

# 成都地区成年健康体检女性生殖道 HPV 感染情况调查及预测模型构建



唐逸娇<sup>1</sup>, 曾德玲<sup>2</sup>, 谭松<sup>3</sup>, 田莉<sup>1</sup>, 邓玉艳<sup>1</sup>

1. 成都市第三人民医院健康管理中心 (成都 610014)
2. 四川省肿瘤医院健康体检中心 (成都 610041)
3. 四川省妇幼保健院妇科 (成都 610031)

**【摘要】目的** 调查成都地区成年健康体检女性生殖道 HPV 感染情况, 并构建风险预测模型。**方法** 选择 2022 年 3 月至 2023 年 3 月在成都市三家医院进行健康体检的女性作为研究对象, 收集生殖道 HPV 感染的相关因素, 以 7:3 比例将研究对象分为训练集、验证集, 并根据有无 HPV 感染将研究对象分为感染组与非感染组。使用 LASSO 回归筛选潜在因素, 采用 Logistic 回归分析构建 HPV 感染风险预测模型并绘制列线图。采用受试者工作特征 (receiver operating characteristic, ROC) 曲线及其曲线下面积 (area under curve, AUC)、校准曲线和决策曲线评估风险预测模型的区分度、校准度和临床适用性。**结果** 共纳入 2 318 例健康体检女性, HPV 感染 481 例 (20.75%), 其中单一感染 316 例 (65.70%)、多重感染 165 例 (34.30%)。训练集 1 622 例, 验证集 696 例; 训练集中感染组 341 例, 非感染组 1 281 例。LASSO 回归及 Logistic 回归结果显示, 年龄 [OR=0.955, 95%CI (0.912, 0.999)]、产次 [OR=4.392, 95%CI (1.420, 13.583)]、初次性交年龄 [OR=0.870, 95%CI (0.774, 0.979)]、安全套使用情况 [OR=0.314, 95%CI (0.109, 0.905)]、性伴侣人数 [OR=6.068, 95%CI (1.825, 20.177)]、性伴侣包皮过长 [OR=3.218, 95%CI (1.042, 9.936)]、性传播疾病患病情况 [OR=3.476, 95%CI (1.071, 11.277)]、生殖道炎症 [OR=3.526, 95%CI (1.185, 10.494)]、宫颈柱状上皮异位情况 [OR=4.375, 95%CI (1.374, 13.934)] 为健康体检女性 HPV 感染的影响因素 ( $P < 0.05$ )。ROC 曲线结果显示, 训练集和验证集中预测模型的 AUC 分别为 0.913 [95%CI (0.866, 0.960)]、0.880 [95%CI (0.818, 0.941)]。训练集和验证集中预测模型的 H-L 拟合优度检验结果均无统计学意义 ( $P > 0.05$ )。校准曲线结果显示, 训练集和验证集中预测曲线与标准曲线基本拟合。决策曲线结果显示, 训练集和验证集中模型风险概率阈值分别为 0.05~0.90、0.05~0.78 时, 患者获益。**结论** 成都地区健康体检女性生殖道 HPV 感染率为 20.75%, 其主要受年龄、产次、初次性交年龄、安全套使用情况、性伴侣人数、性伴侣包皮过长、性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况的影响, 根据上述因素构建的女性生殖道 HPV 感染风险预测模型具有一定的临床使用价值。

**【关键词】** HPV; 感染风险; 预测模型; 健康体检

**【中图分类号】** R 737.33

**【文献标识码】** A

DOI: 10.12173/j.issn.1004-5511.202405134

基金项目: 成都市医学科研课题 (2022585)

通信作者: 曾德玲, 副主任医师, Email: enjoy\_51@sina.com

## Survey and construction of prediction model of HPV infection in the genital tract of women undergoing health checkups in Chengdu

TANG Yijiao<sup>1</sup>, ZENG Deling<sup>2</sup>, TAN Song<sup>3</sup>, TIAN Li<sup>1</sup>, DENG Yuyan<sup>1</sup>

1. Health Management Center, The Third People's Hospital of Chengdu, Chengdu 610014, China

2. Health Examination Center, Sichuan Cancer Hospital, Chengdu 610041, China

3. Department of gynaecology, Sichuan Provincial Maternity and Child Health Care Hospital, Chengdu 610031, China

Corresponding author: ZENG Deling, Email: enjoy\_51@sina.com

**【Abstract】Objective** To investigate and establish a prediction model of HPV infection in the genital tract of women undergoing health checkups in Chengdu. **Methods** Women who underwent occupational physical examination in three hospitals in Chengdu from March 2022 to March 2023 were selected as study subjects. The possible factors associated with genital tract HPV infection were collected. The study subjects were divided into the model group and validation group in a 7:3 ratio, and divided into an infected group and a non-infected group according to the presence or absence of HPV infection. LASSO regression was used to screen the potential factors of HPV infection. Logistic regression model was used to establish a risk prediction model for the risk of HPV infection and to draw a nomogram graph. The receiver operating characteristic (ROC) curve, area under curve (AUC), calibration curve and decision curve were used to assess the discrimination, calibration and clinical applicability of the risk prediction model. **Results** A total of 2,318 women undergoing health checkups were included and 481 (20.75%) were infected with HPV, of which 316 (65.70%) were single infections, 165 (34.30%) were multiple infections. 1,622 cases in the model group and 696 cases in the validation group. 341 cases were in the infected group and 1,281 cases were in the non-infected group in the model group. The LASSO and Logistic regression results showed that age [OR=0.955, 95%CI(0.912, 0.999)], number of births [OR=4.392, 95%CI(1.420, 13.583)], age at first sexual intercourse [OR=0.870, 95%CI(0.774, 0.979)], condom use [OR=0.314, 95%CI(0.109, 0.905)], number of sexual partners [OR=6.068, 95%CI(1.825, 20.177)], circumcision of sexual partners [OR=3.218, 95%CI(1.042, 9.936)], prevalence of sexually transmitted diseases [OR=3.476, 95%CI(1.071, 11.277)], inflammation of the genital tract [OR=3.526, 95%CI(1.185, 10.494)], and cervical columnar epithelial ectasia [OR=4.375, 95%CI(1.374, 13.934)] were the independent correlates of HPV infection in health checkup's women ( $P<0.05$ ). The results of ROC curve showed that the AUC of the predictive models in the model and validation groups were 0.913 [95%CI(0.866, 0.960)] and 0.880 [95%CI(0.818, 0.941)], respectively. The results of the H-L goodness-of-fit test for the column prediction models in both the model and validation groups were not statistically significant ( $P>0.05$ ). The results of calibration curves showed that the prediction curves in the model group and validation group basically fitted the standard curve. The decision curve showed that when the risk probability thresholds of the model and validation groups were 0.05~0.90 and 0.05~0.78, respectively, the patient benefit was greater than 0. **Conclusion** The prevalence of HPV infection in the health checkups of women in Chengdu was 20.75%, which was mainly affected by age, number of births, age at first sexual intercourse, condom use,

number of sexual partners, circumcision of sexual partners, prevalence of sexually transmitted diseases, inflammation of the genital tract, and cervical columnar epithelial ectasia. The risk prediction model for HPV infection in the female genital tract based on the above factors is valuable for clinical use.

**【Keywords】** HPV; Risk of infection; Prediction model; Health checkup

人乳头瘤病毒 (human papilloma virus, HPV) 为 DNA 病毒的一种, 目前已知的 HPV 病毒亚型超 100 种, 其中 30 多种被证实与宫颈疾病的发生高度相关<sup>[1-2]</sup>。HPV 感染可引起生殖器疣、寻常疣, 部分患者随着病情进展可发展为宫颈癌, 数据显示全球每年约有 63 万的新发宫颈癌是因 HPV 感染所致<sup>[3]</sup>。西方发达国家随着 HPV 疫苗的大面积接种宫颈癌的发病率有所下降, 国内因 HPV 疫苗接种率不足, 宫颈癌的发病率及死亡率仍较高<sup>[4]</sup>。因此预测女性生殖道 HPV 感染风险, 重点推荐风险较高人群接种 HPV 疫苗对于降低宫颈癌的发病率具有重要的意义。目前关于女性生殖道 HPV 感染的影响因素研究较多, 但少有研究将各因素整合成预测模型, 虽然王卉等<sup>[5]</sup>建立了女性 HPV 感染预测模型, 但该模型使用 Logistic 回归模型, 模型计算复杂、临床普及难度大, 且该研究筛选潜在因素时仅采用单因素分析, 可能存在多重共线性问题。鉴于此, 本研究通过 LASSO 回归进行潜在预测因素的筛选并构建列线图用于 HPV 感染风险预测, 以期对 HPV 感染风险的预测和预防提供参考。

## 1 资料与方法

### 1.1 研究对象

选择 2022 年 3 月至 2023 年 3 月在成都市第三人民医院、四川省肿瘤医院、四川省妇幼保健院三家医院进行健康体检的女性作为研究对象。纳入标准: ①在三家医院行年度健康体检; ②体检项目包括 HPV 检测; ③智力正常, 读写能力正常, 可配合进行相关调查; ④研究对象知情同意。排除标准: ①有子宫切除或宫颈手术史; ②调查前 72 h 曾使用阴道栓剂或冲洗阴道者; ③有骨盆放射治疗史; ④妊娠期女性; ⑤生殖道畸形者; ⑥宫颈癌患者。本研究获得成都市第三人民医院伦理委员会审批 (批号: 成都三院伦 2022—S-44), 其他 2 家医院认可该伦理审批。

### 1.2 资料收集

收集可能影响健康体检女性生殖道 HPV 感染的相关因素, 包括: ①一般资料: 年龄、BMI 指数、职业、吸烟史 (吸烟指吸食过 20 支及以上香烟或相当的烟丝)、饮酒史 (饮酒指每周饮酒  $\geq 1$  次, 每次  $\geq 150$  mL)、婚姻状况、受教育程度、宫颈癌家族史; ②生殖相关资料: 孕次、产次、初次性交年龄、性交频率、安全套使用情况、输卵管结扎情况、性事后外阴清洗情况; ③性伴侣相关情况: 性伴侣人数、性伴侣性病史、性伴侣生殖系统恶性肿瘤病史、性伴侣是否包皮过长; ④相关疾病情况: 性传播疾病病史、生殖道炎症病史、宫颈柱状上皮异位病史情况。上述指标均采用现场问卷调查的方式在研究对象接受 HPV 检测后的 30 min 内完成填写。

HPV 取样检查与分组方法如下: 嘱研究对象取样前 3 d 不进行阴道上药及冲洗, 24 h 内避免性交, 并避开经期。取样时窥阴器暴露宫颈后, 用 HPV 取样刷在宫颈口顺时针转 5 圈进行宫颈脱落细胞取样, 取样后置于带编号的标本管于 4 °C 冰箱保存待检。使用潮州凯普生物化学有限公司生产的 HPV 亚型检测试剂盒 (PCR+ 膜杂交法) 对研究对象宫颈脱落细胞样本进行 HPV 分型检测。本研究使用的试剂盒可检测出 21 种分型, 出现任意一种及以上亚型阳性则判定为 HPV 感染, 根据有无 HPV 感染将研究对象分为感染组与非感染组。

### 1.3 统计学分析

采用 SPSS 22.0 软件进行数据分析, 计数资料以例数和百分比 ( $n, \%$ ) 表示, 组间比较采用  $\chi^2$  检验, 计量资料以均数和标准差 ( $\bar{x} \pm s$ ) 表示, 组间比较采用  $t$  检验。采用 R 4.1.3 软件和 “caret” 程序包以 7 : 3 比例将研究对象随机分为训练集、验证集。采用 “glmnet” 程序包进行 LASSO 回归分析筛选变量, 在此基础上采用多因素 Logistic 回归模型分析健康体检女性生殖道 HPV 感染的相关因素, 使用 “rms” 程序包构建列线图。采用受

试者工作特征 (receiver operating characteristic, ROC) 曲线及其曲线下面积 (area under curve, AUC) 评价模型的区分度, 采用 Bootstrap 法重复抽样 1 000 次进行内部验证, 采用 Hosmer-Lemeshow 拟合优度检验、校准曲线评价模型校准度, 采用决策曲线评价模型的临床适用性。以  $P < 0.05$  为差异有统计学意义。

## 2 结果

### 2.1 一般资料

研究期间共 2 711 名健康体检女性进行体

检, 经纳入排除标准筛选后最终纳入 2 318 名健康体检女性, 其中训练集 1 622 例, 验证集 696 例。训练集与验证集各影响因素比较差异均无统计学意义 ( $P > 0.05$ )。训练集中, HPV 感染组 341 例, 非感染组 1 281 例, 两组年龄、职业、吸烟史、宫颈癌家族史、孕次、产次、初次性交年龄、性交频率、安全套使用情况、性伴侣人数、性伴侣性病史、性伴侣包皮过长、性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况的差异具有统计学意义 ( $P < 0.05$ ), 见表 1。

表1 不同组别研究对象基本特征比较 (n, %)

Table 1. Comparison of basic characteristics of research subjects in different groups (n, %)

特征	训练集 vs. 验证集				训练集中感染组 vs. 非感染组			
	训练集 (n=1 622)	验证集 (n=696)	$\chi^2/t$ 值	P值	感染组 (n=341)	非感染组 (n=1 281)	$\chi^2/t$ 值	P值
年龄 (岁)*	44.64 ± 5.03	44.28 ± 5.49	1.536	0.125	39.71 ± 4.88	46.03 ± 5.92	-18.140	<0.001
BMI指数 (kg/m <sup>2</sup> )*	20.30 ± 4.17	20.42 ± 4.33	-0.628	0.530	20.17 ± 4.23	20.33 ± 4.19	-0.625	0.532
职业			7.171	0.067			13.531	0.004
公职人员	421 (25.96)	199 (28.60)			74 (21.70)	347 (27.09)		
私企人员	689 (42.48)	314 (45.10)			173 (50.73)	516 (40.28)		
自由职业者	303 (18.68)	102 (14.66)			61 (17.89)	242 (18.89)		
家庭主妇	209 (12.89)	81 (11.64)			33 (9.68)	176 (13.84)		
吸烟史			0.765	0.382			5.551	0.018
无	1 212 (74.72)	508 (72.99)			238 (69.79)	974 (76.03)		
有	410 (25.28)	188 (27.01)			103 (30.21)	307 (23.97)		
饮酒史			1.658	0.198			1.767	0.184
无	1 347 (83.05)	593 (85.20)			275 (80.65)	1 072 (83.68)		
有	275 (16.95)	103 (14.80)			66 (19.35)	209 (16.32)		
婚姻状况			3.561	0.059			0.804	0.370
非在婚	367 (22.63)	133 (19.11)			71 (20.82)	296 (23.11)		
在婚	1 255 (77.37)	563 (80.89)			270 (79.18)	985 (76.89)		
受教育程度			0.033	0.865			0.623	0.430
初中及以下	226 (13.93)	95 (13.65)			52 (15.25)	174 (13.58)		
高中/中专及以上	1 396 (86.07)	601 (86.35)			289 (84.75)	1 107 (86.42)		
宫颈癌家族史			0.033	0.865			5.750	0.016
无	1 447 (89.21)	615 (88.36)			292 (85.63)	1 155 (90.16)		
有	175 (10.79)	81 (11.64)			49 (14.37)	126 (9.84)		
孕次 (次)			0.966	0.326			7.026	0.008
≤1	995 (61.34)	442 (63.51)			188 (55.13)	807 (63.00)		
>1	627 (38.66)	254 (36.49)			153 (44.87)	474 (37.00)		

续表1

特征	训练集 vs. 验证集				训练集中感染组 vs. 非感染组			
	训练集 (n=1 622)	验证集 (n=696)	$\chi^2/t$ 值	P值	感染组 (n=341)	非感染组 (n=1 281)	$\chi^2/t$ 值	P值
产次 (次)			0.059	0.808			36.678	<0.001
≤1	942 (58.08)	408 (58.62)			149 (43.70)	793 (61.90)		
>1	680 (41.92)	288 (41.38)			192 (56.30)	488 (38.10)		
初次性交年龄 (岁)*	23.58 ± 4.71	23.82 ± 4.88	-1.112	0.266	20.97 ± 4.60	24.28 ± 4.91	-11.208	<0.001
性交频率 (次/月)			0.650	0.420			6.926	0.008
≤5	1 086 (66.95)	454 (65.23)			208 (61.00)	878 (68.54)		
>5	536 (33.05)	242 (34.77)			133 (39.00)	403 (31.46)		
使用安全套			0.015	0.902			107.189	<0.001
否	562 (34.65)	243 (34.91)			199 (58.36)	363 (28.34)		
是	1 060 (65.35)	453 (65.09)			142 (41.64)	918 (71.66)		
性伴侣人数 (人)			0.030	0.862			10.567	0.001
1	1 085 (66.89)	463 (66.52)			203 (59.53)	882 (68.85)		
>1	537 (33.11)	233 (33.48)			138 (40.47)	399 (31.15)		
性伴侣性病史			0.688	0.407			5.463	0.019
无	1 278 (78.79)	559 (80.32)			253 (74.19)	1 025 (80.02)		
有	344 (21.21)	137 (19.68)			88 (25.81)	256 (19.98)		
性伴侣生殖系统恶性肿瘤病史			3.203	0.074			1.225	0.268
无	1 595 (98.34)	691 (99.28)			333 (97.65)	1 262 (98.52)		
有	27 (1.66)	5 (0.72)			8 (2.35)	19 (1.48)		
性伴侣包皮过长			0.972	0.324			7.099	0.008
无	984 (60.67)	407 (58.48)			184 (53.96)	793 (61.90)		
有	638 (39.33)	289 (41.52)			157 (46.04)	488 (38.10)		
性传播疾病患病史			3.592	0.058			61.060	<0.001
无	1 292 (79.65)	578 (83.05)			220 (64.52)	1 072 (83.68)		
有	330 (20.35)	118 (16.95)			121 (35.48)	209 (16.32)		
生殖道炎症史			0.461	0.497			83.483	<0.001
无	1 020 (62.89)	448 (64.37)			142 (41.64)	878 (68.54)		
有	602 (37.11)	248 (35.63)			199 (58.36)	403 (31.46)		
宫颈柱状上皮异位情况			0.545	0.460			65.588	<0.001
无	1 124 (69.30)	493 (70.83)			175 (51.32)	949 (74.08)		
有	498 (30.70)	203 (29.17)			166 (48.68)	332 (25.92)		

注: \*计量资料以均数和标准差 ( $\bar{x} \pm s$ ) 表示, 组间比较采用t检验。

## 2.2 HPV病毒分型情况

纳入的 2 318 名健康体检女性中, HPV 感染 481 例 (20.75%), 其中单一感染 316 例 (65.70%), 多重感染 165 例 (34.30%)。训练集 1 622 名健

康体检女性中, HPV 感染 341 例 (21.02%), 其中单一感染 227 例 (66.57%), 多重感染 114 例 (33.43%)。验证集 696 例健康体检女性中, HPV 感染 140 例 (20.11%), 单一感染 89 例

(63.57%)，多重感染 51 例 (35.43%)，其具体分型见表 2。

### 2.3 健康体检女性 HPV 感染潜在因素筛选

以训练集健康体检女性是否感染 HPV 为因变量，将表 1 中各因素作为自变量行 LASSO 回归分析，结果显示，lambda.1se 对应的变量数为 9，因此最终筛选出 9 个潜在变量，分别为年龄、产次、初次性交年龄、安全套使用情况、性伴侣人数、性伴侣包皮过长、性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况，见图 1。

### 2.4 健康体检女性 HPV 感染相关因素 Logistic 回归分析

训练集中，以 LASSO 回归筛选的变量为自变量，以健康体检女性是否感染 HPV 为因变量（感染 =1，非感染 =0）行多因素 Logistic 回归分析，结果显示，年龄、产次、初次性交年龄、安全套使用情况、性伴侣人数、性伴侣包皮过长、

性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况为女性 HPV 感染的独立性影响因素 ( $P < 0.05$ )，见表 3。

### 2.5 健康体检女性 HPV 感染风险列线图模型验证

根据 Logistic 回归分析结果绘制列线图，见图 2。ROC 分析结果显示，训练集中模型的 AUC 为 0.913[95%CI (0.866, 0.960)]，验证集中 AUC 为 0.880[95%CI (0.818, 0.941)]，见图 3。Hosmer-Lemeshow 拟合优度检验结果显示模型与理想模型差异无统计学意义（训练集： $\chi^2=0.319$ ， $P=0.572$ ；验证集： $\chi^2=1.016$ ， $P=0.314$ ）。校准曲线结果显示，训练集与验证集预测曲线与标准曲线基本拟合，见图 4。临床决策曲线显示，当列线图模型训练集风险概率阈值为 0.05~0.90，验证集概率阈值为 0.05~0.78 时，患者获得净收益，见图 5。

表2 健康体检女性的HPV病毒分型情况 (n, %)

Table 2. Types of HPV virus in women undergoing health checkup (n, %)

HPV病毒分型	合计 (n=2 318)	训练集 (n=1 622)	验证集 (n=696)
HPV6	101 (4.36)	71 (4.38)	30 (4.31)
HPV11	90 (3.88)	62 (3.82)	28 (4.02)
HPV40	13 (0.56)	9 (0.55)	4 (0.57)
HPV43	10 (0.43)	5 (0.31)	5 (0.72)
HPV44	38 (1.64)	24 (1.48)	14 (2.01)
HPV81	231 (9.97)	166 (10.23)	65 (9.34)
HPV16	84 (3.62)	61 (3.76)	23 (3.30)
HPV18	62 (2.67)	43 (2.65)	19 (2.73)
HPV31	24 (1.04)	15 (0.92)	9 (1.29)
HPV33	36 (1.55)	24 (1.48)	12 (1.72)
HPV35	12 (0.52)	6 (0.37)	6 (0.86)
HPV39	37 (1.60)	25 (1.54)	12 (1.72)
HPV45	5 (0.22)	3 (0.18)	2 (0.29)
HPV51	33 (1.42)	22 (1.36)	11 (1.58)
HPV52	69 (2.98)	48 (2.96)	21 (3.02)
HPV56	15 (0.65)	9 (0.55)	6 (0.86)
HPV59	12 (0.52)	6 (0.37)	6 (0.86)
HPV58	60 (2.59)	40 (2.47)	20 (2.87)
HPV66	21 (0.91)	13 (0.80)	8 (1.15)
HPV53	39 (1.68)	27 (1.66)	12 (1.72)

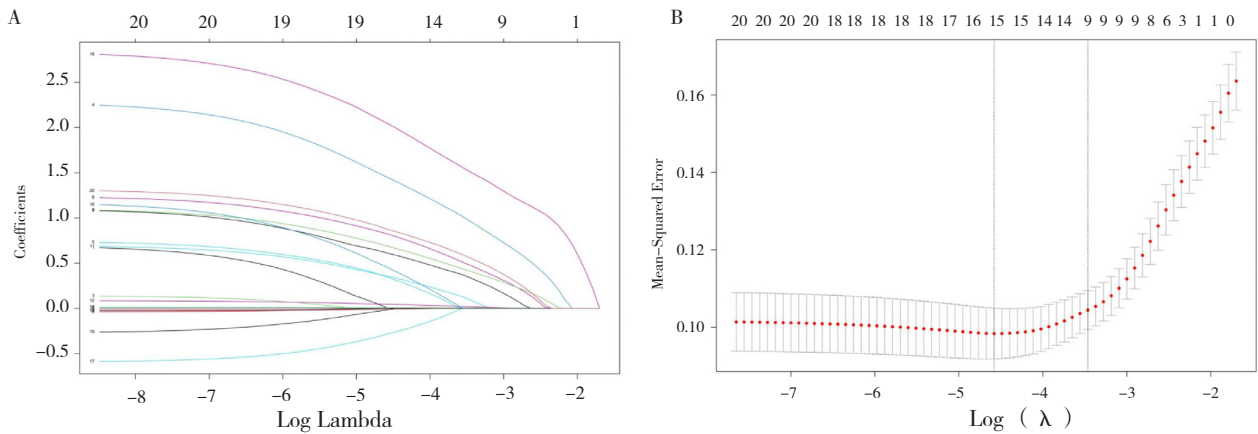


图1 LASSO回归结果图

Figure 1. LASSO regression results

注：A. LASSO回归的系数路径；B. LASSO回归交叉验证结果。

表3 健康体检女性HPV感染相关因素Logistic回归分析结果

Table 3. Logistic regression analysis results of factors related to HPV infection in health checkups women

变量	赋值方式	B值	SE值	Wald $\chi^2$ 值	OR值 (95%CI)	P值
年龄 (岁)	原值代入	-0.046	0.023	3.942	0.955 (0.912, 0.999)	0.047
产次 (次)	$\leq 1=0, > 1=1$	1.480	0.576	6.599	4.392 (1.420, 13.583)	0.010
初次性交年龄 (岁)	原值代入	-0.139	0.060	5.340	0.870 (0.774, 0.979)	0.021
安全套使用情况	否=0, 是=1	-1.159	0.540	4.604	0.314 (0.109, 0.905)	0.032
性伴侣人数	$1=0, > 1=1$	1.803	0.613	8.651	6.068 (1.825, 20.177)	0.003
性伴侣包皮过长	无=0, 有=1	1.169	0.575	4.127	3.218 (1.042, 9.936)	0.042
性传播疾病患病情况	无=0, 有=1	1.246	0.600	4.305	3.476 (1.071, 11.277)	0.038
生殖道炎症	无=0, 有=1	1.260	0.556	5.131	3.526 (1.185, 10.494)	0.024
宫颈柱状上皮异位情况	无=0, 有=1	1.476	0.591	6.237	4.375 (1.374, 13.934)	0.013
常量	-	2.911	1.722	2.858	-	0.091

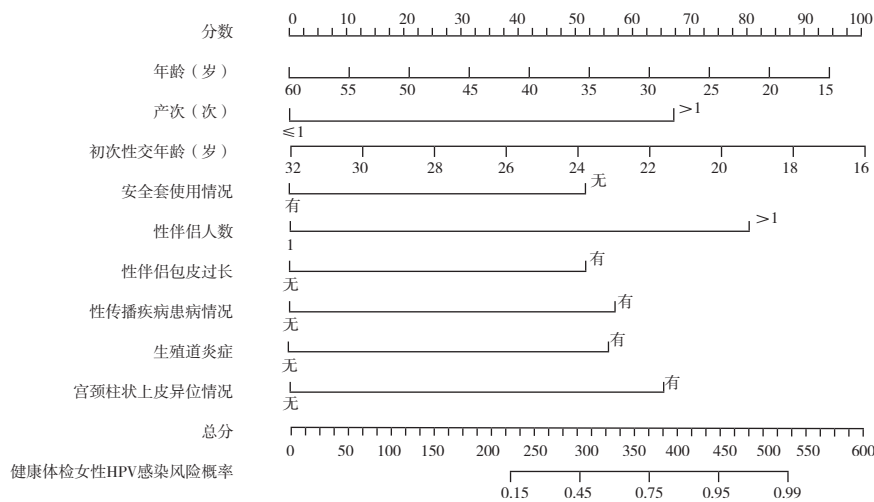


图2 健康体检女性HPV感染风险预测模型列线图

Figure 2. Nomogram of HPV infection risk prediction model for women undergoing health checkups

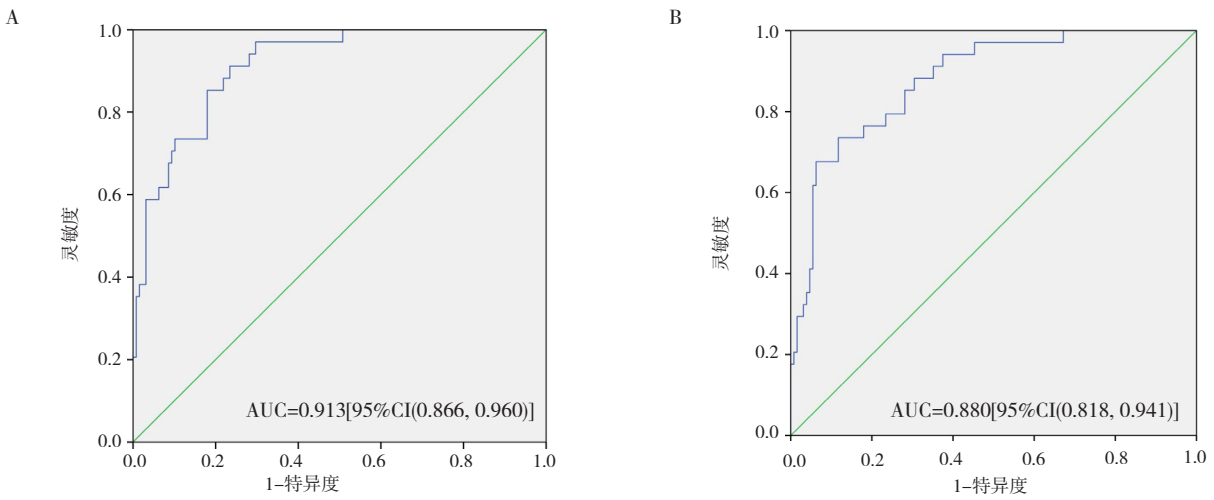


图3 健康体检女性HPV感染风险预测模型ROC分析

Figure 3. ROC analysis of HPV infection risk prediction model for women undergoing health checkups

注：A. 训练集；B. 验证集。

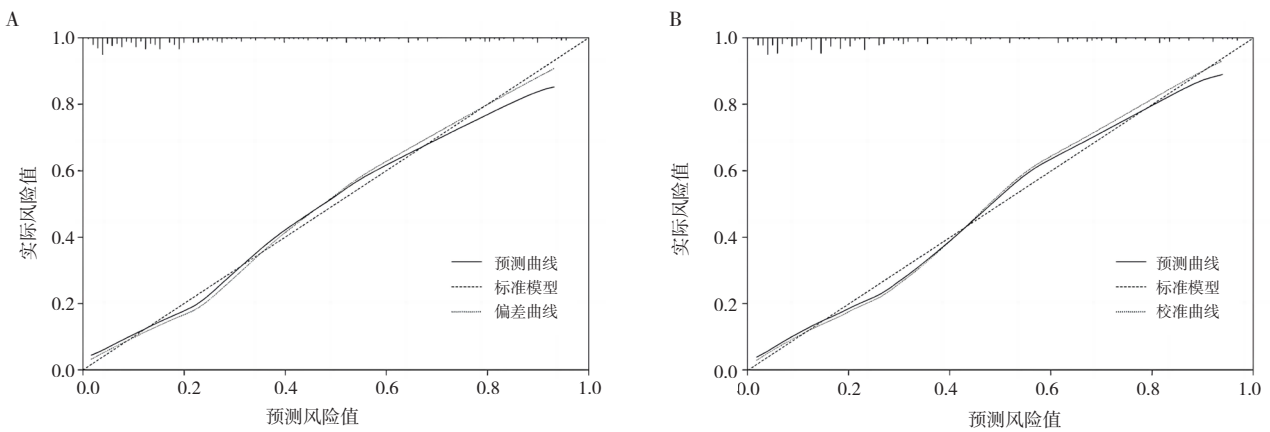


图4 健康体检女性HPV感染风险预测模型校准曲线

Figure 4. Calibration curve of HPV infection risk prediction model for women undergoing health checkups

注：A. 训练集；B. 验证集。

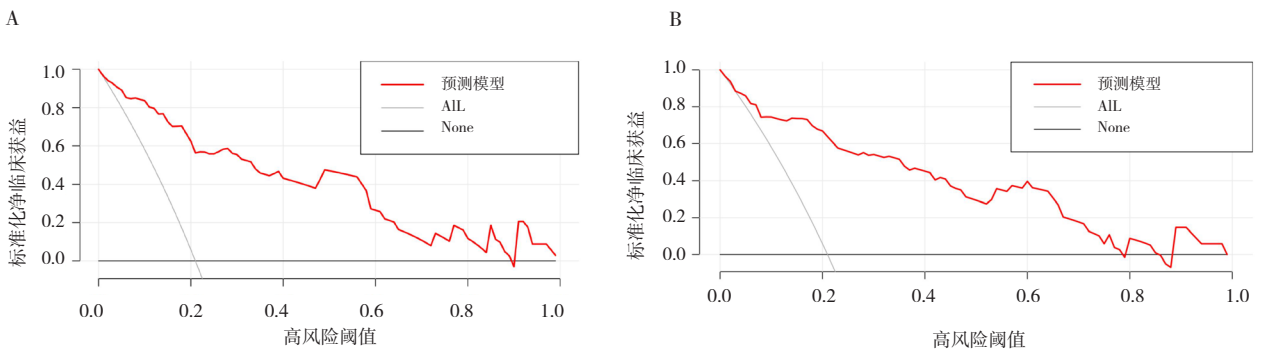


图5 健康体检女性HPV感染风险预测模型决策曲线

Figure 5. Decision curve of HPV infection risk prediction model for women undergoing health checkup

注：A. 训练集；B. 验证集。



### 3 讨论

宫颈癌为女性常见恶性肿瘤类型, 该病也是目前唯一病因明确、三级预防手段完善的恶性肿瘤<sup>[6-8]</sup>。HPV 感染为宫颈癌的主要致病因素, 宫颈癌患者中高危 HPV 感染率几乎可达 100%<sup>[9]</sup>。既往研究显示, 不同地区 HPV 患病率差异较大, 我国地域辽阔, 不同地区 HPV 流行情况也不尽相同, 此前报道的国内 HPV 感染率在 9.9%~31.9% 之间, 本研究中成都地区健康体检女性生殖道 HPV 感染率为 20.75%, 处于中高水平<sup>[10-12]</sup>。

HPV 感染是多方面因素共同作用的结果, 将各影响因素进行整合建立预测模型可实现患者的个性化预测, 既往研究多在单因素分析基础上行多因素回归, 可能将如孕次、产次等存在多重共线性的因素纳入模型, 逐步 Logistic 回归虽通过分析自变量与因变量之间的关系, 逐步引入和剔除变量, 以达到最优的模型拟合效果, 但这种方法基于信息准则 (如 AIC、BIC 等) 来决定何时停止变量的引入或剔除。相比之下, LASSO 回归通过在损失函数中添加一个 L1 正则化项来实现变量的选择, 正则化项会鼓励模型的系数趋向于零, 从而达到特征选择的目的, 这种方法不仅可以帮助构建一个更简洁的模型, 还可以提高模型的预测精度和稳定性<sup>[13-14]</sup>。此外, 本研究在多因素分析基础上建立了列线图, 使用方便并易于解释, 可用于患者个性化预测并将结果向患者展示, 方便患者了解自身病情积极配合临床干预。

本研究结果显示, 年龄、产次、初次性交年龄、安全套使用情况、性伴侣人数、性伴侣包皮过长、性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况为健康体检女性 HPV 感染的独立影响因素。年龄对于女性生殖道 HPV 感染的影响已被多项研究所证实, 考虑主要和年龄较低的人群为性活跃人群, 更容易出现多个性伴侣, 感染性传播疾病的风险更高有关<sup>[15]</sup>。产次对于女性 HPV 的影响可能与以下几方面原因有关, 首先, 多次分娩会对女性的子宫颈造成损伤, 这种损伤为 HPV 提供了更容易进入体内的途径; 同时, 分娩过程中, 女性的免疫系统可能会受到一定的抑制, 导致其暂时性地免疫力下降, 这种免疫力下降状态可能持续一段时间, 使得产妇在这段时间内更容易感染 HPV; 再者, 多次分娩意味着女性

有更多的机会接触到 HPV, 如果产妇的性伴侣是 HPV 的携带者, 那么感染 HPV 风险可进一步增加<sup>[16]</sup>。初次性交年龄较小的女性宫颈上皮细胞尚未发育完全, 在性交时易受刺激而出现损伤, 且此阶段性激素水平分泌明显增加, 一旦出现自身免疫功能下降可影响机体对病毒的清除能力, 增加 HPV 感染风险<sup>[17]</sup>。安全套的使用已被证实可用于阻断大部分性传播疾病, 性交过程全程并正确使用男用安全套可通过物理隔绝的方式阻断生殖道分泌液的交换, 进而降低 HPV 感染的风险。由于 HPV 主要通过性接触传播, 因此性伴侣人数的增加直接导致了 HPV 感染风险的提高; 另外, 频繁更换性伴侣可能导致免疫系统不断面对新的病毒挑战, 从而增加免疫系统的压力, 在免疫系统较为脆弱或功能低下的情况下, HPV 更容易侵入并引发感染; 同时, 有多个性伴侣的人可能更容易忽视个人卫生和性健康的重要性, 如不进行充分的清洁、不使用保护措施等, 这些行为都可能增加 HPV 感染的风险<sup>[18]</sup>。过长的包皮内容易存留包皮垢及分泌物, 清除难度较大, 加上包皮部位局部温热、潮湿环境容易引起致病菌滋生, 也为 HPV 的繁殖提供了温床, 因此包皮过长可导致女性生殖道 HPV 感染风险增加<sup>[19]</sup>。性传播疾病如淋病、梅毒、衣原体感染等会导致阴道小环境的免疫屏障降低, 这种免疫力的下降使得阴道上皮细胞更容易受到 HPV 病毒的侵袭<sup>[20]</sup>。另外, 研究发现, 与无性病者相比, 有性传播疾病的患者在进行性行为时, 更容易接触到 HPV 病毒<sup>[21]</sup>。

除上述因素外, 生殖道炎症及宫颈柱状上皮异位均可增加健康体检女性 HPV 感染风险, 生殖道炎症可使阴道局部环境出现紊乱、自清洁功能下降而致 HPV 感染风险上升<sup>[22]</sup>。另外, 长期生殖道慢性炎症还可使机体细胞免疫功能下降, 阴道内乳酸杆菌减少、pH 值改变而致 HPV 感染后难以被清除, 出现持续感染而增加宫颈癌发病风险<sup>[23]</sup>。宫颈柱状上皮异位主要受体内雌激素水平的影响, 在镜下可见柱状上皮及不成熟鳞状上皮化生, 鳞状上皮化生在 HPV 的自然增殖及分化过程中均起着关键作用, 因此宫颈柱状上皮异位者 HPV 感染风险更高<sup>[24]</sup>。另外, 部分患者在 HPV 感染宫颈上皮细胞后可以通过不同途径引起细胞变化, 导致宫颈上皮异位, 但需要注意的是, 宫颈上皮异位不仅仅由 HPV 感染引起, 还可能由其

他因素如激素水平变化、炎症刺激等导致,因此,仅凭单一指标预测 HPV 感染风险准确性较低。

综上所述,成都地区健康体检女性生殖道 HPV 感染率为 20.75%,其主要受年龄、产次、初次性交年龄、安全套使用情况、性伴侣人数、性伴侣包皮过长、性传播疾病患病情况、生殖道炎症、宫颈柱状上皮异位情况的影响,根据上述因素建立的列线图模型可用于预测女性生殖道 HPV 感染,具有一定的临床应用价值。但因本研究目前聚焦于成都地区的健康体检女性群体,其模型预测效果在其他地区的适用性尚需通过外部数据的验证来明确。为了增强研究结论的普适性和可信度,本研究拟在未来启动开展多中心研究,通过拓宽样本的地域分布和增加样本量,对模型进行更深入的优化与验证。

## 参考文献

- 1 Selinger C, Rahmoun M, Murall CL, et al. Cytokine response following perturbation of the cervicovaginal milieu during HPV genital infection[J]. *Immunol Res*, 2021, 69(3): 255–263. DOI: [10.1007/s12026-021-09196-2](https://doi.org/10.1007/s12026-021-09196-2).
- 2 Lu X, Ji M, Wagner AL, et al. Willingness to pay for HPV vaccine among female health care workers in a Chinese nationwide survey[J]. *BMC Health Serv Res*, 2022, 22(1): 1324. DOI: [10.1186/s12913-022-08716-6](https://doi.org/10.1186/s12913-022-08716-6).
- 3 Kops NL, Caierão J, Bessel M, et al. Behavioral factors associated with multiple-type HPV genital infections: data from a cross-sectional study in young women in Brazil[J]. *Reprod Health*, 2021, 18(1): 201. DOI: [10.1186/s12978-021-01244-2](https://doi.org/10.1186/s12978-021-01244-2).
- 4 Costa AS, Gomes JM, Germani ACCG, et al. Knowledge gaps and acquisition about HPV and its vaccine among Brazilian medical students[J]. *PloS One*, 2020, 15(3): e0230058. DOI: [10.1371/journal.pone.0230058](https://doi.org/10.1371/journal.pone.0230058).
- 5 王卉,雷英,闫晶晶,等.重庆市万州区已婚女性 HPV 感染危险因素和预测模型构建及验证[J]. *中华医院感染学杂志*, 2023, 33(2): 238–242. [Wang H, Lei Y, Yan JJ, et al. Risk factors for HPV infection in married women of Wanzhou District, Chongqing, establishment of prediction model and its validation[J]. *Chinese Journal of Nosocomiology*, 2023, 33(2): 238–242.] DOI: [10.11816/cn.ni.2023-213055](https://doi.org/10.11816/cn.ni.2023-213055).
- 6 Niu J, Pan S, Wei Y, et al. Epidemiology and analysis of potential risk factors of high-risk human papillomavirus (HPV) in Shanghai China: a cross-sectional one-year study in non-vaccinated women[J]. *J Med Virol*, 2022, 94(2): 761–770. DOI: [10.1002/jmv.27453](https://doi.org/10.1002/jmv.27453).
- 7 Helmkamp LJ, Szilagyi PG, Zimet G, et al. A validated modification of the vaccine hesitancy scale for childhood, influenza and HPV vaccines[J]. *Vaccine*, 2021, 39(13): 1831–1839. DOI: [10.1016/j.vaccine.2021.02.039](https://doi.org/10.1016/j.vaccine.2021.02.039).
- 8 Harrison SE, Yelverton V, Wang Y, et al. Examining associations between knowledge and vaccine uptake using the Human Papillomavirus Knowledge Questionnaire (HPV-KQ)[J]. *Am J Health Behav*, 2021, 45(5): 810–827. DOI: [10.5993/AJHB.45.5.2](https://doi.org/10.5993/AJHB.45.5.2).
- 9 Clavé Llavall A, de Wildt G, Meza G, et al. Nurses' and teachers' perceived barriers and facilitators to the uptake of the Human Papilloma Virus (HPV) vaccination program in Iquitos, Peru: a qualitative study[J]. *PloS One*, 2021, 16(7): e0255218. DOI: [10.1371/journal.pone.0255218](https://doi.org/10.1371/journal.pone.0255218).
- 10 张梦培,尹如铁,李克敏,等.中国西部女性生殖道高危人乳头状病毒感染状况的 Meta 分析[J]. *中国循证医学杂志*, 2019, 19(11): 1268–1275. [Zhang MP, Yin RT, Li KM, et al. The prevalence of high risk human papillomavirus of females in western China: a Meta-analysis[J]. *Chinese Journal of Evidence-Based Medicine*, 2019, 19(11): 1268–1275.] DOI: [10.7507/1672-2531.201905076](https://doi.org/10.7507/1672-2531.201905076).
- 11 赵塔娜,曾茵茹,王瑞,等.中国女性 HPV 感染危险因素 Meta 分析[J]. *中国艾滋病性病*, 2022, 28(11): 1334–1338. [Zhao TN, Zeng YR, Wang R, et al. Meta-analysis of risk factors for HPV infection in Chinese women[J]. *Chinese Journal of AIDS & STD*, 2022, 28(11): 1334–1338.] DOI: [10.13419/j.cnki.aids.2022.11.28](https://doi.org/10.13419/j.cnki.aids.2022.11.28).
- 12 洪艳,程正兰,汪俊丽.马鞍山市 12 935 名妇女高危型 HPV 感染现状及影响因素[J]. *华南预防医学*, 2022, 48(12): 1524–1526, 1530. [Hong Y, Cheng ZL, Wang JL. Status and influencing factors of high-risk HPV infection among 12 935 women in Maanshan City[J]. *South China Journal of Preventive Medicine*, 2022, 48(12): 1524–1526, 1530.] DOI: [10.12183/j.sejpm.2022.1524](https://doi.org/10.12183/j.sejpm.2022.1524).
- 13 Isabirye A, Mbonye M, Asiimwe JB, et al. Factors associated with HPV vaccination uptake in Uganda: a multi-level analysis[J]. *BMC Womens Health*, 2020, 20(1): 145. DOI: [10.1186/s12905-020-01014-5](https://doi.org/10.1186/s12905-020-01014-5).

- 14 Aimagambetova G, Babi A, Issa T, et al. What factors are associated with attitudes towards HPV vaccination among Kazakhstani women? Exploratory analysis of cross-sectional survey data[J]. *Vaccines (Basel)*, 2022, 10(5): 824. DOI: [10.3390/vaccines10050824](https://doi.org/10.3390/vaccines10050824).
- 15 Chowdhury S, Ara R, Roy S, et al. Knowledge, attitude, and practices regarding human papillomavirus and its' vaccination among the young medical professionals and students of Bangladesh[J]. *Clin Exp Vac Res*, 2022, 11(1): 63–71. DOI: [10.7774/cevr.2022.11.1.63](https://doi.org/10.7774/cevr.2022.11.1.63).
- 16 Rios Velazquez E, Hoebbers F, Aerts HJ, et al. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging[J]. *Radiother Oncol*, 2014, 113(3): 324–330. DOI: [10.1016/j.radonc.2014.09.005](https://doi.org/10.1016/j.radonc.2014.09.005).
- 17 Albosale AH, Mashkina EV. Association between promoter polymorphisms of IL-1B, IL-4 and IL-6 Genes and a viral load infected women with human papillomavirus[J]. *J Reprod Infertil*, 2021, 22(2): 92–102. DOI: [10.18502/jri.v22i2.5794](https://doi.org/10.18502/jri.v22i2.5794).
- 18 Riza E, Karakosta A, Tsiampalis T, et al. Knowledge, attitudes and perceptions about cervical cancer risk, prevention and human papilloma virus (HPV) in vulnerable women in Greece[J]. *Int J Environ Res Public Health*, 2020, 17(18): 6892. DOI: [10.3390/ijerph17186892](https://doi.org/10.3390/ijerph17186892).
- 19 Bogani G, Sopracordevole F, Di Donato V, et al. High-risk HPV-positive and-negative high-grade cervical dysplasia: analysis of 5-year outcomes[J]. *Gynecol Oncol*, 2021, 161(1): 173–178. DOI: [10.1016/j.ygyno.2021.01.020](https://doi.org/10.1016/j.ygyno.2021.01.020).
- 20 Gopalani SV, Sedani AE, Janitz AE, et al. Barriers and factors associated with HPV vaccination among American Indians and Alaska natives: a systematic review[J]. *J Commun Health*, 2022, 47(3): 563–575. DOI: [10.1007/s10900-022-01079-3](https://doi.org/10.1007/s10900-022-01079-3).
- 21 Olarewaju VO, Jafflin K, Deml MJ, et al. The youth attitudes about vaccines (YAV-5) scale: adapting the parent attitudes about childhood vaccines short scale for use with youth in German, French, and Italian in Switzerland, exploratory factor analysis and mokken scaling analysis[J]. *Hum Vaccin Immunother*, 2021, 17(12): 5183–5190. DOI: [10.1080/21645515.2021.1980314](https://doi.org/10.1080/21645515.2021.1980314).
- 22 Bogani G, Lalli L, Sopracordevole F, et al. Development of a nomogram predicting the risk of persistence/recurrence of cervical dysplasia[J]. *Vaccines (Basel)*, 2022, 10(4): 579. DOI: [10.3390/vaccines10040579](https://doi.org/10.3390/vaccines10040579).
- 23 孟令昊, 胥秋艳, 李科, 等. 1990—2019 年中国女性宫颈癌疾病负担变化的分析 [J]. *中国循证医学杂志*, 2021, 21(6): 648–653. [Meng LH, Xu QY, Li K, et al. The burden of disease of Chinese females cervical cancer from 1990 to 2019[J]. *Chinese Journal of Evidence-Based Medicine*, 2021, 21(6): 648–653.] DOI: [10.7507/1672-2531.202101091](https://doi.org/10.7507/1672-2531.202101091).
- 24 Kotani K, Iwata A, Kukimoto I, et al. Nomogram for predicted probability of cervical cancer and its precursor lesions using miRNA in cervical mucus, HPV genotype and age[J]. *Sci Rep*, 2022, 12(1):16231. DOI: [10.1038/s41598-022-19722-3](https://doi.org/10.1038/s41598-022-19722-3).

收稿日期: 2024 年 05 月 31 日 修回日期: 2024 年 07 月 29 日  
本文编辑: 李绪辉 曹越

引用本文: 唐逸娇, 曾德玲, 谭松, 等. 成都地区成年健康体检女性生殖道 HPV 感染情况调查及预测模型构建[J]. *医学新知*, 2024, 34(10): 1110–1120. DOI: [10.12173/j.issn.1004-5511.202405134](https://doi.org/10.12173/j.issn.1004-5511.202405134).  
Tang YJ, Zeng DL, Tan S, et al. Survey and construction of prediction model of HPV infection in the genital tract of women undergoing health checkups in Chengdu[J]. *Yixue Xinzhi Zazhi*, 2024, 34(10): 1110–1120. DOI: [10.12173/j.issn.1004-5511.202405134](https://doi.org/10.12173/j.issn.1004-5511.202405134).