论 著

2009–2019年泸州市乙型肝炎流行病学特征及其预测

周　瑜[1]，李明珂[2]，汤　艳[1]
1.西南医科大学 公共卫生学院（泸州 646000）；2.西南医科大学 预防医学系（泸州 646000）

【摘　要】**目的** 描述泸州市2009–2019年乙肝的流行病学特征，了解其流行规律和分布特征，预测泸州市乙肝的发病趋势，为乙肝的防控工作提供参考。**方法** 对泸州市乙肝的三间分布特征进行描述，建立乙肝发病ARIMA预测模型。**结果** 2009–2019年乙肝整体发病率呈缓慢下降的趋势，男女发病比例为1.88∶1；发病人群主要为农民，占全部人群的62.5 %。前三位分别为江阳区、泸县和龙马潭区，11年的平均发病率分别为87.29/10万、76.43/10万、70.17/10万。以2009年到2019年的月发病率建立时间序列，最终确定季节乘积模型ARIMA$(1,1,1)(1,1,2)_{12}$为泸州市短期预测乙肝发病的理想模型。**结论** 泸州市近十年乙肝发病下降趋势较为缓慢，无明显季节高峰，发病有地区差异。ARIMA$(1,1,1)(1,1,2)_{12}$模型是泸州市短期预测乙肝发病的最佳模型。

【关键词】乙肝　流行病学特征　ARIMA预测
【中图分类号】R181.2+2　文献标志码 A　DOI：10.3969/j.issn.2096–3351.2021.06.013

Epidemiological characteristics and prediction of hepatitis B in Luzhou, Sichuan Province, China, 2009–2019

ZHOU Yu[1], *LI Ming–ke*[2], *TANG Yan*[1]
1.*School of Public Health, Southwest Medical University, Luzhou 646000, China;* 2.*Department of Preventive Medicine, Southwest Medical University, Luzhou 646000, China*

【**Abstract**】**Objective** To describe the epidemiological characteristics of hepatitis B in Luzhou, Sichuan Province, China from 2009 to 2019; to investigate its epidemic pattern and distribution characteristics; to predict the trend of hepatitis B in Luzhou, and to provide references for hepatitis B prevention and control. **Methods** The three distribution characteristics of hepatitis B in Luzhou were described, and an autoregressive integrated moving average (ARIMA) prediction model for the incidence of hepatitis B was established. **Results** From 2009 to 2019, the overall incidence of hepatitis B showed a slow downward trend, with a male–to–female ratio of 1.88:1. The farmers (62.5%) were the main affected population. The top three areas in terms of incidence rate were Jiangyang District, Luxian County, and Longmatan District, with mean incidence rates in these 11 years of 87.29/100 000, 76.43/100 000 and 70.17/100 000, respectively. A time series was established based on the monthly incidence rates from 2009 to 2019, and the seasonal product model ARIMA$(1,1,1)(1,1,2)_{12}$ was finally determined as an ideal model for short–term prediction of hepatitis B in Luzhou. **Conclusion** The incidence of hepatitis B in Luzhou has declined slowly over the past ten years, and there is no obvious seasonal peak. There are regional differences in incidence. ARIMA$(1,1,1)(1,1,2)_{12}$ model is the optimal model for short–term prediction of hepatitis B in Luzhou.

【**Key words**】Hepatitis B　Epidemiological characteristics　ARIMA prediction

基金项目：泸州市卫生健康委员会（2019–50）
第一作者简介：周　瑜，硕士研究生。
通信作者：汤　艳，教授。E–mail：tangyan200210@163.com

乙型病毒性肝炎是由HBV感染所引起的以肝脏损害为主要病变的慢性传染性疾病，是国家法定的乙类传染病之一[1]，是一种全球性的、危害性较强的传染病。我国乙肝每年发病率一直居甲乙类传染病的前列，我国超过700万人感染乙肝病毒，仍是我国严峻的公共卫生问题[2-3]。感染HBV后可以引起急性肝炎、慢性活动性肝炎、肝硬化和原发性肝癌等疾病。临床上中国治疗肝炎和肝癌的费用每年高达900亿[4-6]，造成了巨大疾病负担。传染病是一种易流行，可控制的疾病，只要掌握了其发病规律就可以极大程度控制其流行，因此对于传染病的预测变得至关重要。自回归移动模型（autoregressive integrated moving average model，ARIMA model）可以对发病率先小幅波动后呈下降趋势的乙肝作出较好的预测[7]。故本文将对泸州市乙肝病例的流行病学特征进行分析，并构建ARIMA模型探索泸州市乙型病毒肝炎的流行规律，为泸州市乙型病毒性肝炎的防控工作提供科学依据。

1 资料来源与方法

1.1 资料来源

2009-2019年泸州市乙肝发病数据来源于中国传染病报告信息管理系统（national notifiable disease report system，NNDRS）网络直报的乙型肝炎个案数据，人口学资料来源于《泸州市统计年鉴》。地图数据来源于中国基础地理信息数据库，选定1∶3 000 000泸州市区县界地图为本研究的主要分析地图。

1.2 乙肝的诊断标准

依据《乙型病毒性肝炎诊断标准》（WS29902008）进行诊断和报告，急性乙肝：有肝炎临床症状且乙肝表面抗原（HBs Ag）阳性，同时符合下列条件之一：①HBs Ag阳性时间小于6个月；②抗-HBc Ig M 1∶1 000阳性；③肝组织学呈急性病毒性肝炎改变；④恢复期血清HBs Ag阴转，抗-HBs IgM阳转。慢性乙肝：血清乙肝病毒e抗原（HBe Ag）阳性或可检出HBV DNA，并排除其他原因引起的丙氨酸氨基转移酶（ALT）升高，同时符合下列条件之一：①HBs Ag阳性时间超过6个月，或者HBs Ag阳性时间持续时间不详，抗-HBc Ig M阴性；②血清ALT反复或者持续升高，或者肝脏病理学有慢性乙肝的特点。未分型：不能明确诊断为急性乙肝和慢性乙肝的病例[8]。

1.3 ARIMA模型的概述及其建立的步骤

1.3.1 模型的概述 ARIMA模型是Box0jekins方法中的重要时间序列分析预测模型，又称为自回归滑动平均混合模型[9]，时间序列模型是可以用数学模型表现出来的一组依赖于时间变化的随机变量，它是假设特征统计量不会随时间的改变而改变，因此该模型要求统计的数据是具有平稳性。ARIMA（p，d，q）（P，D，Q）s模型中p为自回归阶数，d为平稳化过程中趋势差分阶数，q为移动平均阶数，P为季节性自回归阶数，D为季节性趋势差分阶数，Q为季节性移动平均阶数，s为周期[9-10]。

1.3.2 模型建立的步骤 模型的建立需要四步：数据的预处理，模型的识别与参数估计，模型的确定，最后进行预测。建立数据的原始序列图，用单位根（ADF）检验再结合自相关系数函数图（ACF）和偏相关系数函数图（PACF）分析判断其平稳性，如果数据是非平稳性的可以进行对数转换、差分或季节性差分后再用检验其平稳性。做数据一阶差分后的自相关（ACF）、偏自相关（PACF）图，并根据图形的截尾情况对模型做一个预判断。对模型可能的组合进行建模，依据AIC准则、R^2的值还有模型的实际拟合效果，选择最佳的预测模型。最后对选定的模型进行残差白噪声检验一般采用Ljung - Box统计量，$P>0.05$说明模型残差为白噪声序列，所选模型合理，若残差为非白噪声序列则说明残差序列还有信息未提取，需重新建立模型，直到残差序列为白噪声序列为止[11-12]。

1.4 数据分析

采用Excel 2013软件对疫情数据进行统计并绘制相关图表，使用SPSS 20.0和Eviews 10.0软件进行统计分析以及ARIMA模型预测，用MAPGIS绘制传染病专题图形。

2 结果

2.1 乙肝流行的一般情况

2009-2019年乙肝整体发病率呈缓慢下降的趋势，从2009年的71.16/10万到2019年的67.09/10万，整体下降4.07/10万，下降趋势较为缓慢，见图1。

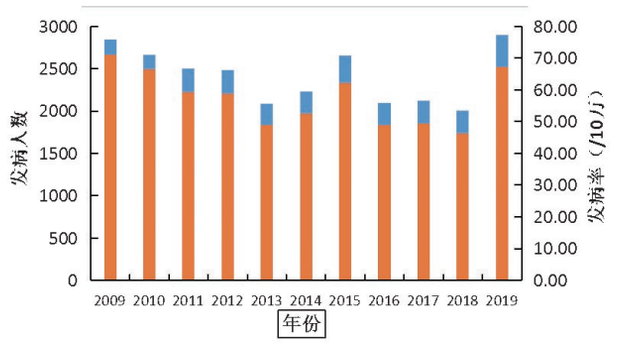


图1 2009-2019年泸州市乙肝发病率随时间变化趋势图

2.1.1 地区分布 泸州市地区乙肝发病率前三位分别为江阳区、泸县和龙马潭区,2009-2019年的平均发病率分别为87.29/10万、76.43/10万、70.17/10万,见图2。



图2 2009-2019年泸州市各区县乙肝发病分布图

2.1.2 时间分布 2009-2019年乙肝病例每个月都有发生,每一年于1月和2月达到最低值,七、八月份达到最高后又开始下降,11月后开始回升,具有一定的季节性但并不是很明显,见图3、图4。

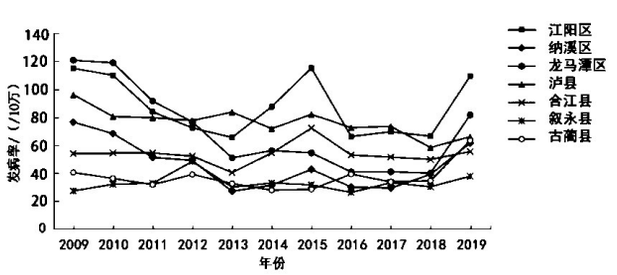


图3 2009-2019年泸州市各区县乙肝发病率随时间变化图

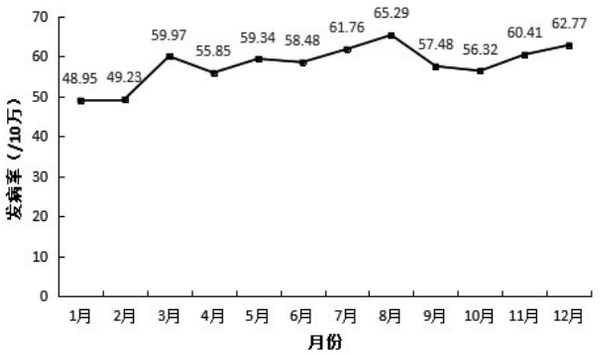


图4 2009-2019年泸州市乙肝发病率月份分布图

2.1.3 人群分布

2.1.3.1 性别分布 26 600例乙肝患者中有17 372例男性患者,女性患者9 228例,男女比例为1.88:1。2009年到2018年男性、女性发病率总体呈下降趋势,但是2019年有回升的现象,见图5。

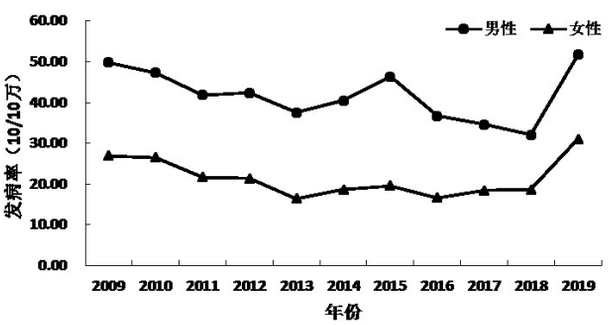


图5 2009-2019年泸州市不同性别乙肝发病率

2.1.3.2 职业分布 农民是发病人数最多的,占了所有职业的62.5%,公共场所服务员是发病人数最少的仅占0.05%,见表1。

表1 2009-2019年泸州市乙肝发病人群的职业分布

职业分类	发病总人数(*n*)	构成比(%)
保育员及保姆	1 619	3.17
不详	1 136	2.22
餐饮食品业	99	0.19
干部职员	585	1.15
工人	1 087	2.13
公共场所服务员	25	0.05
海员及长途驾驶员	169	0.33
家务及待业	9 043	17.71
教师	420	0.82
离退人员	1 606	3.15
民工	445	0.87
农民	31 913	62.50
其它	893	1.75
散居儿童	94	0.18
商业服务	720	1.41
学生	984	1.93
医务人员	180	0.35
幼托儿童	46	0.09

2.2 时间序列预测模型分析

2.2.1 数据的预处理 将数据进行一阶差分和一阶季节性差分后得到了一个平稳的时间序列图，平方根检验（ADF）得到$t=16.80$，$P<0.001$，拒绝原假设，可认为该数据平稳，进行了一次差分和一次季节性差分故d=D=1，故公式可表示为ARIMA（p，1，q）（P，1，Q）$_{12}$，见图6。

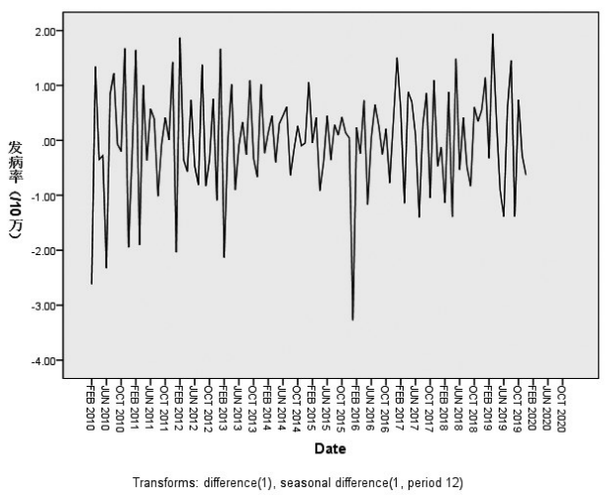


图6 2009–2019年泸州市乙肝月发病率一阶差分和一阶季节性差分后的序列图

2.2.2 模型参数的估计 由表2得自相关图（ACF）、偏自相关图（PACF）下降缓慢，呈现出拖尾的性质，并且自相关图延迟一阶后数值落到了可信区间的范围内、偏自相关图在延迟二阶后数值落到了可信区间的范围内，故可以选择ARIMA（1，1，1）、ARIMA（1，1，2）、ARIMA（2，1，1）、ARIMA（2，2，2）等四个模型进行拟合。

表2 序列经一阶差分和一阶季节性差分后的ACF和PACF图

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.319	-0.319	13.638	0.000
		2	-0.200	-0.336	19.056	0.000
		3	0.136	-0.071	21.571	0.000
		4	0.100	0.082	22.938	0.000
		5	-0.001	0.138	22.938	0.000
		6	-0.214	-0.143	29.351	0.000
		7	-0.068	-0.274	29.992	0.000
		8	0.147	-0.137	33.036	0.000
		9	0.063	0.080	33.611	0.000
		10	-0.138	0.066	36.372	0.000
		11	-0.095	-0.099	37.685	0.000
		12	0.269	0.086	48.247	0.000
		13	-0.012	0.027	48.268	0.000
		14	-0.128	0.001	50.715	0.000
		15	0.028	0.028	50.829	0.000
		16	0.077	0.038	51.731	0.000

2.2.3 模型的建立 并且根据AIC准则和R^2的值[13]即AIC值越小，R^2最大拟合的模型越接近于理论分布。四个预测值的拟合优度结果见表3，可以发现ARIMA（1，1，1）拟合效果最好。经过了一次差分和一次季节性差分故d=D=1，对于季节乘积型ARIMA模型，P、Q一般不超过2阶，可分别取0、1、2逐个检验，最后取BIC值最小和R^2最大以及模型拟合情况进行综合判断[14–15]。最终确定ARIMA（1，1，1）（1，1，2）为最优模型，正态化的BIC为7.164，固定的R^2为0.419。对模型的残差进行白噪声检验，Ljung0Box Q检验显示P=0.849，检验结果不拒绝残差不相关的零假设（$P>0.05$），即判断ARIMA（1，1，1）（1，1，2）$_{12}$模型的残差序列是白噪声序列，所选模型合理。

表3 备选模型的拟合优度统计量

模 型	R^2	调整R^2	AIC	D0W	P
ARIMA（1，1，1）	0.43	0.41	2.41	1.86	<0.001
ARIMA（2，1，1）	0.31	0.30	2.59	2.58	0.59
ARIMA（1，1，2）$_{12}$	0.32	0.31	2.56	2.71	<0.001
ARIMA（2，1，2）	0.25	0.24	2.71	1.28	<0.001

2.2.4 评估预测模型的结果 模型的预测值拟合2018年和2019年的月发病率，并以此判断的模型的拟合实际拟合效果。由表4可得出预测值和实际值的拟合效果较好，相对误差最小值为0.73%，接近真实值，最大的相对误差在27.01%，平均相对误差为10.09%。以此模型为基础得到乙肝2019–2020年的乙肝发病预测图，见图7，2019年1–12月的预测值与实际值拟合效果较好。对2020年1月–12月预测结果显示2020年1–6月整体发病率较低，与实际较为一致，但7–12月发病率还是维持在一个较高的水平，其发病率分别为6.09/10万、6.31/10万、6.23/10万、6.29/10万、7.07/10万、7.78/10万。

3 讨论

3.1 流行病学特征

乙肝一直是急性传染病的防控的重点，其发病率自2014年有上升的趋势，2018年到2019年发病率的上升速度较快，从46.37/10万上涨到了67.09/10万，这可能与泸州市疾控中心对传染病报告卡报告质量要求不一致，也不排除2019年乙肝发病率“实高”的可能。乙肝的描述性流行病学显示：乙肝没有明显的季节高峰，但存在一定的季节差异，1、2月份较低，这可能是因为1、2月正值春季，乙肝的监测量减少所致，而三月份开始增多可能是春季过后人口流动量增多和人们就医增加有关[16]。男性发病率比女性高，大多数研究认为可能与男性社交活动频繁，吸烟饮酒等导致自身肝脏代偿功能较差等有关[17]。还有研究认为由于性激素的差异导致女性的免疫能力高于男性，男性更容易受到病毒的感染与侵害[18-19]，

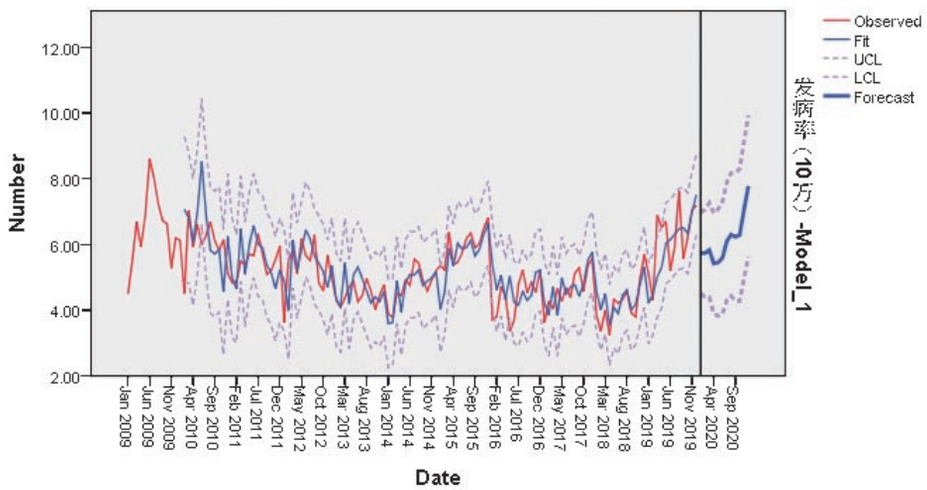


图7 2019-2020年乙肝发病预测图

表4 2018-2019年预测值拟合效果

月份	2019年				2018年			
	实际值(/10万)	预测值(/10万)	实际误差	相对误差(%)	实际值(/10万)	预测值(/10万)	实际误差	相对误差(%)
01	5.11	4.22	0.89	17.43	3.85	4.57	0.19	18.85
02	4.30	4.48	0.18	4.15	3.36	4.01	0.19	19.39
03	6.89	5.03	1.86	27.01	4.01	4.51	0.13	12.55
04	6.52	5.29	1.23	18.89	3.24	3.55	0.09	9.47
05	6.71	6.04	0.67	9.94	4.33	4.09	0.06	5.58
06	5.20	6.15	0.95	18.19	4.22	3.9	0.07	7.49
07	5.92	6.31	0.39	6.58	4.33	4.47	0.03	3.20
08	7.63	6.47	1.16	15.22	4.59	4.62	0.01	0.73
09	5.57	6.51	0.94	16.80	3.91	4.02	0.03	2.69
10	6.20	6.35	0.15	2.45	3.80	4.18	0.10	10.04
11	7.05	6.97	0.08	1.19	4.93	4.78	0.03	3.12
12	7.19	7.52	0.33	4.55	5.70	5.31	0.07	6.81

农民发病率最高是由于我国农民人口基数大,卫生意识较差,人口流动较大等原因导致[20-21]。

地理信息系统可以将空间信息进行提取、分析和输出,并可以将事物进行视觉处理,使研究结果更加清楚[22-23]。研究结果显示江阳区、泸县、龙马潭为乙肝的高发地也是泸州的主要经济带,江阳区、龙马潭区和泸县的人均GDP为全市的前三位。江阳区发病率一直都居高不下,该区县人口密集,其占地面积为650.57平方公里,常住人口为947.63万人,人口密度为1 043.98/万人[24],仅次于龙马潭区,增加了接触性传播的感染风险。

3.2 ARIMA模型预测结果

许多研究显示,ARIMA模型用于乙肝的短时间发病预测精度高,是许多学者常用的预测模型之一,对于传染病的防控起到了重要的作用[25-26]。传染病具有可预测、可预防、可控制的特点。掌握其发展的规律、发病的特点以及影响因素,就可以采取一定的措施进行干预控制,可以为卫生资源的配置提供理论支持。本研究通过反复调试最终确定ARIMA(1,1,1)(1,1,2)$_{12}$为预测泸州市短期预测乙肝发病的最佳模型。传染病的发病受很多因素的影响,如上报的标准的变化,检测量的变化,卫生服务的可及性等,故该模型仅提供一个参考值,并且应该逐年加入新的实际值以调整模型,提高预测精度。

4 结论

泸州市近十年乙肝发病下降趋势较为缓慢,无明显季节高峰,发病有地区差异,发病率最高为江阳区。ARIMA(1,1,1)(1,1,2)$_{12}$模型是泸州市短期预测乙肝发病的最佳模型。

参 考 文 献

[1] 李军，姬艳芳，杨凯朝，等. 2004-2019年河南省乙型肝炎疫情报告及流行特征分析[J]. 疾病监测，2020，35(8): 766-769.

[2] 闫永平，张维璐，苏海霞，等. 我国乙型病毒性肝炎防治研究新进展和面临的挑战[J]. 中国热带医学，2019，19(10): 916-921.

[3] Shan S，You H，Niu JQ，*et al*. Baseline characteristics and treatment patterns of the patients recruited to the China registry of hepatitis B[J]. J Clin Transl Hepatol，2019，7(4): 322-328.

[4] 世界卫生组织. 2016-2021年全球卫生部门病毒性肝炎战略[EB/OL].https://apps.who.int/iris/bitstream/handle/10665/246177/WHO-HIV-2016.06-chi.pdf.

[5] 杨艳红.重庆市病毒性肝炎流行牲及预测模型的研究[D]. 重庆医科大学,2013.

[6] 王平贵，张丽杰，张国民，等. 2005-2016年甘肃省急性乙型肝炎流行特征分析[J]. 疾病监测，2019，34(3): 251-253.

[7] 于智欣. 基于社区的乙肝流行病学仿真建模及预测研究[D]. 杭州: 浙江大学，2019.

[8] 李言飞，张业武，王晓风，等. 2005-2017年全国法定传染病重复报告卡大数据分析与应用[J]. 疾病监测，2019，34(5): 468-472.

[9] 沈忠周，王雅文，马帅，等. ARIMA模型在预测我国乙肝月报告发病数中的应用[J]. 中华疾病控制杂志，2018，22(3): 298-301，307.

[10] 李凤英，马婷，罗文瑞，等. 2005-2017年陕西省咸阳市乙型肝炎流行趋势分析及预测模型的建立[J]. 疾病监测，2019，34(2): 137-140.

[11] 田娟，罗兴静，冯军，等. 贵州省2004-2016年乙型病毒性肝炎发病时空分布特征[J]. 中国公共卫生，2019，35(5): 607-610.

[12] 王彦文. 2010-2018年桂林市手足口病流行特征及发病趋势预测研究[D]. 南宁: 广西医科大学，2019.

[13] 王平. 三种预测模型在主要传染病发病率预测中的应用[D]. 杭州: 浙江大学，2010.

[14] 覃柳麻. 2012-2017年南宁市乙肝流行特征及空间分析[D]. 南宁: 广西医科大学，2019.

[15] 杨燕妮，陶沁，田娟，等. 时间序列分析法在贵州省乙型病毒性肝炎流行趋势预测中的应用[J]. 职业与健康，2018，34(15): 2102-2105.

[16] 刘意. 湖南省2006-2014年戊肝时空特征分析及预测研究[D]. 长沙: 湖南师范大学，2017.

[17] 杨晓丽. 2007-2016年辽宁省乙型病毒性肝炎流行特征及预测[D]. 长春: 吉林大学，2019.

[18] 袁磊，张水军. 性别与性激素对乙肝病毒易感性及其免疫应答的影响[J]. 河南医学研究，2019，28(10): 1919-1921.

[19] 周雪宁，万延川，权志博. 性别对乙肝血清标志物模式人群分布率的影响及其探讨[J]. 实用医技杂志，2008，15(14): 1802-1803.

[20] 许长沙，杭栋，周新亚，等. 2004—2018年宿迁市乙型病毒性肝炎流行病学分析[J]. 江苏预防医学，2019，30(6): 652-654.

[21] 周骥，张文华，张静敏，等. 2012-2016年乐山市乙肝流行病学分析[J]. 预防医学情报杂志，2018，34(4): 503-506.

[22] Kauhl B，Heil J，Hoebe CJPA，*et al*. The spatial distribution of hepatitis C virus infections and associated determinants--an application of a geographically weighted poisson regression for evidence-based screening interventions in hotspots[J]. PLoS One，2015，10(9): e0135656.

[23] 赖铿，吴桂锋，雷宇，等. 2014-2018年广州市肺结核发病空间分布特征分析[J]. 中国防痨杂志，2020，42(4): 327-334.

[24] 范玉梅,王小娟,蒋德勇. 2014-2018年泸州市乙型肝炎流行病学特征分析[J].实用预防医学,2020,27(03):346-348.

[25] 毛向群，熊小庆，涂秋凤，等. 江西省乙型肝炎发病趋势的时间序列和预测模型分析[J]. 中国预防医学杂志，2013，14(6): 435-438.

[26] 黄淑琼，陈红缨，蔡晶，等. 应用ARIMA模型分析湖北省乙型病毒性肝炎时间分布特征[J]. 公共卫生与预防医学，2016，27(4): 14-18.

（2021-01-24 收稿）