylor KC, et al. National secular trends in ambient air volatile organic compound levels and biomarkers of exposure in the United States[J]. *Environ Res*, 2020, 182: 108991.

(编辑:余诗诗)