

DOI: 10.3969/j.issn.1005-8982.2021.15.007
文章编号: 1005-8982 (2021) 15-0036-06

妇科疾病专题·论著

基于 Logistic 回归的外阴上皮内瘤变 危险度评价模型的构建

金钰妮, 李亚敏, 刘苑, 许静, 鲁潇凝

(昆明医科大学第一附属医院 妇科, 云南 昆明 650031)

摘要: **目的** 应用临床资料和流行病学资料建立外阴上皮内瘤变危险度评价模型。**方法** 模型样本为 101 例外阴上皮内瘤变患者、134 例外阴良性病变者及 150 例外阴完全正常者。对其临床资料和流行病学资料进行单因素分析, 通过受试者工作特征(ROC)曲线获取影响因素的截断值, 构建 Logistic 回归模型, 采用混淆矩阵回代分析验证模型的正确率。**结果** 经 Logistic 回归分析, 年龄、吸烟、人类乳头瘤病毒(HPV)感染、外阴瘙痒、外阴触血、外阴肿物、外阴疼痛、绝经与否、中性粒细胞绝对值(GRA)、中性粒细胞与淋巴细胞的比值(NLR)、血清糖类抗原 125(CA125) 指标进入回归方程。回归模型最大似然比(LR) 检验具有差异($P < 0.05$)。混淆矩阵回代分析结果发现该模型总体预测正确率为 78.44%(302/385)。**结论** 应用临床资料和流行病学资料构建外阴上皮内瘤变危险度评价模型, 能有效提高外阴上皮内瘤变正确筛查率。

关键词: 外阴上皮内瘤变; 早期筛查; Logistic 回归; 危险度评价

中图分类号: R737.35

文献标识码: A

Construction of risk assessment model for vulvar intraepithelial neoplasia based on logistic regression

Yu-ni Jin, Ya-min Li, Wan Liu, Jing Xu, Xiao-ning Lu

(Department of Gynecology, The First Affiliated Hospital of Kunming Medical University, Yunnan, Kunming 650031, China)

Abstract: Objective To risk assessment model of vulvar intraepithelial neoplasia established by clinical and epidemiological data. **Methods** The clinical and epidemiological data of 101 cases with intraepithelial neoplasia, 134 cases with benign lesions, and 150 cases with completely normal vulva were collected. The clinical data were analyzed by single factor analysis, the cut-off values of influencing factors were obtained by ROC curve, the logistic regression model was constructed, and the accuracy of the model was verified by using confusion matrix back analysis. **Results** Logistic regression analysis: Age, smoking, human papillomavirus (HPV) infection, pruritus of vulva, blood contact of vulva, tumor of vulva, pain of vulva, menopause or not, absolute value of neutrophils (GRA), ratio of neutrophils to lymphocytes (NLR), serum carbohydrate antigen 125 (CA125) index were concluded in the regression equation. The maximum likelihood ratio (LR) test of regression model showed significant difference ($P < 0.05$). The results of obfuscation matrix analysis showed that the overall prediction accuracy of the model is 78.44% (302/385). **Conclusions** The risk assessment model of vulvar intraepithelial neoplasia was established by using epidemiological and clinical data, and can effectively improve the correct screening rate of vulvar intraepithelial neoplasia.

Keywords: intraepithelial neoplasia, vulva; early detection of cancer; logistic regression; risk assessment

收稿日期: 2020-12-12

[通信作者] 李亚敏, E-mail: 394663203@qq.com; Tel: 15368048955

外阴上皮内瘤变是指女性外阴表皮内未向周围产生间质浸润和转移的癌前病变,全球每年新发病例数约 500 万例^[1]。不过外阴上皮内瘤变具有较长且可逆转的癌前病变期,大部分早期患者经治疗后可完全消退^[2-3],但未能在早期接受治疗的癌变率却高达 87.5%^[4]。可见,外阴上皮内瘤变的治疗及预后主要取决于早期诊断。临床普遍认为该病发生与年龄、吸烟、HPV 感染、疼痛、绝经等因素相关^[5]。外阴位置较特殊,女性对日常的一些外阴病变,如小结节、外阴瘙痒、外阴红斑、外阴触血、外阴疼痛等病理症状不够重视。这些症状恰好也是外阴良性疾病的一些病理症状,同时健康女性在日常生活中也常遇到一些外阴瘙痒、红斑的病理现象,无疑增加外阴上皮内瘤变早期筛查难度,导致部分早期外阴上皮内瘤变患者不能被确诊,错失癌前病变期的逆转机会。Logistic 回归分析能找出某种疾病的危险因素^[6]。危险因素在预测疾病发生概率具有较高敏感性和特异性,基于此,本研究以临床资料和流行病学资料构建外阴上皮内瘤变危险度评价模型,用于外阴上皮内瘤变的早期筛查,为外阴上皮内瘤变早期筛查提供一种新思路。

1 资料与方法

1.1 研究对象

建模样本选取 2011 年 1 月—2020 年 3 月在昆明医科大学第一附属医院确诊的外阴上皮内瘤变患者 101 例作为外阴上皮内瘤变组;同时选取外阴皮炎、外阴巴氏腺囊肿、外阴巴氏腺脓肿、外阴鲍文样丘疹病等具有自愈倾向的外阴良性疾病患者 134 例作为外阴良性疾病组;以同期健康体检外阴

完全正常女性 150 例作为正常对照组。

1.2 研究方法

收集外阴上皮内瘤变、外阴良性疾病及正常对照组的流行病学和临床资料,找出与外阴上皮内瘤变发生相关的各种因素,包括年龄、吸烟、妊娠次数、分娩次数、人类乳头瘤病毒(HPV)感染、外阴瘙痒、外阴触血、外阴肿块、外阴疼痛、绝经、白细胞(WBC)、单核细胞、嗜酸性粒细胞、嗜碱性粒细胞、中性粒细胞绝对值(GRA)、淋巴细胞绝对值(ALC)、中性粒细胞与淋巴细胞的比值(NLR)、血清糖类抗原 125(CA125),其中 $NLR = GRA/ALC$ 。

1.3 统计学方法

数据分析采用 SPSS 19.0 统计软件。计量资料以均数 \pm 标准差($\bar{x} \pm s$)表示,比较用方差分析;计数资料以例(%)表示,比较用 χ^2 检验;相关性因素采用单因素分析,并通过 ROC 曲线下面积(AUC)获取相关影响因素的最佳截断值;多因素采用 Logistic 回归模型,自变量回归系数采用 Wald χ^2 检验,模型的拟合情况采用似然比(LR)检验;以建模样本进行混淆矩阵回代分析,检验模型的正确率。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 外阴上皮内瘤变的单因素分析

单因素分析发现,年龄、吸烟、HPV 感染、外阴瘙痒、外阴触血、外阴肿块、外阴疼痛、绝经、GRA、NLR、CA125 与外阴上皮内瘤变发生有关($P < 0.05$)。见表 1。

表 1 外阴上皮内瘤变的单因素分析

组别	n	年龄 例(%)		妊娠 例(%)		分娩 例(%)		吸烟 例(%)	
		< 45 岁	≥ 45 岁	≥ 4 次	< 4 次	≥ 3 次	< 3 次	是	否
外阴上皮内瘤变组	101	33(32.67)	68(67.33)	42(41.58)	59(58.42)	39(38.61)	62(61.39)	25(24.75)	76(75.25)
外阴良性疾病组	134	43(32.09)	91(67.91)	48(35.82)	86(64.18)	46(34.33)	88(65.67)	21(15.67)	113(84.33)
正常对照组	150	68(45.33)	82(54.67)	58(38.67)	92(61.33)	57(38.00)	93(62.00)	16(10.67)	134(89.33)
χ^2/F 值		6.610		0.814		0.586		8.892	
P 值		0.037		0.666		0.746		0.012	

续表 1

组别	HPV 感染 例(%)		外阴瘙痒 例(%)		外阴触血 例(%)		外阴肿块 例(%)	
	阳性	阴性	有	无	有	无	有	无
外阴上皮内瘤变组	64(63.37)	37(36.63)	78(77.23)	23(22.77)	66(65.35)	35(34.65)	41(40.59)	60(59.41)
外阴良性疾病组	78(58.21)	56(41.79)	81(60.45)	53(39.55)	56(41.79)	78(58.21)	38(28.36)	96(71.64)
正常对照组	0(0.00)	120(100.00)	0(0.00)	150(100.00)	0(0.00)	150(100.00)	0(0.00)	150(100.00)
χ^2/F 值	144.262		179.580		128.758		68.731	
P 值	0.000		0.000		0.000		0.000	

组别	外阴疼痛 例(%)		绝经 例(%)		WBC/($\times 10^9/L$, $\bar{x} \pm s$)	单核细胞/($\times 10^9/L$, $\bar{x} \pm s$)
	有	无	已绝经	未绝经		
外阴上皮内瘤变组	59(58.42)	42(41.58)	68(67.33)	33(32.67)	6.43 \pm 1.29	0.58 \pm 0.16
外阴良性疾病组	52(38.81)	82(61.19)	65(48.51)	69(51.49)	6.42 \pm 1.31	0.61 \pm 0.18
正常对照组	0(0.00)	150(100.00)	68(45.33)	82(54.67)	6.75 \pm 1.56	0.62 \pm 0.11
χ^2/F 值	110.347		12.829		2.451	2.198
P 值	0.000		0.002		0.088	0.112

组别	嗜酸性粒细胞/ ($\times 10^9/L$, $\bar{x} \pm s$)	嗜碱性粒细胞/ ($\times 10^9/L$, $\bar{x} \pm s$)	ALC/($\times 10^9/L$, $\bar{x} \pm s$)	GRA/($\times 10^9/L$, $\bar{x} \pm s$)	NLR ($\bar{x} \pm s$)	CA125/(u/ml, $\bar{x} \pm s$)
外阴上皮内瘤变组	0.19 \pm 0.05	0.17 \pm 0.04	1.88 \pm 0.57	4.93 \pm 1.58	3.31 \pm 0.84	35.18 \pm 11.51
外阴良性疾病组	0.18 \pm 0.04	0.16 \pm 0.04	1.98 \pm 0.61	4.76 \pm 1.52	2.55 \pm 0.73	30.48 \pm 9.89
正常对照组	0.19 \pm 0.05	0.16 \pm 0.03	2.04 \pm 0.64	4.46 \pm 1.35	2.19 \pm 0.66	26.36 \pm 7.78
χ^2/F 值	1.765	2.807	2.066	3.321	70.750	25.580
P 值	0.173	0.062	0.128	0.037	0.000	0.000

2.2 GRA、NLR、CA125 预测外阴上皮内瘤变发生风险的 ROC 曲线分析

GRA、NLR、CA125 指标预测外阴上皮内瘤变发生的最佳截断值分别为 $4.92 \times 10^9/L$ 、 $2.57 \times 10^9/L$ 、 34.65 u/ml ，3 个指标均具有较高的敏感性和特异性。见表 2 和图 1。

2.3 多因素 Logistic 回归分析

以单因素分析具有统计学意义的(年龄、吸烟、HPV 感染、外阴瘙痒、外阴触血、外阴肿块、外阴疼

痛、绝经、GRA、NLR、CA125)，结合表 2 截断值，进行相关赋值(见表 3)。多因素 Logistic 回归分析中最大似然比(LR)向前逐步回归法分析，得到 Logistic 回归方程为： $\text{Logit}(p) = -7.589 + 1.497 \times \text{年龄} + 1.278 \times \text{吸烟} + 1.745 \times \text{HPV 感染} + 0.979 \times \text{外阴瘙痒} + 0.645 \times \text{外阴触血} + 1.285 \times \text{外阴肿块} + 1.094 \times \text{外阴疼痛} + 1.318 \times \text{绝经} + 1.647 \times \text{GRA} + 4.078 \times \text{NLR} + 3.093 \times \text{CA125}$ 。回归模型 LR 检验具有差异($\chi^2 = 128.79, P = 0.000$)。见表 4。

表 2 GRA、NLR、CA125 指标的 ROC 曲线分析结果

参数	截断值	AUC	95% CI		P 值	敏感性	95% CI		特异性	95% CI	
			下限	上限			下限	上限		下限	上限
GRA	$4.92 \times 10^9/L$	0.709	0.578	0.797	0.023	0.695	0.428	0.824	0.805	0.696	0.878
NLR	$2.57 \times 10^9/L$	0.907	0.694	0.978	0.026	0.753	0.686	0.875	0.863	0.742	0.914
CA125	34.65 u/ml	0.789	0.639	0.852	0.029	0.723	0.627	0.803	0.821	0.764	0.865

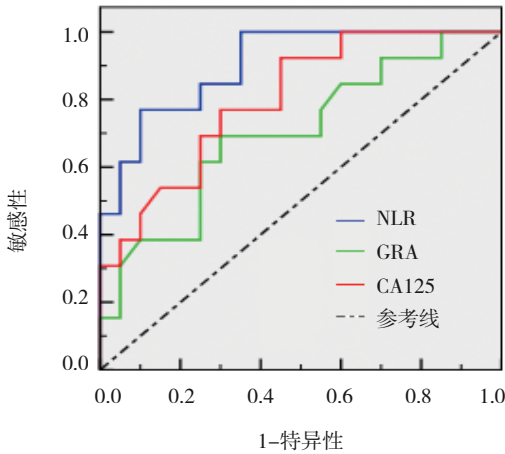


图 1 GRA、NLR、CA125 预测外阴上皮内瘤变的 ROC 曲线

2.4 模型样本预测的混淆矩阵回代检验

该模型预测外阴上皮内瘤变的正确率为 72.27%

表 3 相关指标的赋值说明

因素	变量名	赋值说明
年龄	X_1	0=<45 岁; 1= \geq 45 岁
吸烟	X_2	0=否; 1=是
HPV 感染	X_3	0=阴性; 1=阳性
外阴瘙痒	X_4	0=无; 1=有
外阴触血	X_5	0=无; 1=有
外阴肿块	X_6	0=无; 1=有
外阴疼痛	X_7	0=无; 1=有
绝经与否	X_8	0=未绝经; 1=绝经
GRA	X_9	0=< $4.92 \times 10^9/L$; 1= $\geq 4.92 \times 10^9/L$
NLR	X_{10}	0=< $2.98 \times 10^9/L$; 1= $\geq 2.98 \times 10^9/L$
CA125	X_{11}	0=< 34.65 u/ml; 1= ≥ 34.65 u/ml

(73/101); 外阴良性疾病的正确率为 69.40% (93/134); 正常对照组的正确率为 90.67% (136/150)。该模型总体预测正确率为 78.44% (302/385)。见表 5。

表 4 外阴上皮内瘤变的多因素 Logistic 回归分析参数

自变量	b	S_b	Wald χ^2	\hat{OR}	P 值	95% CI	
						上限	下限
常量	-7.589	0.819	85.862	0.001	0.000	-	-
年龄 ≥ 45 岁	1.497	0.573	6.825	4.468	0.011	2.046	9.368
吸烟	1.278	0.563	5.153	3.589	0.018	1.584	7.985
HPV 感染	1.745	0.476	13.436	5.726	0.000	2.595	15.278
外阴瘙痒	0.979	0.343	8.147	2.662	0.003	1.239	4.847
外阴触血	0.645	0.285	5.122	1.906	0.018	1.092	5.986
外阴肿块	1.285	0.457	7.906	3.615	0.005	2.174	9.132
外阴疼痛	1.094	0.493	4.924	2.986	0.019	1.145	12.715
已经绝经	1.318	0.765	2.968	3.736	0.048	1.705	9.983
GRA $\geq 4.92 (\times 10^9/L)$	1.647	0.886	3.456	5.191	0.044	2.234	12.894
NLR $\geq 2.98 (\times 10^9/L)$	4.078	0.542	56.610	59.027	0.000	9.424	79.692
CA125 ≥ 34.65 u/ml	3.093	0.725	18.201	22.043	0.000	6.745	38.719

表 5 模型样本预测的混淆矩阵 例 (%)

组别	n	预测值		
		外阴上皮内瘤变	外阴良性疾病	外阴完全正常
外阴上皮内瘤变组	101	73 (72.27)	22 (21.78)	6 (5.94)
外阴良性疾病组	134	29 (21.64)	93 (69.40)	12 (8.96)
正常对照组	150	1 (0.67)	13 (8.67)	136 (90.67)

3 讨论

近年来, 罹患外阴上皮内瘤变患者越来越年轻化^[7]。因外阴上皮内瘤变早期症状复杂多样, 部分症状与外阴良性疾病和女性日常中遇到的一些病理症状相同, 鉴别难度较大, 易产生误诊、漏诊, 造成部分外阴上皮内瘤变在早期不能被识别, 错失治疗时机, 随病情恶化, 最终发展成癌症^[8-10], 影响患者预后。因此, 研究外阴上皮内瘤变的早

期诊断方法,是外阴上皮内瘤变临床研究的重要方向。

某种疾病早期筛查准确率关键取决于高危人群筛选,对可能罹患某种疾病的概率进行风险评估,然后进行似然比检验,可在一定程度上提高疾病早期诊断的正确率。众所周知,在高危人群的筛选过程中,需根据现有的数据建立模型,才能预测目标对象走势和影响因素,而建立模型的好坏也将直接关系到预测精度^[11]。Logistic 回归分析可对多个自变量进行综合性分析,能较好地筛选出事件的危险因素,并预测事件的发生概率^[12]。目前 Logistic 回归在医学上用于分析某种疾病的治疗结果、诊断结果、影响因素、试验评价,均取得令人满意的效果。

基于 Logistic 回归模型优势,本研究从流行病学和临床资料中通过多因素 Logistic 回归分析筛选年龄 ≥ 45 岁、吸烟、HPV 感染、外阴瘙痒、外阴触血、外阴肿块、外阴疼痛、绝经、 $GRA \geq 4.92 \times 10^9/L$ 、 $NLR \geq 2.57 \times 10^9/L$ 、 $CA125 \geq 34.65$ u/mL 等有意义变量,且 $\beta > 0$,认为以上变量与外阴上皮内瘤变关系密切,这一发现与部分研究报道相同^[13-15]。其中 $NLR \geq 2.98 \times 10^9/L$ 的 \hat{OR} 值均高于其他变量,说明 NLR 在外阴上皮内瘤变的鉴别诊断中较其他变量更可靠,危险度更高。外阴上皮内瘤变原本属于自身免疫性系统疾病^[16-19],在机体炎症反应时,淋巴细胞减少,中性粒细胞呈上升趋势,NLR 作为中性粒细胞与淋巴细胞的比值,易于获取,可同时反映两者的变化,反映机体免疫状态。有研究表明,NLR 与外阴上皮内瘤变的病灶大小存在正相关性^[20],与本研究结果相符,说明 NLR 与外阴上皮内瘤变发生发展具有密切的关系,对外阴上皮内瘤变的早期鉴别具有较好的辅助作用。

回归模型最大似然比检验发现模型检验具有差异,说明从临床资料和流行病学资料中能更全面获取外阴上皮内瘤变发生的相关因素,使得模型检验具有差异性。在建模样本的混淆矩阵回代检验发现,模型的总体预测正确率高达 78.33%,说明该模型预测效果较优,对识别可能罹患外阴上皮内瘤变的高危人群具有较好效果。为可能罹患外阴上皮内瘤变人群的诊断提供参考依据。基于 Logistic 回归模型为临床医生早期筛查外阴上皮

内瘤变提供方便,也为临床医生快速诊断提供定量决策依据,且该项技术操作难度小,便捷,可重复性强,便于推广应用,也符合社会效益。

综上所述,通过流行病学和临床资料获取外阴上皮内瘤变相关影响因素,基于 Logistic 回归的外阴上皮内瘤变危险度评价模型能有效提高外阴上皮内瘤变早期诊断的正确率,为临床快速诊断提供一种新思路。但本研究样本量较少,模型预测效果有可能存在偏倚,为使该模型能取得更好的预测效果,后续将扩大样本量进行更全面的分析。

参 考 文 献 :

- [1] 薛雯娟,崔旭,杨宁,等.人乳头瘤病毒的 p53 蛋白表达与外阴上皮内瘤变的相关性[J].实用癌症杂志,2018,33(8):1267-1269.
- [2] RIVERO R C, GARCIA D, HAMMES L S, et al. Carcinogenesis of vulvar lesions: morphology and immunohistochemistry evaluation[J]. J Low Genit Tract Dis, 2017, 21(1): 73-77.
- [3] JIN C, LIANG S. Differentiated vulvar intraepithelial neoplasia: a brief review of clinicopathologic features[J]. Arch Pathol Lab Med, 2019, 143(6): 768-771.
- [4] 刘万清,赖曾珍,魏琴.中性粒细胞与淋巴细胞比值在外阴鳞癌与外阴鳞状上皮内瘤变术前鉴别诊断中的临床意义[J].河北医药,2017,39(23):3546-3549.
- [5] 谢玲玲,林荣春,林仲秋.«FIGO 2018 癌症报告»—外阴癌诊治指南解读[J].中国实用妇科与产科杂志,2019,35(6):660-665.
- [6] 廖孟霞,沈严严,刘灿,等.Logistic 回归模型在乳腺小结节超声鉴别诊断中的应用[J].中国现代医学杂志,2019,29(10):62-66.
- [7] AYAKANNU T, MURUGESU S, TAYLOR A H, et al. The impact of focality and centrality on vulvar Intraepithelial neoplasia on disease progression in HIV⁺ patients: a 10-year retrospective study[J]. Dermatology, 2019, 235(4): 327-333.
- [8] GARLAND S M, PAAVONEN J, JAISAMRARN U, et al. Prior human papillomavirus-16/18 AS04-adjuvanted vaccination prevents recurrent high grade cervical intraepithelial neoplasia after definitive surgical therapy: Post-hoc analysis from a randomized controlled trial[J]. Int J Cancer, 2016, 139(12): 2812-2826.
- [9] HORN L C, BRAMBS C E, HANDZEL R, et al. Nomenclature of squamous cell precursor lesions of the lower female genital tract: current aspects[J]. Pathologie, 2016, 37(6): 526-533.
- [10] BIGBY S M, EVA L J, FONG K L, et al. The natural history of vulvar intraepithelial neoplasia, differentiated type: evidence for progression and diagnostic challenges[J]. Int J Gynecol Pathol, 2016, 35(6): 574-584.
- [11] 林雨婷,庄虹莉,李立婷,等.基于 Logistic 回归组合预测的疾病诊断研究[J].中国卫生统计,2018,35(1):146-151.

- [12] 聂竞飞, 李玉玲. Logistic 模型及其在疾病诊断领域的应用[J]. 统计与管理, 2018(4): 72-75.
- [13] WAKEHAM K, KAVANAGH K, CUSCHIERI K, et al. HPV status and favourable outcome in vulvar squamous cancer[J]. Int J Cancer, 2017, 140(5): 1134-1146.
- [14] 肖晶晶, 隋龙. HPV 感染与肛门及外阴上皮内瘤变[J]. 实用妇产科杂志, 2017, 33(2): 84-86.
- [15] LI T, NIU X X, ZHANG X, et al. Baofukang suppository promotes the repair of vaginal epithelial cells in response to candida albicans[J]. AMB Express, 2016, 6(1): 109.
- [16] 冯宗昊, 刘朝晖, 张旭, 等. 外阴硬化性苔藓外阴上皮局部免疫状态初探[J]. 现代妇产科进展, 2019, 28(7): 509-513.
- [17] LI T, NIU X, ZHANG X, et al. Recombinant human IFN α -2b response promotes vaginal epithelial cells defense against[J]. Front Microbiol, 2017, 20(8): 697.
- [18] JIANG D, GE J, LIAO Q, et al. IgG and IgA with potential microbial-binding activity are expressed by normal human skin epidermal cells[J]. Int J Mol Sci, 2015, 16(2): 2574-2590.
- [19] 黄丹, 崔树娜, 李士华. 中性粒细胞与淋巴细胞比值影响妇科恶性肿瘤预后的研究进展[J]. 现代免疫学, 2019, 39(6): 500-504.
- [20] ABU-ZAID A, ALOMAR O, SALEM H. Preoperative elevated platelet count as a prognostic factor in vulvar squamous cell cancer: a mini-review[J]. Cureus, 2018, 10(3): e2279.

(张西倩 编辑)

本文引用格式: 金钰妮, 李亚敏, 刘菀, 等. 基于 Logistic 回归的外阴上皮内瘤变危险度评价模型的构建[J]. 中国现代医学杂志, 2021, 31(15): 36-41.

Cite this article as: JIN Y N, LI Y M, LIU W, et al. Construction of risk assessment model for vulvar intraepithelial neoplasia based on logistic regression[J]. China Journal of Modern Medicine, 2021, 31(15): 36-41.