

# 糖尿病肾病风险预测模型研究现状

潘琼妮<sup>1</sup>, 胡雯勤<sup>1</sup>, 李雪萍<sup>1</sup>, 宋梅<sup>1</sup>, 李亚<sup>2</sup>, 许刚柱<sup>3</sup>

(1. 西安医学院临床医学院, 陕西 西安 710021;

2. 西安医学院第一附属医院内分泌科, 陕西 西安 710003;

3. 西安医学院第一附属医院神经外科, 陕西 西安 710003)

**摘要:**随着社会经济的发展和人民生活水平的提高, 糖尿病患病率逐年增高, 糖尿病肾病等糖尿病并发症给患者、家庭和社会带来了沉重的负担。临床预测模型可以评估受试者当前患有某种疾病或将来发生某种结局的可能性。本文对目前糖尿病肾病风险的预测模型作一综述, 其中主要包括列线图模型、Logistic 回归模型、评分表、基于机器学习不同算法构建模型, 旨在为构建实用型糖尿病肾病风险预测模型提供理论依据。

**关键词:**糖尿病肾病; 风险预测模型; 列线图模型; Logistic 回归模型

**中图分类号:** R587.2; R692.9

**文献标识码:** A

**DOI:** 10.3969/j.issn.1006-1959.2022.08.011

**文章编号:** 1006-1959(2022)08-0043-04

## Research Status of Risk Prediction Model of Diabetic Nephropathy

PAN Qiong-ni<sup>1</sup>, HU Wen-qin<sup>1</sup>, LI Xue-ping<sup>1</sup>, SONG Mei<sup>1</sup>, LI Ya<sup>2</sup>, XU Gang-zhu<sup>3</sup>

(1. The Clinical Medicine College, Xi'an Medical University, Xi'an 710021, Shaanxi, China;

2. Department of Endocrinology, the First Affiliated Hospital of Xi'an Medical University, Xi'an 710003, Shaanxi, China;

3. Department of Neurosurgery, the First Affiliated Hospital of Xi'an Medical University, Xi'an 710003, Shaanxi, China)

**Abstract:** With the development of economy and the improvement of people's living standard, the prevalence of diabetes is increasing year by year. Diabetic complications such as diabetic nephropathy have brought a heavy burden to patients, families and society. A clinical prediction model can assess the likelihood of subjects suffering from a disease or a future outcome. In this paper, the current risk prediction models of diabetic nephropathy are reviewed, including the line graph model, Logistic regression model, scoring table, and different machine learning algorithms. The purpose is to provide a theoretical basis for the construction of practical diabetic nephropathy risk prediction models.

**Key words:** Diabetic nephropathy; Risk prediction model; Line graph model; Logistic regression model

根据国际糖尿病联盟估计, 全球糖尿病患病率到 2030 年将上升至 10.2% (约 5.78 亿人), 到 2045 年将上升至 10.9% (约 7 亿人)<sup>[1]</sup>。我国的糖尿病患病率已上升至 11.2%, 糖尿病知晓率、控制率和治疗率虽有所改善, 但仍处于较低水平<sup>[2]</sup>。糖尿病肾病 (diabetic kidney disease, DKD) 是糖尿病的微血管并发症之一, 约 30% 的糖尿病患者可发展为 DKD, 进而出现终末期肾病, 其与较高的心脏病发病率和死亡率密切相关, 故及早发现和筛查 DKD 的高危患者是防治的关键<sup>[3]</sup>。临床预测模型是使用数学模型来评估受试者当前患有某种疾病或将来发生某种结局的可能性<sup>[4]</sup>。因此, 了解 DKD 的风险预测模型, 有利于医护人员早期识别危险人群, 以便及时采取预防和治疗措施, 从而降低 DKD 的发生率。本文主要对 DKD 相关风险预测模型进行综述, 以期今后构建及使用 DKD 风险预测模型提供参考。

### 1 糖尿病肾病及其预测模型概述

#### 1.1 糖尿病肾病 糖尿病肾病是由糖尿病所致的慢

性肾脏疾病, 是根据尿蛋白升高程度和 (或) 估算肾小球滤过率 (estimated glomerular filtration rate, eGFR) 的下降程度, 同时排除其他原因所致的慢性肾脏疾病 (chronic kidney disease, CKD) 而作出的诊断, 是许多终末期肾病 (end stage kidney disease, ESKD) 患病人群中的最常见原因<sup>[5,6]</sup>。2020 年《亚太肾脏病学会糖尿病肾病临床实践指南》发布, 其中建议在 1 型糖尿病 (type 1 diabetes mellitus, T1DM) 患病后的第 5 年和 2 型糖尿病 (type 2 diabetes mellitus, T2DM) 患病后每年分别利用尿蛋白肌酐比值 (urine albumin to creatinine ratio, UACR) 进行 DKD 的筛查<sup>[7]</sup>。2021 年版的《基层糖尿病微血管病变筛查与防治专家共识》建议将尿常规、eGFR、UACR 作为 DKD 筛查的首选方法<sup>[8]</sup>。

#### 1.2 疾病风险预测模型

疾病风险预测模型最早被应用在美国氟莱明翰的心脏病研究中, 其使用分类变量开发出一种简单的冠状动脉粥样硬化性心脏病 (coronary heart disease, CHD) 的预测算法, 可预测未发生 CHD 患者的 CHD 患病风险<sup>[9]</sup>。近年来研究者在糖尿病<sup>[10]</sup>、肿瘤<sup>[11]</sup>和脑血管疾病<sup>[12]</sup>等领域逐渐应用疾病风险预测模型来预测疾病。

### 2 糖尿病肾病的危险因素

DKD 发病机制复杂, 有多种危险因素共同参与, 国内外研究<sup>[13-15]</sup>中关于糖尿病肾病的危险因素

基金项目: 陕西省科技计划项目 (编号: 2021ZDLSF02-09)

作者简介: 潘琼妮 (1990.3-), 女, 陕西渭南人, 硕士, 住院医师, 主要从事慢性病的研究

通讯作者: 李雪萍 (1967.10-), 女, 陕西西安人, 博士, 教授, 硕士生导师, 主要从事慢性病的诊疗工作

主要包括:①人口学资料及生活方式:性别、年龄、血压、体重指数、吸烟;②代谢相关指标:血糖、糖化血红蛋白、血脂、尿酸等;③糖尿病病程;④体重方面:超重和肥胖是其危险因素,随着体重指数的增加,DKD 的检出率呈上升趋势,体重指数与 DKD 呈正相关<sup>[16]</sup>;⑤其他:睡眠时间等。有研究发现长睡眠时间( $\geq 8.5$  h)与 DKD 相关,而我国研究发现只有短睡眠时间( $< 6$  h)与 DKD 相关。Tan NYQ 等<sup>[17]</sup>研究了新加坡成年糖尿病患者睡眠时间与 DKD 的关系,发现睡眠过长或过短均与 DKD 有关。

### 3 糖尿病肾病风险预测模型

**3.1 列线图模型** 列线图<sup>[18]</sup>又称诺莫图,是一种在多因素回归分析基础上同时将多个预测指标整合后使用带有刻度的线段绘制在同一平面上的图形,常被用于表达预测模型中各变量之间的关系。目前列线图已被应用于胃癌、浸润性导管癌、前列腺癌和骨肉瘤等肿瘤性疾病的研究中<sup>[19]</sup>,也有学者将其应用在糖尿病及其并发症的研究中。Jiang S 等<sup>[20]</sup>的研究纳入中日友好医院连续行肾活检的 T2DM 患者 302 例,病例的相关数据被随机分成包含 70% 患者的训练集和剩余 30% 患者的验证集,分别用于模型构建和外部验证。纳入包括性别、糖尿病病程、糖尿病视网膜病(diabetic retinopathy, DR)、血尿、糖化血红蛋白(HbA1c)、血红蛋白、血压、尿蛋白排泄和 eGFR 在内的 9 个变量构建列线图。模型 c 指数为 0.934, 内外验证 c 指数值为 0.91、0.875。Xi C 等<sup>[21]</sup>的研究对桂林市 1095 例 T2DM 患者进行问卷调查、体格检查、血常规和生化指标评价,筛选出的危险因素为性别、年龄、高血压、药物使用、糖尿病持续时间、体重指数、血尿素氮水平、血清肌酐水平、中性粒细胞与淋巴细胞的比率和红细胞分布宽度,结合筛选出的危险因素进行逻辑回归分析并建立预测列线图模型,模型 c 指数为 0.819, ROC 曲线下面积(AUC)为 0.813, 内部验证 c 指数为 0.796; 决策曲线分析显示,当风险阈值在 1%~83% 时,糖尿病肾病风险列线图在临床上适用的。这两项研究结果均表明列线图模型对 DKD 的风险评估有一定作用,但这两项研究的样本量均较小,且第 2 项研究未进行外部验证。当然,也有基于社区的样本量较大的研究。Shi R 等<sup>[22]</sup>的研究纳入 4219 例 T2DM 患者,分为单纯 T2DM 组、DKD 组、DR 组和 DR+DKD 组,预测模型中的预测因子包括病程、体重指数、甘油三酯、收缩压、餐后血糖、HbA1c 和尿素氮,模型 c 指数为 0.807, AUC 为 0.807, 内部验证 c 指数达到 0.804, 决策曲线的分析风险阈值为 16%~75%, 说明此模型可预测 DKD 风险。此外,Wang G 等<sup>[23]</sup>的研究选取糖尿病住院患者 2163 例,基于诺莫图建立 4 种不同的筛选方程

(全模型、基于实验室的模型 1、基于实验室的模型 2 和简化模型),4 种模型纳入的因素有所差异,全模型纳入 10 个因素:饮酒状况、高血压、糖尿病持续时间、冠心病史、SBP、总胆固醇、空腹血浆 C 肽、尿酸和糖尿病视网膜病变;模型 1 纳入 7 个因素:性别、SBP、TC、饮酒、高血压、冠心病和糖尿病持续时间;模型 2 在模型 1 的基础上加入 HbA1c;简化模型有 6 个因素:性别、SBP、饮酒、高血压、冠心病和糖尿病持续时间;4 种模型的 c 指数分别为 0.8450、0.8149、0.8171、0.8083,根据 Hosmer-Lemeshow 拟合优度检验,分别为 3.2756、7.749、10.023、12.294,结果表明模型 1、模型 2 模型均具有良好的预测性能和有效性,可用于中国 DKD 病例的筛查。需要说明的是,这两项研究虽然纳入的样本量大,但均未进行外部验证,故其适用性需要进一步考量。

**3.2 基于机器学习的不同算法** 机器学习是一门涉及到概率论、统计学等多门学科的交叉学科。机器学习从当前现有的数据中发现相应的规律,并利用这些规律对未来作出预测,已在当前各个领域中得到广泛应用。在医学领域中应用机器学习算法可有效地降低财力、物力,提高当前医疗系统的运作效率,缓解当前人民群众就医难的问题<sup>[24]</sup>。乔高星<sup>[25]</sup>研究基于机器学习适用不同算法构建预测模型,纳入 2197 例 T2DM 患者,将整体数据随机分组,训练集和验证集的比例约为 7:3,在训练集中,采用单因素 Logistic 回归筛选糖尿病肾病的预测因素,建立 Logistic 回归、人工神经网络(ANN)、朴素贝叶斯分类器(NBC)和分类回归树(CART)糖尿病肾病诊断预测模型,结果发现 CART 模型的准确度和区分度最好,AUC 为 0.7454,校准度  $\chi^2$  为 4.2899,对糖尿病肾病发病的绝对诊断预测与糖尿病肾病实际发生情况一致性较好。林鑫等<sup>[26]</sup>的研究选择 894 条 T2DM 住院数据,利用单因素 Logistic 回归筛选出 24 个有效检查指标作为特征,并基于随机森林、BP 神经网络、支持向量机分别构建模型,结果发现随机森林预测模型的总体性能最优。上述两项研究比较了不同算法构建模型的性能,得出的结论不同,故需要进一步研究探讨不同算法在糖尿病肾病预测方面的价值。在危险因素筛选方面,这两项研究均采用 Logistic 回归,对于危险因素的筛选还有其它方法,如分类树模型。两者在研究因子间的交互作用和混杂因素方面均有应用,Logistic 回归分析作为应用频率较高的模型,其变量共线性较差,而分类树模型因其因变量为分类变量,作为一种非参数回归分析方法,能很好的解决变量间的共线性问题<sup>[27]</sup>。有研究者比较了两种方法在筛选糖尿病危险因素方面的性能,Lou J 等<sup>[28]</sup>分别采用 Logistic 回归和分类树模型

分析 T2DM 患者肾病的危险因素,发现两种模型的结果相似。但廖志波等<sup>[29]</sup>的研究显示,分类树模型在预测糖尿病肾病进展方面效果较好,能较好地分析因素间的交互作用。因此,Logistic 回归和分类树模型的差别可能是前者侧重于糖尿病肾病危险因素的分析,而后者侧重糖尿病肾病进展因素的研究。

**3.3 Logistic 回归模型** Logistic 回归是流行病学和医学中最广泛使用的二元结果建模方法,该模型是一个具有典型连接函数的广义线性模型,通过使用解释变量的值的线性函数来计算事件的概率。当因变量为分类变量时,可使用此模型。此外,对解释变量没有限制被认为是该模型的一个优点<sup>[30]</sup>。

有研究为筛选 DKD 的高危因素<sup>[31]</sup>,利用优化后的 Logistic 回归模型构建 DN 的发病风险预测模型,共纳入 7 项危险因素:LDL-C、SBP、FPG、HbA1c、高血压、糖尿病病程和吸烟,建立的模型为: $0.44x_1+0.06x_2+0.34x_3+0.36x_4+0.53x_5+0.18x_6+0.66x_7$ ,并通过内分泌科的糖尿病患者住院数据进行外部验证,结果显示该模型的 AUC 及 95% CI 分别为 0.662、0.591~0.773,其结果更为可靠,但该模型中的有关系数是根据当地疾病发生率估算的,故而应用于其他地区时应进行适当的调整。此外,也有基于临床资料和超声资料的模型。李楠等<sup>[32]</sup>研究结合三维超声获得的肾脏体积参数和临床资料参数(收缩压、尿蛋白、肾小球滤过率、血浆尿素氮、血肌酐、糖尿病病史、是否有 DR、是否有血尿),使用 Logistic 回归建立预测模型,结果显示该模型的 AUC 为 0.9217,阳性预测值为 0.8824,阴性预测值为 0.8214,但因为此模型需要考虑肾脏超声检查的结果,故在基层医疗卫生机构开展存在一定程度的限制。

**3.4 评分表** 评分表<sup>[33]</sup>以结局作为因变量,筛选出的危险因素作为自变量,根据回归系数  $\beta$  确定不同变量的积分值,以累积积分值的大小判断个体患病的危险性。各变量 10 倍的回归系数作为各变量记分的具体分值,即风险评分表中各变量的具体分值。

Jiang W 等<sup>[34]</sup>纳入了 41271 例 T2DM 患者,筛选出包括年龄、BMI、吸烟、DR、HbA1c、收缩压、HDL-c、TG 和 UACR 在内的 9 个危险因素,对所有危险因素按权重进行评分。根据 Sullivan LM 等<sup>[35]</sup>建议的方法开发出评分表,对所有风险因素进行分类,为每个类别分配权重,然后计算预测模型中所有风险因素的总得分,最后将该模型代入验证组,选择 16 分作为最佳截断值,结果显示该模型的 AUC 为 0.762,灵敏度为 0.847,特异度为 0.667。该研究在回顾性队列中进行了外部验证,并且纳入的危险因素亦属于常用检查指标,故可用于 DKD 的预测和诊断,适用性强,使用方便,非常适合基层医生使用,值得推广。

**3.5 其他** 除了上述的列线图模型、Logistic 回归模型及积分表等,疾病风险计算器亦是一种模型展示形式。如 Heikes KE 等<sup>[36]</sup>使用国内健康和营养检查调查数据开发的糖尿病风险计算器,可计算一个人患未确诊的糖尿病或糖尿病前期的概率。Choi Y 等<sup>[37]</sup>开发的无症状糖尿病患者心血管风险计算器,可预测无症状 T2DM 患者中阻塞性冠状动脉疾病的存在和主要不良心脑血管事件的风险。上述两项研究开发出的疾病风险计算器可预测糖尿病、糖尿病前期及糖尿病患者心血管风险,但目前尚无针对 DKD 的风险计算器,相信今后会有针对 DKD 的风险计算器。

#### 4 总结

目前,国内外研究者已经构建出多种适用于 DKD 患者的风险预测模型,可用于糖尿病患者的筛查、评估等,但大部分预测模型未进行外部验证,且一些模型纳入的危险因素可能不适用于社区卫生服务中心等基层医疗机构。今后可对部分风险预测模型进行外部验证,进一步评估模型的效能,也可研发出针对不同级别医疗机构的模型,优化筛查措施,必要时进行早期干预,在降低医疗成本的同时也为患者、家庭和社会减轻负担。

#### 参考文献:

- [1]Saeedi P,Petersohn I,Salpea P,et al.Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition [J].Diabetes Res Clin Pract,2019(157):107843.
- [2]中华医学会糖尿病学分会.中国 2 型糖尿病防治指南(2020 年版)[J].中华内分泌代谢杂志,2021,37(4):311-398.
- [3]Ahn HS, Kim JH, Jeong H, et al. Differential Urinary Proteome Analysis for Predicting Prognosis in Type 2 Diabetes Patients with and without Renal Dysfunction [J]. Int J Mol Sci, 2020, 21(12):4236.
- [4]Zhou ZR, Wang WW, Li Y, et al. In-depth mining of clinical data: the construction of clinical prediction model with R [J]. Ann Transl Med, 2019, 7(23):796.
- [5]中华医学会糖尿病学分会微血管并发症学组.中国糖尿病肾脏疾病防治临床指南[J].中华糖尿病杂志,2019,11(1):15-28.
- [6]Selby NM, Taal MW. An updated overview of diabetic nephropathy: Diagnosis, prognosis, treatment goals and latest guidelines [J]. Diabetes Obes Metab, 2020, 22(Suppl 1):3-15.
- [7]Liew A, Bavanandan S, Prasad N, et al. Asian pacific society of nephrology clinical practice guideline on diabetic kidney disease [J]. Nephrology (Carlton), 2020, 25(Suppl 2):12-45.
- [8]中国微循环学会糖尿病与微循环专业委员会,中国医疗保健国际交流促进会基层卫生分会基层糖尿病学部,江苏省基层内分泌特色科室孵化联盟.基层糖尿病微血管病变筛查与防治专家共识(2021 年版)[J].中国医学前沿杂志(电子版),2021,13(6):16-38.

- [9]Wilson PW,D'Agostino RB,Levy D,et al.Prediction of coronary heart disease using risk factor categories [J].Circulation,1998,97(18):1837-1847.
- [10]Cheng Y,Shang J,Liu D,et al.Development and validation of a predictive model for the progression of diabetic kidney disease to kidney failure[J].Ren Fail,2020,42(1):550-559.
- [11]Miller EA,Pinsky PF,Heckman-Stoddard BM,et al.Breast cancer risk prediction models and subsequent tumor characteristics[J].Breast Cancer,2020,27(4):662-669.
- [12]Tang EYH,Price CI,Robinson L,et al.Assessing the Predictive Validity of Simple Dementia Risk Models in Harmonized Stroke Cohorts[J].Stroke,2020,51(7):2095-2102.
- [13]Teliti M,Cogni G,Sacchi L,et al.Risk factors for the development of micro-vascular complications of type 2 diabetes in a single-centre cohort of patients [J].Diab Vasc Dis Res,2018,15(5):424-432.
- [14]Molla G,Ismail-Beigi F,Larijani B,et al.Smoking and Diabetes Control in Adults With Type 1 and Type 2 Diabetes: A Nationwide Study From the 2018 National Program for Prevention and Control of Diabetes of Iran [J].Can J Diabetes,2020,44(3):246-252.
- [15]侯新月,胡松,费春晓,等.预测糖尿病患者并发糖尿病肾病风险的列线图模型的建立 [J].临床肾脏病杂志,2020,20(3):207-212.
- [16]Zhang Y,Guo Y,Shen X,et al.Lower body mass index is not of more benefit for diabetic complications[J].J Diabetes Investig,2019,10(5):1307-1317.
- [17]Tan NYQ,Chan J,Cheng CY,et al.Sleep Duration and Diabetic Kidney Disease [J].Front Endocrinol (Lausanne),2019(9):808.
- [18]王广利,章法香,黄欢,等.个体化预测2型糖尿病患者并发高尿酸血症列线图模型的建立及验证 [J].医学综述,2021,27(10):2040-2044,2049.
- [19]Zhang X,Zhao X,Huo L,et al.Risk prediction model of gestational diabetes mellitus based on nomogram in a Chinese population cohort study[J].Sci Rep,2020,10(1):21223.
- [20]Jiang S,Fang J,Yu T,et al.Novel Model Predicts Diabetic Nephropathy in Type 2 Diabetes [J].Am J Nephrol,2020,51(2):130-138.
- [21]Xi C,Wang C,Rong G,et al.A Nomogram Model that Predicts the Risk of Diabetic Nephropathy in Type 2 Diabetes Mellitus Patients: A Retrospective Study [J].Int J Endocrinol,2021(2021):6672444.
- [22]Shi R,Niu Z,Wu B,et al.Nomogram for the Risk of Diabetic Nephropathy or Diabetic Retinopathy Among Patients with Type 2 Diabetes Mellitus Based on Questionnaire and Biochemical Indicators: A Cross-Sectional Study[J].Diabetes Metab Syndr Obes,2020(13):1215-1229.
- [23]Wang G,Wang B,Qiao G,et al.Screening Tools Based on Nomogram for Diabetic Kidney Diseases in Chinese Type 2 Diabetes Mellitus Patients[J].Diabetes Metab J,2021,45(5):708-718.
- [24]肖薇.机器学习算法在糖尿病预测中的应用[J].数字技术与应用,2021,39(4):104-106.
- [25]乔高星.糖尿病肾病诊断预测模型构建[D].郑州:郑州大学,2019.
- [26]林鑫,李晋,刘蕾,等.二型糖尿病肾病风险预测模型比较[J].中华医学图书情报杂志,2019,28(4):41-45.
- [27]杨宏辉,朱利杰,高传玉.分类树模型与 logistic 回归在高血压危险因素预测中的应用分析[J].中国卫生标准管理,2017,8(24):7-10.
- [28]Lou J,Jing L,Yang H,et al.Risk factors for diabetic nephropathy complications in community patients with type 2 diabetes mellitus in Shanghai: Logistic regression and classification tree model analysis [J].Int J Health Plann Manage,2019,34(3):1013-1024.
- [29]廖志波,刘迅,成彩联,等.慢性肾脏病3-4期糖尿病肾脏疾病患者进展到终末期肾病的危险因素分类树模型分析 [J].中山大学学报(医学版),2019,40(5):767-773.
- [30]Golpour P,Ghayour-Mobarhan M,Saki A,et al.Comparison of Support Vector Machine, Nave Bayes and Logistic Regression for Assessing the Necessity for Coronary Angiography[J].International Journal of Environmental Research and Public Health,2020,17(18):6449.
- [31]刘小钰.基于 Meta-分析 II 型糖尿病并发症发病风险的 Logistic 回归模型研究[D].重庆:第三军医大学,2016.
- [32]李楠,唐杰,王一茹,等.基于三维超声构建糖尿病肾病诊断预测模型的初步研究[J].中华医学超声杂志(电子版),2019,16(9):702-708.
- [33]朱高培,孙娜,许小珊,等.2型糖尿病风险评估模型和风险评估分表的建立[J].中国卫生统计,2019,36(6):843-846,850.
- [34]Jiang W,Wang J,Shen X,et al.Establishment and Validation of a Risk Prediction Model for Early Diabetic Kidney Disease Based on a Systematic Review and Meta-Analysis of 20 Cohorts[J].Diabetes Care,2020,43(4):925-933.
- [35]Sullivan LM,Massaro JM,D'Agostino RB Sr.Presentation of multivariate data for clinical use: The Framingham Study risk score functions[J].Stat Med,2004,23(10):1631-1660.
- [36]Heikes KE,Eddy DM,Arondekar B,et al.Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes[J].Diabetes Care,2008,31(5):1040-1045.
- [37]Choi Y,Yang Y,Hwang BH,et al.Practical cardiovascular risk calculator for asymptomatic patients with type 2 diabetes mellitus: PRECISE-DM risk score[J].Clin Cardiol,2020,43(9):1040-1047.

收稿日期:2021-07-19;修回日期:2021-07-29

编辑/成森