·流行病与统计方法·

# 基于贝叶斯网络模型的非酒精性脂肪肝发病路径分析

汤建敏[1],王璇[1],陈硕[2],韩玉梅[2],孔邻润[2],杨兴华[1]
1.首都医科大学公共卫生学院,北京 100069;2.北京市体检中心

**摘要:目的**　识别非酒精性脂肪肝(NAFLD)关键影响因素,并构建贝叶斯网络模型探索其发病路径。**方法**　研究基于北京健康管理队列。通过 LASSO 回归筛选 NAFLD 的相关影响因素,使用 10 折交叉验证选择最优惩罚参数 λ。基于筛选出的变量,应用贝叶斯网络模型进行结构学习,并通过贝叶斯信息准则(BIC)和交叉验证评估模型拟合优度,进行路径分析。**结果**　5 年随访期内,30 001 名研究对象中有 22.83%发生 NAFLD($n$=6 850),发病密度为 45.7/1 000 人年。筛选出 6 个关键影响因素:AGE、BMI、HDL-C、FPG、ALT、DBP,并基于这些因素构建贝叶斯网络模型。贝叶斯网络分析揭示了 23 条 NAFLD 发病路径,以 BMI 为根节点的路径最多,有 12 条,其次年龄组有 5 条。其中 AGE→DBP→FPG→NAFLD 路径在 45 岁以下人群中的发病概率显著高于 45 岁以上人群;BMI 为偏低状态下,BMI→HDL→ALT→NAFLD 等路径显示 NAFLD 的发病概率高于其他 BMI 组。**结论**　贝叶斯网络模型揭示了 NAFLD 的发病路径,明确了关键因素对其发病的影响,为 NAFLD 的早期筛查和干预提供了理论依据。

**关键词**:非酒精性脂肪性肝病;LASSO 回归;贝叶斯网络;发病路径
**中图分类号**:R575.5　　**文献标志码**:A　　**文章编号**:1003-8507(2025)01-21-06
**DOI**:10.20043/j.cnki.MPM.202409406

## Analysis of the pathogenesis of non-alcoholic fatty liver disease based on Bayesian network model

TANG Jian-min*, WANG Xuan, CHEN Shuo, HAN Yu-mei, KONG Lin-run, YANG Xing-hua
*School of Public Health, Capital Medical University, Beijing 100069, China*

**Abstract: Objective** To identify key influencing factors of non-alcoholic fatty liver disease (NAFLD) and construct a Bayesian network model to explore its pathogenesis. **Methods** The study was based on the Beijing Health Management Cohort. The LASSO regression was employed to screen for relevant influencing factors of NAFLD, and 10-fold cross-validation was used to select the optimal penalty parameter λ. Based on the selected variables, structural learning was conducted using the Bayesian network model, and the model fit was evaluated through the Bayesian Information Criterion (BIC) and cross-validation for path analysis. **Results** During the 5-year follow-up period, 22.83% of the 30 001 study participants developed NAFLD ($n$=6 850), with an incidence density of 45.7 per 1 000 person-years. Six key influencing factors were identified: AGE, BMI, HDL-C, FPG, ALT, and DBP, and a Bayesian network model was constructed based on these factors. The Bayesian network analysis revealed 23 pathways for the development of NAFLD, with the pathway having BMI as the root node being the most frequent (12 pathways), followed by age groups (5 pathways). The pathway AGE→DBP→FPG→NAFLD showed a significantly higher incidence probability in individuals under 45 years compared to those over 45. In cases of low BMI, the pathway BMI→HDL→ALT→NAFLD indicated a higher probability of NAFLD compared to other BMI groups. **Conclusion** The Bayesian network model elucidated the pathways of NAFLD development and clarified the impact of key factors on its onset, providing a theoretical basis for early screening and intervention of NAFLD.

**Keywords**: Non-alcoholic fatty liver disease (NAFLD); LASSO regression; Bayesian network; Pathogenesis

非酒精性脂肪性肝病 (non-alcoholic fatty liver disease, NAFLD)是一种常见的慢性肝病[1],成年人患病率达 25%～30%[2],在中国部分城市已接近 30%[3]。

NAFLD 不仅会导致肝病进展,还与 2 型糖尿病、心血管疾病等代谢综合征相关[4],并增加肝癌风险,严重影响生活质量[5]。其发病与肥胖、胰岛素抵抗、2 型糖尿病、高血压、血脂异常、遗传、饮食和环境等因素有关[6]。改善这些风险因素可显著降低 NAFLD 风险。

传统研究多采用横断面研究和回归分析,但这些方法难以揭示复杂的因果路径和交互作用。贝叶斯网络因其捕捉变量间复杂依赖关系和因果推断的能力,已广泛应用于流行病学研究中[7]。它不仅能构建风险因素的因果路径,还能量化每个因素的贡献,从而为个体化预防提供依据。本研究基于北京健康管理队列

**基金项目**:北京市自然科学基金(7202010)
**作者简介**:汤建敏(1999—),女,硕士在读,研究方向:流行病与卫生统计学
**通信作者**:杨兴华,E-mail:xinghuayang@ccmu.edu.cn

(BHMC)数据,采用 LASSO 回归筛选 NAFLD 的影响因素,并通过贝叶斯网络模型探索发病路径。LASSO 可有效筛选与 NAFLD 相关的因素,避免过拟合[8]。贝叶斯网络模型通过构建有向无环图揭示 NAFLD 的发病机制,助力更有效的预防策略。

## 1　资料与方法

**1.1　研究对象**　本研究基于 BHMC，为长期动态队列研究,采用整群抽样,所有受试者签署知情同意书。以 2017 年为基线，纳入 2017 和 2022 年均参与体检的个体。排除标准:(1)年龄<18 岁;(2)过量饮酒;(3)未报告疾病史;(4)其他脂肪肝或恶性肿瘤;(5)服用与脂肪肝相关药物;(6)基线患有 NAFLD;(7)数据缺失或无随访。最终纳入 30 001 人,中位随访时间为 3.05 年(95%*CI*: 3.04 ~ 4.04 年)。数据涵盖问卷调查、体检和实验室检查。本研究获得首都医科大学伦理委员会批准(编号:2019SY088)。见图 1。

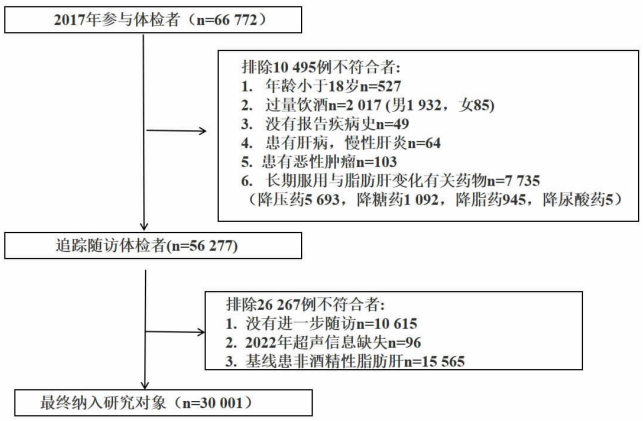



**图 1**　纳入排除流程图

**Figure 1**　Flowchart of inclusion and exclusion criteria

**1.2　诊断标准及定义**　NAFLD 的诊断依据《中国非酒精性脂肪性肝病防治指南(2018 年更新版)》[9]。高尿酸血症:空腹血尿酸>420 pmol/L[10]。血脂异常的标准为:总胆固醇(TC)≥5.2 mmol/L,低密度脂蛋白胆固醇(LDL-C)≥3.4 mmol/L，甘油三酯(TG)≥1.7 mmol/L,高密度脂蛋白胆固醇(HDL-C)≤1.0 mmol/L,或正在接受降脂药物治疗[11]。身体质量指数(BMI)分类标准[12]:偏低为 BMI<18.5 kg/m$^2$;正常为 18.5≤BMI<24 kg/m$^2$;超重为 24 ≤ BMI<28 kg/m$^2$;肥胖为 BMI≥28 kg/m$^2$。腹型肥胖[12]定义为男性腰围(WC)≥90 cm,女性腰围≥85 cm。其他指标包括丙氨酸氨基转移酶(ALT)>40 U/L,HDL-C<1.0 mmol/L,空腹血糖(FPG)>6.1 mmol/L,舒张压(DBP)≥90 mm Hg。

**1.3　统计分析方法**　数据分析使用 SPSS 26.0 和 R 4.4.1 软件,双侧检验,$\alpha$=0.05。计数资料以频数和百分比表示,计量资料以(均数 ± 标准差)表示。

采用 LASSO 回归筛选 NAFLD 的关键影响因素,利用 10 折交叉验证确定最优 λ 值,避免过拟合。基于筛选因素构建贝叶斯网络模型,结构学习使用最大最小爬山算法(MMHC),并结合黑白名单减少错误,参数学习采用极大似然法。模型拟合通过贝叶斯信息准则(BIC)和交叉验证评估,使用 Netica 软件进行可视化,并计算条件概率和边缘概率分析 NAFLD 的关键发病路径。

## 2　结　果

**2.1　研究对象的基线特征描述**　研究纳入 30 001 人,平均年龄(37.53 ± 10.72)岁,男性占 39.34%,女性占 60.66%。随访 5 年内,6 850 人(22.83%)发生 NAFLD,发病密度为 45.7/1 000 人年,基线信息见表 1。

**表 1**　研究对象基线特征[$n$(%),($\bar{x} \pm s$)]

**Table 1**　Baseline characteristics of study participants [$n$(%),($\bar{x} \pm s$)]

| 变量 | 总人群($n$=30 001) | 男性($n$=11 802) | 女性($n$=18 199) | 统计量 | $P$ 值 |
|---|---|---|---|---|---|
| 吸烟 | | | | 6 659.84 | <0.001 |
| 　是 | 4 295 (14.3) | 4 108 ( 34.8) | 187 ( 1.0) | | |
| 　否 | 25 706 (85.7) | 7 694 ( 65.2) | 18 012 ( 99.0) | | |
| BMI 分组 | | | | 2 838.27 | <0.001 |
| 　偏低 | 15 487 (51.6) | 4 133 ( 35.0) | 11 354 ( 62.4) | | |
| 　超重 | 6 933 (23.1) | 3 644 ( 30.9) | 3 289 ( 18.1) | | |
| 　肥胖 | 5 837 (19.5) | 3 732 ( 31.6) | 2 105 ( 11.6) | | |
| 　正常 | 1 744 ( 5.8) | 293 ( 2.5) | 1 451 ( 8.0) | | |
| 高血压 | | | | 2 210.29 | <0.001 |
| 　是 | 6 062 (20.2) | 3 982 ( 33.7) | 2 080 ( 11.4) | | |
| 　否 | 23 939 (79.8) | 7 820 ( 66.3) | 16 119 ( 88.6) | | |
| 血脂异常 | | | | 770.41 | <0.001 |
| 　是 | 10 331 (34.4) | 5 180 ( 43.9) | 5 151 ( 28.3) | | |
| 　否 | 19 670 (65.6) | 6 622 ( 56.1) | 13 048 ( 71.7) | | |

（续表）

| 变量 | 总人群($n$=30 001) | 男性($n$=11 802) | 女性($n$=18 199) | 统计量 | $P$ 值 |
|---|---|---|---|---|---|
| 高尿酸血症 | | | | 4 075.28 | <0.001 |
| 是 | 2 922 ( 9.7) | 2 751 ( 23.3) | 171 ( 0.9) | | |
| 否 | 27 079 (90.3) | 9 051 ( 76.7) | 18 028 ( 99.1) | | |
| 腹型肥胖 | | | | 1 637.82 | <0.001 |
| 是 | 2 743 ( 9.1) | 2 066 ( 17.5) | 677 ( 3.7) | | |
| 否 | 27 258 (90.9) | 9 736 ( 82.5) | 17 522 ( 96.3) | | |
| 年龄 (岁) | 37.53 ± 10.72 | 38.06 ± 11.66 | 37.19 ± 10.05 | 365.55 | <0.001 |
| BMI (kg/m$^2$) | 22.65 ± 2.84 | 23.85 ± 2.75 | 21.88 ± 2.62 | 26.73 | <0.001 |
| ALT (U/L) | 16.54 ± 12.81 | 21.29 ± 15.61 | 13.46 ± 9.40 | 862.89 | <0.001 |
| LDL-C (mmol/L) | 2.74 ± 0.73 | 2.86 ± 0.74 | 2.67 ± 0.72 | 17.18 | <0.001 |
| TC (mmol/L) | 4.66 ± 0.87 | 4.67 ± 0.87 | 4.65 ± 0.86 | 0.09 | 0.025 |
| TG(mmol/L) | 1.02 ± 0.65 | 1.24 ± 0.83 | 0.88 ± 0.46 | 1 093.82 | <0.001 |
| HDL-C(mmol/L) | 1.42 ± 0.32 | 1.26 ± 0.26 | 1.52 ± 0.31 | 238.12 | <0.001 |
| BUN (mmol/L) | 4.47 ± 1.12 | 4.94 ± 1.13 | 4.16 ± 1.00 | 151.96 | <0.001 |
| UA(μmol/L) | 306.35 ± 82.39 | 371.01 ± 75.22 | 264.41 ± 55.24 | 1 085.19 | <0.001 |
| SBP(mm Hg) | 117.06 ± 13.86 | 123.46 ± 13.35 | 112.91 ± 12.54 | 53.59 | <0.001 |
| DBP(mm Hg) | 71.95 ± 9.99 | 75.73 ± 10.04 | 69.51 ± 9.15 | 125.03 | <0.001 |
| AST (U/L) | 17.76 ± 8.17 | 19.51 ± 9.32 | 16.62 ± 7.11 | 193.73 | <0.001 |
| FPG((mmol/L) | 5.05 ± 0.59 | 5.17 ± 0.70 | 4.97 ± 0.48 | 194.34 | <0.001 |
| WC (cm) | 75.79 ± 9.87 | 82.44 ± 7.69 | 71.48 ± 8.66 | 111.93 | <0.001 |

## 2.2　基于 LASSO 回归筛选 NAFLD 发病影响因素

LASSO 回归筛选出 6 个与 NAFLD 发病显著相关的因素：年龄(AGE)、BMI、HDL-C、FPG、ALT 和 DBP，这些变量为贝叶斯网络模型提供了依据，见图 2。

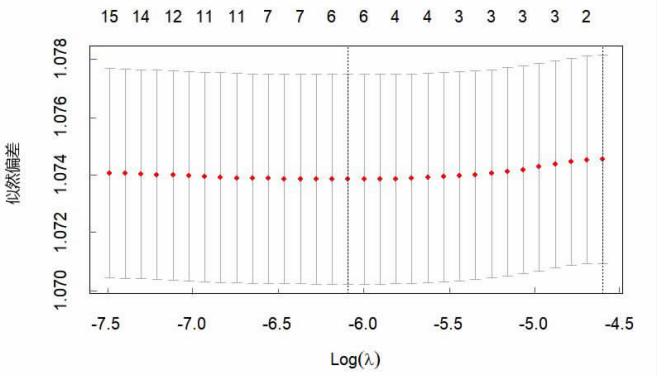



**图 2**　NAFLD 发病影响因素 LASSO 回归交叉验证图

**Figure 2**　Cross-validation plot for Lasso regression of NAFLD risk factors

## 2.3　NAFLD 发病贝叶斯网络学习

基于 LASSO 回归筛选出的 6 个变量，采用 MMHC 构建了 NAFLD 发病的贝叶斯网络模型。最终模型由 7 个节点及 23 条有向无环图(DAG)路径构成，见图 3 和表 2。通过交叉验证评估模型的稳定性和泛化能力，BIC 评分范围为 −9 554.90 ~ −8 169.23，均值为 −8 590.78，表明模型具有良好的稳定性和泛化能力。

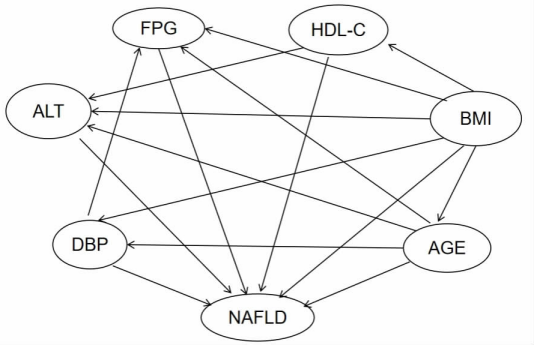



**图 3**　NAFLD 发病网络有向无环图

**Figure 3**　Directed acyclic graph (DAG) of the NAFLD onset network

## 2.4　NAFLD 发病贝叶斯参数学习

通过极大似然估计法对贝叶斯网络进行参数学习，计算各节点在不同父节点状态下的条件概率。构建的 NAFLD 贝叶斯网络包含 23 条因果路径，图 4 显示了各条件概率的可视化分布。模型揭示了年龄(AGE)和 BMI 为关键节点的主要发病路径，具体路径及其相对重要性见表 3 和 4。不同节点在 NAFLD 发病中的影响复杂且相互作用显著。

表 2 和 3 展示了各根节点到 NAFLD 发病的因果路径。将其他节点设为最高水平，比较根节点不同分组的发病概率。结果显示，AGE→ALT→NAFLD 路径在≥45 岁组的发病率最高，为 33.2%；AGE→DBP→FPG→NAFLD 路径在<45 岁组发病率最高，为 39.0%；HDL-C 降低时，HDL-C→ALT→NAFLD 的

发病概率较高，为 31.7%；DBP 升高时，DBP→FPG→NAFLD 的发病概率校高，为 28.2%；ALT 异常和 FPG 升高时，两条路径的概率均为 26.5%；BMI→HDL-C→ALT→NAFLD 路径在正常组和超重组的发病率均为最高；BMI→AGE→ALT→NAFLD 路径在偏低组发病率最高，为 69.0%；BMI→ AGE→ DBP→ FPG→ NAFLD 路径在肥胖组发病率最高，为 32.0%。

**2.5　NAFLD 发病贝叶斯网络推理**　分析结果显示，NAFLD 发病概率在不同年龄、BMI、HDL-C、FPG、ALT 和 DBP 组别存在差异。高年龄、肥胖、HDL-C 降低、FPG 升高、ALT 异常和 DBP 升高的组别患病风险更高，具体数值见表 4。

**表 2**　各根节点到 NAFLD 发病的因果路径

**Table 2**　Causal pathways from root nodes to NAFLD onset

| 路径 | *P*(%) | |
|---|---|---|
| | <45 岁 / 正常 | ≥45 岁 / 异常 |
| AGE→ ALT→ NAFLD | 25.5 | 33.2 |
| AGE→ DBP→ NAFLD | 23.3 | 27.0 |
| AGE→ FPG→ NAFLD | 27.2 | 26.1 |
| AGE→ DBP→ FPG→ NAFLD | 39.0 | 25.0 |
| AGE→ NAFLD | 22.6 | 23.8 |
| HDL-C→ ALT→ NAFLD | 25.6 | 31.7 |
| HDL-C→ NAFLD | 23.0 | 23.1 |
| DBP→ FPG→ NAFLD | 26.1 | 28.2 |
| DBP→ NAFLD | 22.8 | 25.6 |
| ALT→ NAFLD | 22.4 | 26.6 |
| FPG→ NAFLD | 22.9 | 26.5 |

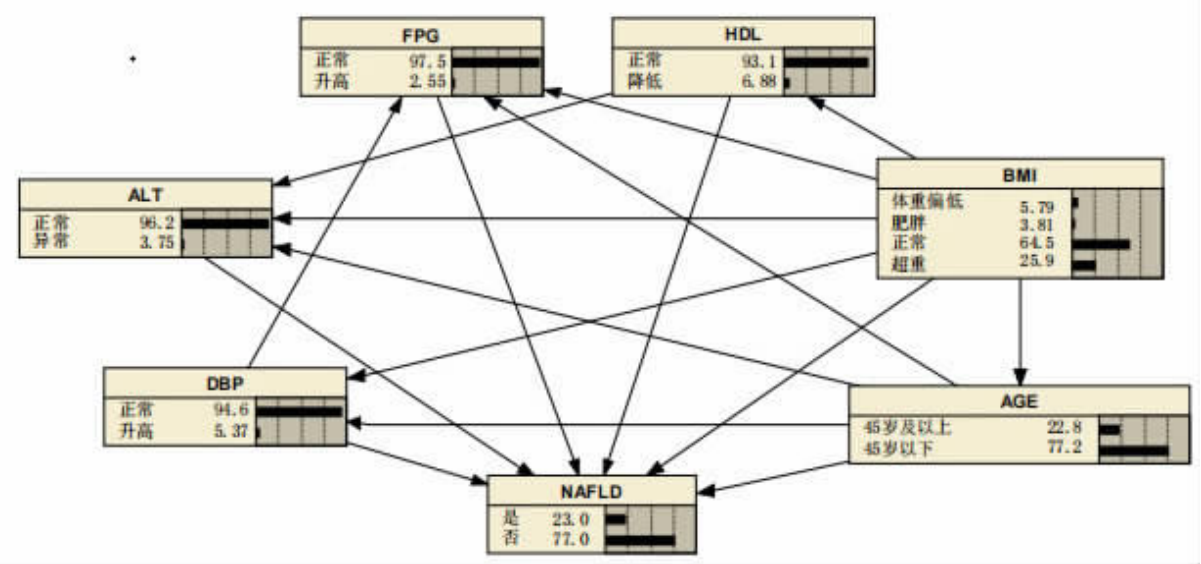



**图 4**　NAFLD 发病的贝叶斯网络模型的概率分布

**Figure 4**　Probability distribution of the bayesian network model for NAFLD onset

**表 3**　BMI 为根节点到 NAFLD 发病的因果路径

**Table 3**　Causal pathways from BMI to NAFLD onset

| 路径 | *P*(%) | | | |
|---|---|---|---|---|
| | 正常 | 偏低 | 超重 | 肥胖 |
| BMI→ HDL-C→ ALT→ NAFLD | 32.6 | 37.6 | 32.3 | 29.0 |
| BMI→ AGE→ ALT→ NAFLD | 31.1 | 69.0 | 29.2 | 31.4 |
| BMI→ AGE→ DBP→ FPG→ NAFLD | 27.1 | 50.0 | 21.6 | 32.0 |
| BMI→ DBP→ FPG→ NAFLD | 25.9 | 50.0 | 28.9 | 29.3 |
| BMI→ ALT→ NAFLD | 25.5 | 41.3 | 27.3 | 25.7 |
| BMI→ AGE→ DBP→ NAFLD | 24.9 | 26.0 | 29.7 | 26.0 |
| BMI→ AGE→ FPG→ NAFLD | 24.4 | 15.9 | 27.7 | 29.6 |
| BMI→ DBP→ NAFLD | 24.4 | 16.6 | 27.5 | 25.7 |
| BMI→ AGE→ NAFLD | 24.3 | 21.1 | 24.5 | 25.2 |
| BMI→ FPG→ NAFLD | 24.2 | 24.1 | 28.4 | 29.5 |
| BMI→ NAFLD | 23.0 | 22.9 | 23.0 | 23.3 |
| BMI→ HDL-C→ NAFLD | 22.4 | 32.6 | 23.0 | 24.6 |

## 3　讨　论

本研究基于 BMHC 数据，使用贝叶斯网络模型分析 NAFLD 的发病路径。通过 LASSO 回归筛选影响因素，构建了多个关键路径，揭示了各节点对 NAFLD 发病的影响。随访期间，6 850 名研究对象发生

**表 4**　各个固定节点不同分组的 NAFLD 发病条件概率比较

**Table 4**　Comparison of conditional probabilities for NAFLD onset across different groupings of fixed nodes

| 固定节点 | NAFLD 发病概率(%) | 概率增加值(%) | 增加幅度(%) |
|---|---|---|---|
| AGE(岁) | | | |
| <45 | 22.6 | | |
| ≥45 | 24.3 | 1.7 | 7.52 |
| BMI | | | |
| 正常 | 23.0 | | |
| 偏低 | 22.9 | –0.1 | –0.43 |
| 超重 | 23.1 | 0.1 | 0.43 |
| 肥胖 | 23.3 | 0.3 | 1.30 |
| HDL–C | | | |
| 正常 | 23.0 | | |
| 降低 | 23.7 | 0.7 | 3.40 |
| FPG | | | |
| 正常 | 22.9 | | |
| 升高 | 26.5 | 3.6 | 15.72 |
| ALT | | | |
| 正常 | 22.8 | | |
| 异常 | 26.6 | 3.8 | 16.67 |
| DBP | | | |
| 正常 | 22.8 | | |
| 升高 | 25.8 | 3.0 | 13.16 |

NAFLD，总发病率为 22.83%，发病密度为 45.92/1 000 人年。通过 LASSO 回归筛选出 6 个与 NAFLD 密切相关的变量，并基于此构建贝叶斯网络模型。结果显示，BMI、FPG 和 ALT 直接影响 NAFLD，而 HDL–C 和 DBP 则通过间接路径影响发病。贝叶斯网络模型直观展示了变量间的因果关系，10 折交叉验证的 BIC 评分表明模型在各子集中的拟合效果一致，具有良好的泛化能力和可靠性[13]。

在以年龄为根节点的路径中，我们发现≥45 岁且伴有 ALT 升高、DBP 升高、FPG 异常的人群 NAFLD 发病概率显著升高，这表明随着年龄的增长，肝脏修复能力下降，而血压升高和血糖异常会显著增加 NAFLD 的患病风险。已有研究表明，早期的代谢异常和血糖调节机制与 NAFLD 发病相关[14]。因此，年轻人应对代谢健康问题保持高度警惕，以有效预防 NAFLD 的发生。在以 BMI 为根节点的路径中，BMI 通过影响多个代谢指标（如 AGE、DBP、FPG 和 ALT）增加了 NAFLD 的风险，且多条路径显示<45 岁组的 NAFLD 发病率较高（例如 AGE→ DBP→ FPG→ NAFLD 路径的发病率为 39.0%），这说明在年轻肥胖人群中，代谢异常是 NAFLD 发病的重要原因。研究还发现，在 BMI → HDL → ALT → NAFLD 等路径中 BMI 为偏低状态时 NAFLD 的发病概率高于其他 BMI 分组。这可能与低体重人群中的营养不良、代谢异常等因素有关[15]。低体重可能导致身体对能量储存和利用异常，影响脂质代谢、增加肝脏的脂肪沉积[16]。此外，BMI → AGE → DBP → FPG → NAFLD 等路径的 NAFLD 发病概率在偏低体重人群中也较高，说明低 BMI 个体在 NAFLD 发病中的特殊风险。因此，体重管理不仅要关注超重和肥胖，也应重视偏低体重的管理。

在 HDL、FPG、ALT 和 DBP 的状态分析中，我们发现 HDL 降低、FPG 升高、ALT 异常和 DBP 升高均显著增加 NAFLD 的发病概率。FPG 升高时 NAFLD 的发病概率增加幅度最大，ALT 升高也与 NAFLD 的发病有强关联，提示肝脏功能异常与胰岛素抵抗是 NAFLD 发病的主要机制之一[17]。ALT 作为肝脏损伤的标志物，其升高反映了肝细胞的炎症反应，进一步促进脂肪肝的病程进展[18]。HDL 通过与 BMI 和 ALT 的交互作用间接影响 NAFLD，表明高密度脂蛋白不仅仅是单一的心血管风险标志物，其在脂肪肝的发病中可能具有复杂的代谢调节作用，HDL–C 水平下降可能通过引发脂质代谢紊乱和氧化应激，增加肝脏负担，进而加剧 NAFLD[19]。DBP 的影响主要体现在其对 BMI 和 FPG 的间接路径上，提示血压管理在 NAFLD 的防治中可能具有重要作用，血压升高可能通过改变血流动力学影响肝脏的血液供应和代谢环境，从而间接加重 NAFLD 的发生[20]。

在所有节点中，BMI 对 NAFLD 发病的贡献率最高，其次是 FPG 和 ALT。BMI 的变化直接影响肝脏的脂质代谢环境，过高或过低的 BMI 均可增加 NAFLD 风险，可能是由于脂肪组织过多或不足对肝脏代谢功能的调节异常。FPG 的升高提示胰岛素抵抗是 NAFLD 的重要致病因素，长期的胰岛素抵抗会增加肝脏脂肪的沉积，从而导致 NAFLD 的发生[21]。ALT 的升高反映了肝细胞损伤和炎症反应，可能进一步促进脂肪肝病程的恶化[22]。HDL 的影响主要表现为对脂质代谢的调节，低水平的 HDL 可能导致脂肪酸运输和清除受阻，进一步加重肝脏的脂肪沉积[23]。而 DBP 可能通过血流动力学改变影响肝脏代谢功能，这些因素的综合作用可能解释了其在 NAFLD 中的间接影响[20]。

本研究利用贝叶斯网络模型探索 NAFLD 的发病路径，这种方法不仅能够识别影响 NAFLD 的多个关键因素，还能揭示这些因素之间的复杂相互作用关系。LASSO 回归的应用使得我们在众多潜在影响因素中精准筛选出最具临床意义的变量。贝叶斯网络模型的因果推断能力和路径分析功能，使得我们能够提供更直观、更全面的 NAFLD 发病机制图谱。本研究的创新之处在于首次将贝叶斯网络与 NAFLD 发病机制的深入分析结合，为揭示疾病的复杂因果关系

提供了新的视角。

本研究也存在一定局限性。研究依赖于观察性数据，虽然通过统计方法控制了部分混杂偏倚，但仍不能完全排除未测量混杂因素的影响。由于随访时间较短，可能无法充分评估长期因素的累积效应。模型的预测结果基于现有数据集，其在其他种群中的适用性仍需进一步验证。此外，本研究未能深入分析药物、饮食和运动等生活方式干预对 NAFLD 路径的调节作用。

综上所述，本研究通过贝叶斯网络模型成功构建了 NAFLD 的发病路径图谱，阐明了关键代谢指标对 NAFLD 发病的复杂机制。这些发现为 NAFLD 的早期筛查和干预提供了重要的理论依据，尤其是在临床实践中针对 BMI、FPG 和 ALT 等指标的管理，可能有效降低 NAFLD 的发病风险。未来的研究应进一步验证这些发现，探索更长随访时间对 NAFLD 发病的影响，以优化防治策略。

**利益冲突声明** 本研究不存在任何利益冲突

**参考文献**

[1] Powell EE, Wong VWS, Rinella M. Non-alcoholic fatty liver disease [J]. Lancet, 2021, 397(10290): 2212-2224.

[2] Quek J, Chan KE, Wong ZY,et al.Global prevalence of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in the overweight and obese population: a systematic review and meta-analysis [J]. Lancet Gastroenterol Hepatol, 2023, 8(1): 20-30.

[3] Li J, Zou BY, Yeo YH, et al. Prevalence, incidence, and outcome of non-alcoholic fatty liver disease in Asia, 1999-2019: a systematic review and meta-analysis [J]. Lancet Gastroenterol Hepatol, 2019, 4 (5): 389-398.

[4] Targher G, Tilg H, Byrne CD. Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach[J]. Lancet Gastroenterol Hepatol, 2021, 6(7): 578-588.

[5] Thomas JA, Kendall BJ, El-Serag HB, et al. Hepatocellular and extrahepatic cancer risk in People with non-alcoholic fatty liver disease[J]. Lancet Gastroenterol Hepatol, 2024, 9(2): 159-169.

[6] Xie JR, Huang HK, Liu ZN, et al. The associations between modifiable risk factors and nonalcoholic fatty liver disease: A comprehensive Mendelian randomization study[J]. Hepatology, 2023, 77(3): 949-964.

[7] Lau CL, Smith CS. Bayesian networks in infectious disease eco-epidemiology[J]. Reviews on Environmental Health, 2016, 31(1): 173-177.

[8] Ali H, Shahzad M, Sarfraz S, et al. Application and impact of Lasso regression in gastroenterology: A systematic review[J]. Indian Journal of Gastroenterology: Official Journal of the Indian Society of Gastroenterology, 2023, 42(6): 780-790.

[9] 中华医学会肝病学分会脂肪肝和酒精性肝病学组，中国医师协会脂肪性肝病专家委员会. 非酒精性脂肪性肝病防治指南(2018 年更新版)[J]. 实用肝脏病杂志，2018，21(2)：177-186.
Liver Disease Group of Chinese Society of Hepatology, Fatty Liver and Alcoholic Liver Disease Group of Chinese Medical Association, Fatty Liver Expert Committee of Chinese Medical Doctor Association. Guidelines of prevention and treatment for nonalcoholic fatty liver disease: a 2018 update [J]. Journal of Practical Hepatology, 2018, 21 (2): 177-186.(In Chinese)

[10] Ningyuan F, Liwei L, Xiaoxi L, et al. Multidisciplinary expert consensus on diagnosis and treatment of hyperuricemia related diseases in China (2023 edition)[J]. Chinese Journal of Practical Internal Medicine, 2023, 43(6): 461-480.

[11] Zengwu W, Yuanlin G. Chinese blood lipid management guidelines (grassroots edition 2024)[J]. Chinese Circulation Journal, 2024, 39 (4): 313-321.

[12] 中华医学会糖尿病学分会. 中.中国成人超重和肥胖症预防与控制专家共识 (2021 年版)[J]. 中华糖尿病杂志，2021，13(5)：315-324.
Chinese Diabetes Society. Consensus of Chinese experts on prevention and control of overweight and obesity in adults (2021 edition)[J]. Chinese Journal of Diabetes, 2021, 13 (5): 315-324.(In Chinese)

[13] Choi T. Bayesian networks with examples in R [J]. Biometrics, 2015, 71(3): 864-865.

[14] Lee CH, Lui DT, Lam KS. Non-alcoholic fatty liver disease and type 2 diabetes: An update[J]. Journal of Diabetes Investigation, 2022, 13 (6): 930-940.

[15] Ahadi M, Molooghi K, Masoudifar N, et al. A review of non-alcoholic fatty liver disease in non-obese and lean individuals [J]. Journal of Gastroenterology and Hepatology, 2021, 36(6): 1497-1507.

[16] Lonardo A, Nascimbeni F, Ballestri S, et al. Sex differences in nonalcoholic fatty liver disease: state of the art and identification of research Gaps[J]. Hepatology, 2019, 70(4): 1457-1469.

[17] Fujii H, Kawada N, Japan study group of nafld J-N. The role of insulin resistance and diabetes in nonalcoholic fatty liver disease[J]. International Journal of Molecular Sciences, 2020, 21(11): 3863.

[18] Lazarus JV, Mark HE, Anstee QM, et al. Advancing the global public health agenda for NAFLD: a consensus statement[J]. Nature Reviews Gastroenterology & Hepatology, 2022, 19(1): 60-78.

[19] Duarte SMB, Stefano JT, Miele L, et al. Gut microbiome composition in lean patients with NASH is associated with liver damage Independent of caloric intake: A prospective pilot study[J]. Nutrition, Metabolism, and Cardiovascular Diseases: NMCD, 2018, 28 (4): 369-384.

[20] Yuan MQ, He J, Hu X, et al. Hypertension and NAFLD risk: Insights from the NHANES 2017-2018 and Mendelian randomization analyses[J]. Chin Med J (Engl), 2024, 137(4): 457-464.

[21] Tanase DM, Gosav EM, Costea CF, et al. The intricate relationship between type 2 diabetes mellitus (T2DM), insulin resistance (IR), and nonalcoholic fatty liver disease (NAFLD)[J]. Journal of Diabetes Research, 2020, 2020: 3920196.

[22] Amernia B, Moosavy SH, Banookh F, et al. FIB-4, APRI, and AST/ALT ratio compared to FibroScan for the assessment of hepatic fibrosis in patients with non-alcoholic fatty liver disease in Bandar Abbas, Iran[J]. BMC Gastroenterology, 2021, 21(1): 453.

[23] Vergès B. Pathophysiology of diabetic dyslipidaemia: where are we? [J]. Diabetologia, 2015, 58(5): 886-899.

**收稿日期**：2024-09-23