

DOI: 10. 12138/j. issn. 1671-9638. 20246108

· 论 著 ·

基于 ARIMA 模型的天津地区单中心 HPV 感染趋势及基因型特征

李 杨, 谭桂兰, 李 怡, 谢晓媛, 李 姝, 吴 芳, 刘 霞

(天津市妇女儿童保健中心妇女保健科, 天津 300299)

[摘要] **目的** 采用自回归移动平均(ARIMA)模型构建时间序列, 分析天津地区单中心人乳头瘤病毒(HPV)感染趋势及基因型特征。**方法** 选择 2018 年 1 月—2022 年 12 月某院进行 HPV 检测的 7 236 例女性患者, 比较 2018—2022 年天津地区 HPV 感染情况及基因型分布。建立 ARIMA 模型时间序列, 分析模型拟合。预测 2023 年 HPV 感染数, 并与实际发生数进行比较, 评价模型的预测效果。**结果** 2018—2022 年天津地区 HPV 感染率为 14.41%; HPV 感染率在 31~40 岁年龄段最高, 感染率为 15.47%。阳性标本中 HPV 单一型别感染比率最高, 占比为 73.54%(767/1 043), 以高危型 HPV 为主。低危型感染占比最高的是 HPV-6 型, 为 2.59%, 高危型感染占比最高的是 HPV-16 型, 为 16.06%。建立 ARIMA 模型, 确定最佳模型为 ARIMA(0, 1, 2)(0, 1, 1)₁₂, 其 AIC 值和 BIC 值分别为 3.877、4.005, 经白噪声检验 Ljung-Box $Q = 8.828$ 差异无统计学意义($P > 0.05$)。利用模型预测 2023 年 HPV 感染数, 实际值、预测值的总体趋势基本保持一致, 模型 RMSE、MAPE、MAE 分别为 6.289、34.149、4.706, 提示模型的预测效果较好。**结论** 天津地区女性人群中, HPV 病毒感染类型以单一高危型感染为主, 其中 HPV-16 型感染率最高。天津地区 HPV 感染存在季节性, ARIMA 模型在 HPV 感染流行趋势的预测中效果较好, 适用于短期预测。

[关键词] 自回归移动平均模型; 人乳头瘤病毒; 基因型分布; 感染趋势; HPV

[中图分类号] R181.3⁺2

Trend and genotype characteristics of single-center HPV infection in Tianjin area based on ARIMA model

LI Yang, TAN Gui-lan, LI Yi, XIE Xiao-yuan, LI Shu, WU Fang, LIU Xia (Department of Women's Health, Tianjin Women and Children's Health Center, Tianjin 300299, China)

[Abstract] **Objective** To construct time-series by adopting autoregressive integrated moving average (ARIMA) model for analyzing the trend and genotype characteristics of single-center human papillomavirus (HPV) infection in Tianjin area. **Methods** A total of 7 236 female patients who underwent HPV testing in a hospital from January 2018 to December 2022 were selected. HPV infection status and genotype distribution in Tianjin area from 2018 to 2022 were compared. ARIMA model time-series was constructed, and model fitting was analyzed. The number of HPV infections in 2023 was predicted and compared with the actual occurrence, the predictive performance of the model was evaluated. **Results** HPV infection rate in Tianjin area from 2018 to 2022 was 14.41%, with the highest rate (15.47%) in the age group of 31-40 years. Among the positive specimens, the proportion of single type HPV infection was the highest, accounting for 73.54% (767/1 043), with high-risk HPV being the main type. The highest infection rates of low-risk and high-risk types were type HPV-6 (2.59%) and type HPV-16 (16.06%), respectively. ARIMA model was constructed, and the optimal model was ARIMA(0, 1, 2)(0, 1, 1)₁₂, with akaike information criterion (AIC) and bayesian information criterion (BIC) values of 3.877 and 4.005, respectively. There was no statistical significance in Ljung-Box $Q = 8.828$ showed by white noise test ($P > 0.05$). The number of HPV infection in 2023 was predicted by the model. The overall trend of the actual value and the predicted value was basi-

[收稿日期] 2024-02-18

[作者简介] 李杨(1980-), 女(汉族), 天津市人, 副主任医师, 主要从事妇产科宫颈病变研究。

[通信作者] 刘霞 E-mail: Lx13579@126.com

cally consistent, RMSE, MAPE and MAE of the model were 6.289, 34.149 and 4.706, respectively, suggesting that the model had a good prediction effect. **Conclusion** Among the female population in Tianjin area, HPV infection is mainly caused by single, high-risk type, with HPV-16 having the highest infection rate. There is seasonal variation in HPV infection in Tianjin. ARIMA model has good prediction effect on the prevalence trend of HPV infection, which is suitable for short-term prediction.

[Key words] autoregressive integrated moving average model; human papillomavirus; genotype distribution; infection trend; HPV

人乳头瘤病毒(human papillomavirus, HPV)是一种双链、环状 DNA 病毒,人类是 HPV 唯一宿主^[1]。HPV 主要通过性传播感染,有性生活的女性在其一生中 75%~80% 概率会感染 HPV,虽大多数 HPV 感染无症状并会在 1~2 年内自愈,但持续感染或再感染会导致感染者发病甚至死亡^[2-3]。研究^[4-5]表明,HPV 感染具有一定季节性,但研究中不同地区感染高峰时间并不一致,故准确判断当地 HPV 季节性感染趋势对疾病的防控具有积极意义。目前,已经发现的 HPV 基因型约有 200 种不同分型,其中 20 多种是确定或可能致癌的基因型,根据其致病性,HPV 被分为高危型和低危型两类。研究^[6-7]显示,低危型 HPV 感染主要导致外生殖器疣;而高危型感染可引发癌前病变甚至癌变,除了宫颈癌外,还可导致阴道癌、肛门癌及外阴癌等多种癌变。不同地区 HPV 流行趋势存在着较大差异^[8-10]。为了解天津地区 HPV 亚型分布及感染趋势,本研究对当地 HPV 基因型分布进行分析,并建立自回归移动平均 (autoregressive integrated moving average, ARIMA) 乘积季节模型对 HPV 感染趋势进行预测,研究旨在为天津地区 HPV 感染的早期预警和防控策略的选择提供有力依据。

1 对象与方法

1.1 研究对象 选择 2018 年 1 月—2022 年 12 月某院 7 236 例进行 HPV 检测的女性作为研究对象,年龄 20~85 岁,平均年龄(33.98 ± 8.38)岁。纳入标准:①年龄 ≥ 18 岁;②3 d 内未对阴道进行冲洗;③24 h 内无性行为;④自愿接受 HPV 检测。排除标准:①3 d 内阴道用药者;②处于妊娠期内;③处于月经期内;④宫颈或生殖道畸形者。研究对象均签署知情同意书。

1.2 研究方法

1.2.1 检测方法 HPV 检测标本采集,使用阴道窥器充分暴露宫颈,宫颈刷采集宫颈脱落上皮细胞

标本,将宫颈刷浸入细胞保存液中,检测时将标本离心 1 min(转速 10 000 r/min,半径 3 cm)弃上清液,加入裂解酶液,提取 DNA。溶液配置:PCR mix 23.25 μL + DNA taq 酶 0.75 μL + 1 μL 模板。反应条件:95℃ 预变性 5 min;95℃ 变性 1 min,55℃ 退火 30 s,72℃ 延伸 30 s,共 40 个循环;72℃ 终延伸 5 min。导流杂交:①取 PCR 产物 20 μL,95℃ 加热 5 min,冰水浴至少 2 min。②1 μL 杂交液预热 45℃,温育至少 2 min。③0.5 mL 杂交液加至 DNA 样品溶液中预热 45℃,温育 10 min;使用 0.8 mL 杂交液反复洗膜 3 次。④杂交仪设定 25℃,加入 0.5 mL 酶标液温育 4 min 开泵,36℃ 彻底洗膜 4 次,加入 NBT/BCIP 溶液 0.5 mL,盖上盖板显色 3~6 min。⑤阳性点为蓝紫色圆点,根据 HPV 分型比对图(图 1)判断 HPV 分型。

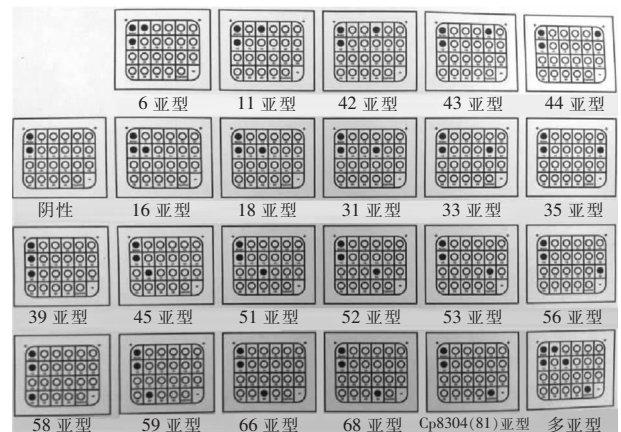


图 1 HPV 分型比对图

Figure 1 Comparison of HPV typing

1.2.2 模型说明 ARIMA 模型分为简单 ARIMA (p, d, q) 模型、季节 ARIMA (P, D, Q)_s 模型和乘积季节 ARIMA (p, d, q) (P, D, Q)_s 模型, d 表示时序数据变为平稳序列的差分次数, D 为季节性差分次数; p 为自回归系数, P 为季节性自回归系数; q 为滑动平均系数, Q 为季节性滑动平均系数; s 为季节周期和循环长度。

1.2.3 构建 ARIMA 模型 按照 2018 年 1 月—2022 年 12 月时间轴制作 HPV 感染例数序列图。评判序列的平稳性,由于 HPV 感染例数随周期、季节性变化存在不稳定的序列。对其进行平稳性处理,检查是否白噪声,建立拟合 ARIMA 模型。利用自相关分析 (autocorrelation function, ACF) 和偏自相关分析 (partial autocorrelation function, PACF) 为目标序列定阶。参数 d, D 根据差分次数确定,参数 p, q, P, Q 值均分别取 0、1、2,筛选出赤池信息准则 (akaike information criterion, AIC) 和贝叶斯信息准则 (bayesian information criterion, BIC) 最小值的模型。利用 Ljung-Box Q 统计值对残差序列进行白噪声检验。利用选定模型对 2018—2022 年 HPV 感染检出数据进行回代模拟,并对 2023 年的数据进行预测。

1.3 统计学分析 应用 SPSS 22.0 软件进行数据分析,2018—2022 年 HPV 感染整体比较采用 χ^2 检验;构建 2018—2022 年天津地区 HPV 感染 ARIMA 模型,通过比较 AIC 和 BIC 筛选模型,使用 Ljung-Box Q 检验检验序列是否为白噪声,若 Ljung-Box Q 检验 $P > 0.05$,则确定为白噪声序列。

2 结果

2.1 HPV 基因型分布 2018—2022 年天津地区 1 043 例 HPV 阳性患者中,低危型感染占比最高的是 HPV-6 型 (2.59%),其次为 HPV-11 型 (1.89%)。高危型感染占比最高的是 HPV-16 型 (16.06%),之后 4 位依次为 HPV-52 型 (13.04%)、HPV-58 型 (12.48%)、HPV-66 型 (9.05%)、HPV-51 型 (7.22%),感染占比最低的是 HPV-68 型 (2.73%)。在单一感染和多重感染患者中,低危型感染占比最高的均为 HPV-6 型,分别为 1.33%、1.26%。单一感染患者中高危型感染占比最高的为 HPV-16 型,为 10.38%;多重感染患者中高危型感染占比最高的为 HPV-52 型,为 6.03%,见表 1。

2.2 HPV 基因分型感染情况 7 236 例被检标本中 1 043 例检测阳性,总感染率 14.41%。HPV 单一型别感染阳性率为 73.54% (767 例),单一型别感染中 HPV 单一高危型阳性率最高,为 70.37% (734 例),其余依次为 HPV 多重感染高危型阳性率 23.49% (245 例) 和 低危型感染阳性率 3.36% (35 例),见表 2、3。

表 1 1 043 例 HPV 感染者各亚型分布情况 [例 (%)]

Table 1 Distribution of subtypes in 1 043 cases of HPV infection (No. of cases [%])

类型	感染数	单一感染	多重感染
HPV-16	229(16.06)	148(10.38)	81(5.68)
HPV-52	186(13.04)	100(7.01)	86(6.03)
HPV-58	178(12.48)	97(6.80)	81(5.68)
HPV-66	129(9.05)	78(5.47)	51(3.58)
HPV-51	103(7.22)	54(3.79)	49(3.43)
HPV-18	101(7.08)	48(3.36)	53(3.72)
HPV-31	76(5.33)	36(2.53)	40(2.80)
HPV-56	70(4.91)	36(2.53)	34(2.38)
HPV-39	60(4.21)	27(1.89)	33(2.32)
HPV-45	51(3.58)	25(1.75)	26(1.83)
HPV-33	50(3.51)	26(1.83)	24(1.68)
HPV-59	49(3.44)	23(1.61)	26(1.83)
HPV-35	41(2.88)	15(1.05)	26(1.83)
HPV-68	39(2.73)	21(1.47)	18(1.26)
HPV-6	37(2.59)	19(1.33)	18(1.26)
HPV-11	27(1.89)	14(0.98)	13(0.91)

表 2 2018—2022 年某院 7 236 例 HPV 检测者感染情况

Table 2 Infection status of 7 236 persons who underwent HPV testing in a hospital, 2018 - 2022

类型	阳性例数	阳性比率 (%)	总感染率 (%)
单一感染	767	73.54	10.60
双重感染	192	18.41	2.65
三重感染	63	6.04	0.87
四重感染	19	1.82	0.26
五重感染及以上	2	0.19	0.03
合计	1 043	100	14.41

表 3 2018—2022 年某院 7 236 例 HPV 检测者高危型、低危型 HPV 分布及感染情况

Table 3 Distribution and infection status of high-risk and low-risk HPV in 7 236 persons who underwent HPV testing in a hospital, 2018 - 2022

类型	阳性例数	阳性比率 (%)	总感染率 (%)
单一高危型感染	734	70.37	10.14
高危型 + 高危型	245	23.49	3.39
低危型感染	35	3.36	0.48
高危型 + 低危型	29	2.78	0.40
合计	1 043	100	14.41

2.3 不同年龄者 HPV 感染情况 将 7 236 例被检标本按年龄进行分组,各组检测结果见表 4。HPV 感染率在 31~40 岁年龄段最高,感染率为 15.47%,且以高危型 HPV 为主。年龄≥61 的患者 HPV 单

一感染比率最高,占感染总数的 91.67%;感染者中 41~50 岁年龄段患者 HPV 多重感染比率最高,占比为 30.08%。

表 4 不同年龄被检测者 HPV 感染及感染类型分布情况

Table 4 Distribution of infection status and infection types of HPV in tested persons of different ages

年龄(岁)	总标本数(份)	阳性		高危型		单一感染		多重感染	
		例数	感染率(%)	例数	感染率(%)	例数	感染率(%)	例数	感染率(%)
≤30	2 188	336	15.36	329	15.04	245	11.20	91	4.16
31~40	2 696	417	15.47	409	15.17	312	11.57	105	3.89
41~50	1 890	246	13.02	231	12.22	172	9.10	74	3.92
51~60	268	32	11.94	30	11.19	27	10.07	5	1.87
≥61	194	12	6.19	9	4.64	11	5.67	1	3.81

2.4 2018—2022 年天津地区 HPV 感染情况 2018—2022 年天津地区 5 年间 HPV 感染率的总体差异不具有统计学意义($\chi^2 = 0.853, P = 0.837$)。

HPV 感染存在季节性。每年 9—12 月为天津地区 HPV 感染高峰期,其中 11 月比率最高(16.01%)。见表 5。

表 5 2018—2022 年不同月份天津地区 HPV 感染情况(例)

Table 5 HPV infection status in Tianjin area in different months, 2018 - 2022 (Cases)

月份	2018 年 (n=229)	2019 年 (n=222)	2020 年 (n=205)	2021 年 (n=194)	2022 年 (n=193)	合计 [例(%)]	月份	2018 年 (n=229)	2019 年 (n=222)	2020 年 (n=205)	2021 年 (n=194)	2022 年 (n=193)	合计 [例(%)]
1	12	13	8	16	10	59(5.66)	7	13	13	16	15	10	67(6.42)
2	9	16	13	10	14	62(5.94)	8	21	16	10	19	13	79(7.57)
3	13	10	8	12	10	53(5.08)	9	29	16	23	13	18	99(9.49)
4	16	13	10	7	16	62(5.94)	10	31	36	31	28	26	152(14.57)
5	8	16	10	12	13	59(5.66)	11	33	34	39	26	35	167(16.01)
6	18	21	8	13	7	67(6.42)	12	26	18	29	23	21	117(11.22)

2.5 时间序列分析 将 HPV 感染每月阳性病例数作为因变量,时间作为自变量生成数据序列图。观察序列图,随着时间有一定波动,序列具有一定的周期性和季节性特点,不满足构建模型的序列平稳性要求,见图 2。

2.6 序列平稳化 时间序列数据的稳定性较差,需对原始数据进行 1 阶差分;HPV 感染具有一定的季节性,再次采用 1 阶长度为 12 的季节差分来消除季节性影响。经过差分处理后,序列呈现较为平稳的状态,因此可以确定模型参数 $d = 1, D = 1, s = 12$, 见

图 3。

2.7 模型的识别与定阶 根据差分后的平稳时间序列,绘制 ACF 图和 PACF 图,并对残差的 ACF 和 PACF 进行分析,见图 4。初步确定模型为 $ARIMA(p, 1, q)(P, 1, Q)_{12}$, 对其进行残差相关性检验,检验值基本落在 95% 置信限内。对 p, q, P 和 Q 的取值进行尝试,比较各模型 AIC 值和 BIC 值,筛选模型。 $ARIMA(0, 1, 2)(0, 1, 1)_{12}$ 模型 AIC 值和 BIC 值最小,分别为 3.877、4.005(见表 6),因此研究将 $ARIMA(0, 1, 2)(0, 1, 1)_{12}$ 模型选为最优模型。

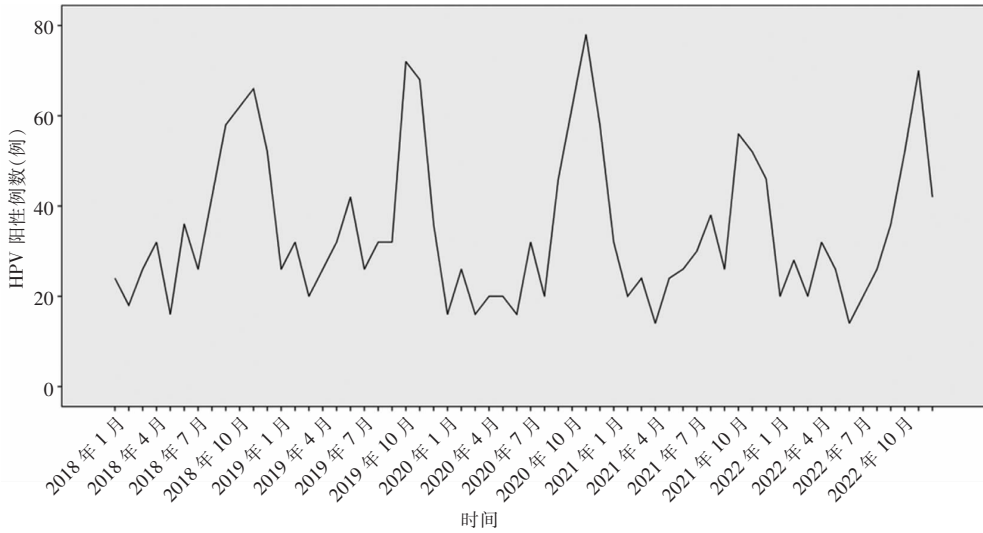


图 2 2018—2022 年天津地区 HPV 阳性例数时序图

Figure 2 Time-series of HPV-positive cases in Tianjin area, 2018 – 2022

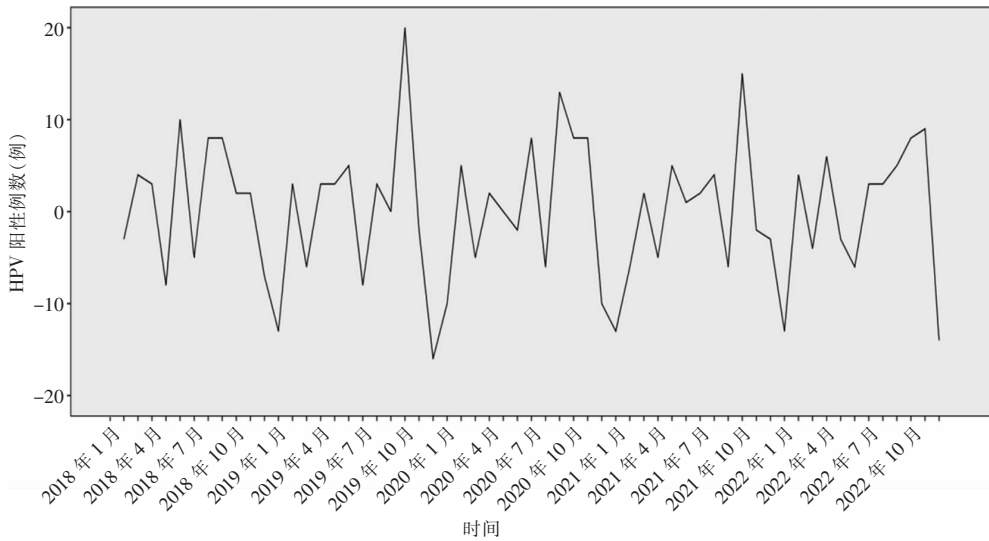


图 3 2018—2022 年天津地区 HPV 阳性例数 1 阶差分 and 1 阶季节差分后的时序图

Figure 3 Time-series of HPV-positive cases in Tianjin area after first-order difference and first-order seasonal difference, 2018 – 2022

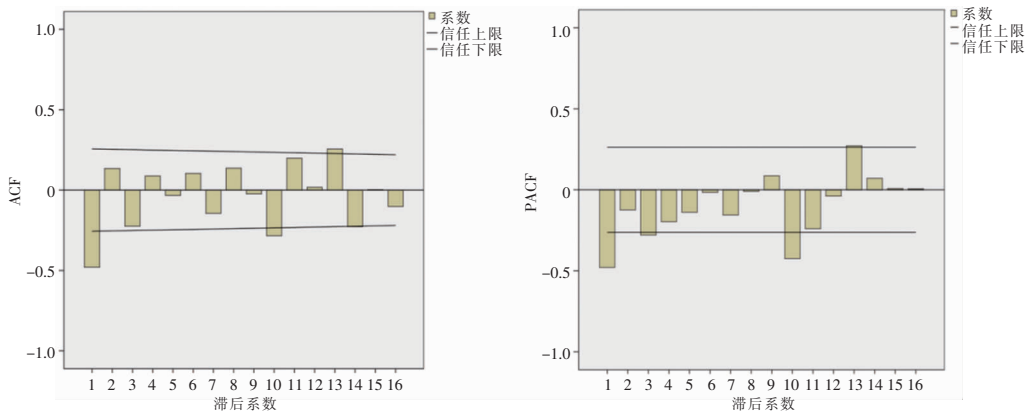


图 4 HPV 阳性序列残差自相关图

Figure 4 Autocorrelation of residual HPV positive sequences

表 6 各 ARIMA 乘积模型 AIC 值和 BIC 值对比

Table 6 Comparison of AIC value and BIC value of each ARIMA multiplicative model

AIC/BIC	(p,d,q)	(P,D,Q)								
		(0,1,0)	(0,1,1)	(0,1,2)	(1,1,0)	(1,1,1)	(1,1,2)	(2,1,0)	(2,1,1)	(2,1,2)
AIC	(0,1,0)	4.211	3.979	3.988	3.934	3.961	4.151	4.213	4.191	4.042
	(0,1,1)	3.927	3.938	4.197	4.206	4.174	4.001	4.181	3.902	4.104
	(0,1,2)	4.025	3.877	3.941	4.021	4.169	4.081	3.952	3.944	3.930
	(1,1,0)	4.199	3.914	4.081	3.891	3.903	3.988	4.054	4.071	4.085
	(1,1,1)	3.888	4.168	3.903	3.948	4.019	4.118	3.903	4.142	3.982
	(1,1,2)	4.009	4.036	4.188	4.224	4.001	4.099	3.915	4.023	4.065
	(2,1,0)	4.110	3.898	4.173	3.966	4.051	4.128	3.997	4.082	3.984
	(2,1,1)	3.949	3.980	3.966	4.043	4.056	4.097	4.047	4.228	4.014
	(2,1,2)	4.083	4.111	4.179	3.897	4.213	4.112	3.941	4.159	4.121
BIC	(0,1,0)	4.351	4.109	4.112	4.063	4.088	4.286	4.345	4.331	4.168
	(0,1,1)	4.056	4.070	4.330	4.342	4.312	4.133	4.315	4.024	4.232
	(0,1,2)	4.154	4.005	4.068	4.153	4.299	4.213	4.084	4.070	4.056
	(1,1,0)	4.339	4.041	4.215	4.016	4.028	4.115	4.184	4.201	4.214
	(1,1,1)	4.016	4.299	4.026	4.075	4.145	4.254	4.033	4.280	4.109
	(1,1,2)	4.140	4.163	4.327	4.364	4.128	4.232	4.046	4.153	4.199
	(2,1,0)	4.242	4.028	4.303	4.091	4.182	4.262	4.127	4.215	4.112
	(2,1,1)	4.078	4.108	4.091	4.169	4.191	4.227	4.177	4.366	4.144
	(2,1,2)	4.220	4.248	4.312	4.026	4.345	4.240	4.072	4.296	4.252

2.8 参数估计和模型诊断 经比较 ARIMA(0,1,2)(0,1,1)₁₂ 模型 AIC 值和 BIC 值相对较小,且平稳 R² 相对较大,选为最优模型。绘制 ARIMA(0,1,2)(0,1,1)₁₂ 模型的残差序列 ACF 图和 PACF 图,如图 5 所示,残差基本处于 95% 置信区间,经白噪声检验 Ljung-Box Q = 8.828 差异无统计学意义(P > 0.05),表明该模型残差序列为白噪声(见表 7)。初步确定该模型可以用来预测 HPV 感染的月检出数。

表 7 ARIMA 模型拟合度和白噪声检验

Table 7 ARIMA model fitting and white noise test

参数	ARIMA(0,1,2)(0,1,1) ₁₂ 模型
模型拟合度	平稳 R ² 0.879
	RMSE 6.289
	MAPE 34.149
	MAE 4.706
	AIC 3.877
	BIC 4.005
Ljung-Box Q	检测值 8.828
	P 0.886

2.9 模型预测效果评估 利用 2018—2022 年天津地区 HPV 感染数据回代入 ARIMA(0,1,2)(0,1,1)₁₂ 模型进行拟合,并预测 2023 年数据。选取 2023 年 HPV 阳性检出数作为样本,将实际检出数与预测值进行比较,结果表明实际检出数均处于预测值的 95% 置信区间内,实际值、预测值的总体趋势基本保持一致,见图 6。模型 RMSE、MAPE、MAE 分别为 6.289、34.149、4.706,提示模型的预测效果较好,可用于 HPV 感染流行趋势的短期预测和动态分析。

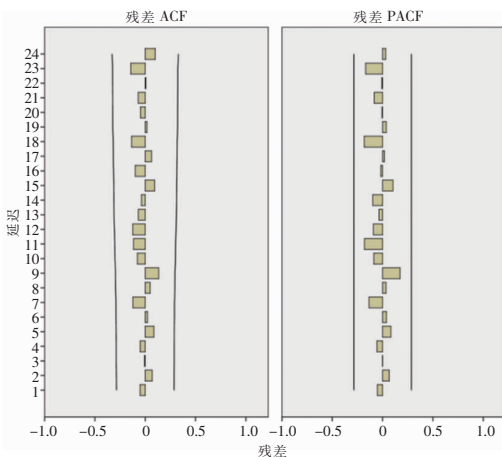


图 5 ARIMA(0,1,2)(0,1,1)₁₂ 模型残差序列的 ACF 图和 PACF 图

Figure 5 ACF and PACF graphs of residual sequence of ARIMA(0,1,2)(0,1,1)₁₂ model

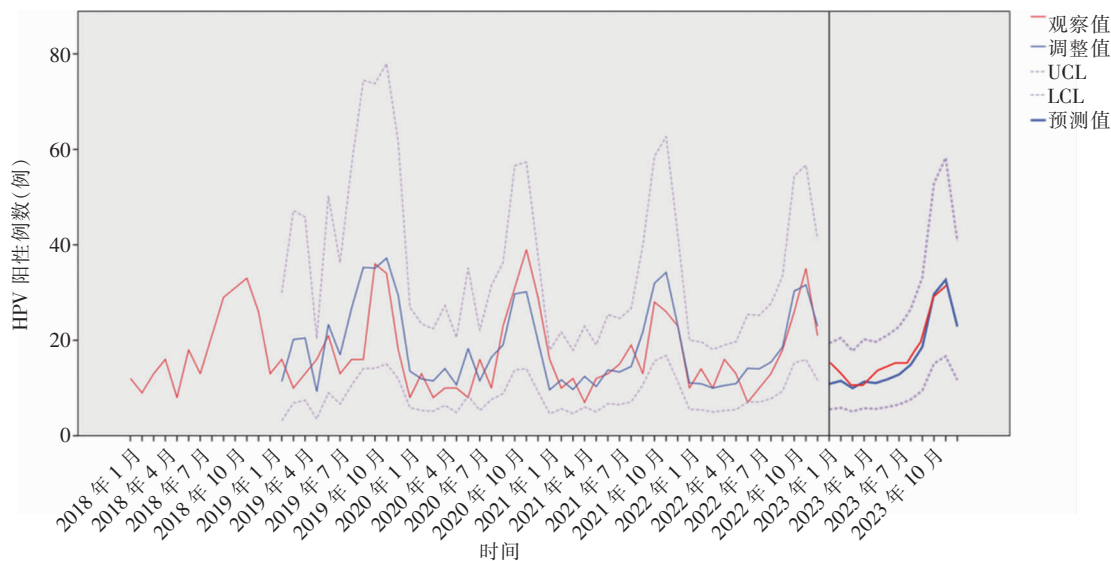


图 6 2018—2022 年天津地区 HPV 感染数的拟合及预测

Figure 6 Fitting and predicting of HPV infection number in Tianjin area, 2018 – 2022

3 讨论

HPV 病毒基因型繁多,主要分为高危型和低危型两类。低危型 HPV 主要引起生殖器官的尖锐湿疣,对身体健康的威胁相对较小。然而,高危型 HPV 的持续感染则可能会导致女性外阴、阴道和宫颈的损伤,甚至发生癌变,严重威胁女性的生命健康^[11-13]。因此,早期检测 HPV 病毒基因型,尤其是高危型,对于预防和治疗宫颈癌等疾病至关重要。通过及早预防和干预,可以有效降低宫颈癌和相关疾病的发病率和病死率^[14]。HPV 的基因型分布和流行具有显著的地域和人群特征。在我国,不同地域和民族之间的 HPV 感染率存在差异。如维吾尔族女性的感染率高于同一地区的回族和汉族^[15-17]。在中国西部地区,女性 HPV 总感染率为 18.24%,而中国东部地区的女性 HPV 总感染率为 15.53%^[18]。在本研究中天津地区 7 236 例样本中 14.41% 感染 HPV,其中低危 HPV 基因型以 HPV6、HPV11 为主,此结论与既往报道中最常见的低危 HPV 类型为 HPV6 和 HPV11 结果一致^[19];高危型 HPV 中感染占比居前 3 位依次为 HPV-16、HPV-52、HPV-58,与近期在我国洛阳、延安及杭州等地区对女性 HPV 感染率的研究结论基本一致^[20-22]。在 1 043 例阳性样本中,HPV 单一高危型感染构成比(73.54%)最高,HPV 多重感染高危型构成比为 26.27%(其中高危型 + 低危险占 2.78%),高危型 HPV 与宫颈

癌发生发展密切相关^[23-24],对于此类感染者,应定期复查,并在必要时采取适当的干预措施。

HPV 感染具有一定季节性,王铁等^[4]研究认为,长春地区 HPV 阳性率有季节性变化,在 4—6 月份最高,7—9 月份最低;Chen 等^[5]研究显示,贵州地区 HPV 感染率与季节存在一定的相关性。与其他季节相比,春季的感染率呈现出相对的峰值。本研究中 2018—2022 每年 9—12 月为 HPV 感染高峰期,其中 11 月比例最高,本研究地处天津地区,与贵州、长春等地区经纬度不同,即使是同月份温度、湿度也有一定的差异。可能由于该地区秋季昼夜气温变化较大,为病毒复制创造有利条件,感染者的症状会加重导致就诊量的增加,也与低温导致 HPV 感染者或易感人群清洁次数的减少有关^[25-26],在 HPV 预防策略的制定中,建议将季节性因素纳入考量。影响一个地区疾病流行的因素众多且复杂,如经济状况、气候条件、行为习惯以及环境流行因素的变化等。这些因素与时间紧密关联,随时间变化而变化,对疾病的流行趋势产生影响,可运用 ARIMA 时间序列分析方法进行定量预测^[27-29],本研究中利用季节性 ARIMA(0,1,2)(0,1,1)₁₂模型,较好地模拟了天津地区 HPV 感染趋势,模型 MAE、RMSE、MAPE 值分别为 4.706、6.289、34.149,对短期内的预测也取得了较好效果,误差均在预测模型的 95% 置信限内,预测的 HPV 感染流行趋势与实际发病数量大致相符。本研究尚存在一定的局限性,纳入的样本数据来自同一医疗中心,结果难免存

在一定偏倚;研究对象均为天津汉族,未涉及其他民族、地区女性,后续研究还需进一步扩大样本量。

综上所述,天津地区女性人群中,HPV 病毒感染类型以单一高危型感染为主,高危型检出率最高的 5 种基因型依次为 HPV-16 型、HPV-52 型、HPV-58 型、HPV-66 型和 HPV-51 型,HPV 感染率在 31~40 岁年龄段最高。天津地区 HPV 感染存在季节性,每年 9—12 月为天津地区 HPV 感染高峰期,其中 11 月比例最高,ARIMA 模型在预测 HPV 感染流行趋势方面表现较好,适用于短期预测。

利益冲突:所有作者均声明不存在利益冲突。

[参考文献]

- [1] Szymonowicz KA, Chen JJ. Biological and clinical aspects of HPV-related cancers[J]. *Cancer Biol Med*, 2020, 17(4): 864–878.
- [2] Li YY, Yu T, Yan H, et al. Vaginal microbiota and HPV infection: novel mechanistic insights and therapeutic strategies [J]. *Infect Drug Resist*, 2020, 13: 1213–1220.
- [3] Yang D, Zhang J, Cui XL, et al. Risk factors associated with human papillomavirus infection, cervical cancer, and precancerous lesions in large-scale population screening[J]. *Front Microbiol*, 2022, 13: 914516.
- [4] 王铁, 李春子, 许建成. 长春地区 38 260 例女性高危型人乳头瘤病毒感染现状研究[J]. *中国实验诊断学*, 2021, 25(12): 1774–1777.
Wang T, Li CZ, Xu JC. Analysis of the infection status of high-risk HPV in Changchun 38 260 females [J]. *Chinese Journal of Laboratory Diagnosis*, 2021, 25(12): 1774–1777.
- [5] Chen ZY, Li QY, Huang Q, et al. Characteristics of human papillomaviruses distribution in Guizhou Province, China[J]. *Virology*, 2019, 16(1): 123.
- [6] Wencel-Wawrzęczyk A, Lewitowicz P, Lewandowska A, et al. Sexual behavior and the awareness level of common risk factors for the development of cervical, anogenital and oropharyngeal cancer among women subjected to HR HPV DNA-testing[J]. *Int J Environ Res Public Health*, 2022, 19(15): 9580.
- [7] Trama JP, Trikannad C, Yang JJ, et al. High-risk HPV genotype distribution according to cervical cytology and age[J]. *Open Forum Infect Dis*, 2022, 9(11): ofac595.
- [8] Taskin MH, Nursal AF, Oruc MA, et al. Genotype distribution and prevalence of high-risk human papillomavirus infection among women in Samsun province of Turkey[J]. *Asian Pac J Cancer Prev*, 2022, 23(7): 2477–2482.
- [9] Lin XH, Chen L, Zheng YY, et al. Age-specific prevalence and genotype distribution of human papillomavirus in women from northwest China[J]. *Cancer Med*, 2022, 11(22): 4366–4373.
- [10] Li XX, Xiang FF, Dai JH, et al. Prevalence of cervicovaginal human papillomavirus infection and genotype distribution in Shanghai, China[J]. *Virology*, 2022, 19(1): 146.
- [11] Singh D, Vignat J, Lorenzoni V, et al. Global estimates of incidence and mortality of cervical cancer in 2020: a baseline analysis of the WHO global cervical cancer elimination initiative[J]. *Lancet Glob Health*, 2023, 11(2): e197–e206.
- [12] Dong Z, Hu RJ, Du Y, et al. Immunodiagnosis and immunotherapeutics based on human papillomavirus for HPV-induced cancers[J]. *Front Immunol*, 2021, 11: 586796.
- [13] Kaczmarek M, Baj-Krzyworzeka M, Bogucki Ł, et al. HPV-related cervical cancer and extracellular vesicles[J]. *Diagnosics (Basel)*, 2022, 12(11): 2584.
- [14] Shapiro GK. HPV vaccination: an underused strategy for the prevention of cancer[J]. *Curr Oncol*, 2022, 29(5): 3780–3792.
- [15] 王旭晖, 王萍. 6 352 例妇女宫颈人乳头瘤病毒感染及基因分型的分析研究[J]. *标记免疫分析与临床*, 2020, 27(8): 1328–1332.
Wang XH, Wang P. An analysis of human papillomavirus infection and genotyping in 6 352 women[J]. *Labeled Immunoassays and Clinical Medicine*, 2020, 27(8): 1328–1332.
- [16] Zeng ZY, Austin RM, Wang L, et al. Nationwide prevalence and genotype distribution of high-risk human papillomavirus infection in China[J]. *Am J Clin Pathol*, 2022, 157(5): 718–723.
- [17] Zhang HJ, Zhang SL. Prevalence and genotype distribution of human papillomavirus infection among female outpatients in northeast China: a population-based survey of 110,927 women [J]. *Arch Gynecol Obstet*, 2023, 308(1): 35–41.
- [18] Chen L, Dong Y, Li J, et al. The genomic distribution map of human papillomavirus in western China[J]. *Epidemiol Infect*, 2021, 149: e135.
- [19] 杜珍武, 艾冬晴, 李智慧, 等. 长春地区 2 051 例女性人乳头瘤病毒感染基因类型及感染人群特征分析[J]. *中国体视学与图像分析*, 2021, 26(4): 346–354.
Du ZW, Ai DQ, Li ZH, et al. Genotype distribution and population characteristics of human papillomavirus among 2 051 females in Changchun [J]. *Chinese Journal of Stereology and Image Analysis*, 2021, 26(4): 346–354.
- [20] 季泽益, 诸靖宇, 李瑞鹏, 等. 杭州地区 22 454 例人乳头瘤病毒感染患者基因分型的研究[J]. *中国性科学*, 2022, 31(2): 141–144.
Ji ZY, Zhu JY, Li RP, et al. Genotyping of 22 454 patients with HPV infection in Hangzhou[J]. *Chinese Journal of Human Sexuality*, 2022, 31(2): 141–144.
- [21] 石芳, 赵建芳. 延安地区女性宫颈 HPV 感染特征及基因分型研究[J]. *延安大学学报(医学科学版)*, 2021, 19(3): 55–58, 65.
Shi F, Zhao JF. Study on the characteristics and genotypes of

- female cervical HPV infection in Yan'an area[J]. Journal of Yan'an University(Medical Science Edition), 2021, 19(3): 55-58, 65.
- [22] Wang XL, Han SZ, Li XW, et al. Prevalence and distribution of human papillomavirus (HPV) in Luoyang city of Henan province during 2015-2021 and the genetic variability of HPV16 and 52[J]. Virol J, 2022, 19(1): 37.
- [23] 王丽, 何玉, 许驰. HPV 分型检测在宫颈病变诊断中的价值[J]. 中国妇幼保健, 2021, 36(6): 1232-1234.
Wang L, He Y, Xu C. The value of HPV typing in the diagnosis of cervical lesions[J]. Maternal & Child Health Care of China, 2021, 36(6): 1232-1234.
- [24] 刘一获, 吴欢, 陈德福, 等. 女性人乳头瘤病毒感染与宫颈鳞状上皮内病变的关系[J]. 解放军医学院学报, 2021, 42(3): 251-255, 261.
Liu YD, Wu H, Chen DF, et al. Association between human papillomavirus infection and cervical squamous intraepithelial lesion in women[J]. Academic Journal of Chinese PLA Medical School, 2021, 42(3): 251-255, 261.
- [25] 魏碧娜, 彭臻菲, 过丹丹, 等. 宫颈病变患者 hr-HPV 感染基因分型及其影响因素模型构建[J]. 中华医院感染学杂志, 2023, 33(11): 1700-1704.
Wei BN, Peng ZF, Guo DD, et al. Genotyping of hr-HPV infection in patients with cervical lesions and construction of a model of influencing factors[J]. Chinese Journal of Nosocomiology, 2023, 33(11): 1700-1704.
- [26] 赵塔娜, 曾茵茹, 王瑞, 等. 中国女性 HPV 感染危险因素的 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.
- [27] 蒋建国, 孙定勇, 张艳秋, 等. 应用 ARIMA 模型对河南省肺结核流行趋势预测分析[J]. 现代疾病预防控制, 2023, 34(7): 495-499, 563.
Jiang JG, Sun DY, Zhang YQ, et al. Applying ARIMA model to predict the epidemic trend of pulmonary tuberculosis in Henan[J]. Modern Disease Control and Prevention, 2023, 34(7): 495-499, 563.
- [28] 李平, 黄澳迪, 包黎明, 等. 基于 ARIMA 乘积季节模型的中国流行性腮腺炎发病趋势预测分析[J]. 中国疫苗和免疫, 2023, 29(2): 174-179.
Li P, Huang AD, Bao LM, et al. Analysis and prediction of mumps incidence in China using a multiplicative seasonal ARIMA model[J]. Chinese Journal of Vaccines and Immunization, 2023, 29(2): 174-179.
- [29] 赵棋锋, 马珊珊, 王吉玲, 等. 指数平滑法与 ARIMA 模型对流感样病例流行趋势的预测效果比较[J]. 预防医学, 2020, 32(4): 381-383, 387.
Zhao QF, Ma SS, Wang JL, et al. Comparison of the effect of exponential smoothing method and ARIMA model on predicting the epidemic trend of influenza-like cases[J]. Preventive Medicine, 2020, 32(4): 381-383, 387.

(本文编辑:左双燕)

本文引用格式:李杨,谭桂兰,李怡,等.基于 ARIMA 模型的天津地区单中心 HPV 感染趋势及基因型特征[J].中国感染控制杂志,2024,23(10):1249-1257. DOI: 10.12138/j.issn.1671-9638.20246108.

Cite this article as: LI Yang, TAN Gui-lan, LI Yi, et al. Trend and genotype characteristics of single-center HPV infection in Tianjin area based on ARIMA model[J]. Chin J Infect Control, 2024, 23(10): 1249-1257. DOI: 10.12138/j.issn.1671-9638.20246108.