

基于组合模型的太原市丙型肝炎发病趋势预测研究

郭思瑶^{1,2,3,4} 赵启玉^{1,2,3,4,5} 张越^{2,3,4} 张萍^{2,3,4} 车晓文⁶ 郑金鸽⁶ 王蕾⁶

¹山西医科大学医学科学院,太原 030001;²煤炭环境致病与防治教育部重点实验室,太原 030001;³山西医科大学公共卫生学院流行病学教研室,太原 030001;⁴山西医科大学环境污染与重大慢病流行病学研究中心,太原 030001;⁵太原市疾病预防控制中心放射卫生科,太原 030012;⁶太原市疾病预防控制中心艾滋病防制科,太原 030012

通信作者:赵启玉,Email:maradona164@163.com

【摘要】目的 根据差分自回归移动平均(ARIMA)模型、反向传播神经网络(BPNN)、ARIMA-BPNN模型,针对太原市丙型肝炎(丙肝)发病数据特征选择适合发病趋势预测的最优模型。**方法** 选取2008–2021年现住址为太原市的丙肝报告病例数据,使用季节性趋势分解图分析期间太原市丙肝月发病率的季节性特征,并建立ARIMA模型、BPNN模型、ARIMA-BPNN模型进行预测。采用平均绝对误差(MAE)、均方误差(MSE)、均方根误差(RMSE)以及平均绝对百分比误差(MAPE)指标衡量模型的性能。**结果** 累计报告丙肝病例20 025例,总体发病趋势稳定。BPNN模型在MSE、MAE和RMSE指标上表现较好,ARIMA-BPNN模型在MAPE指标上表现较好,ARIMA模型表现相对一般。**结论** ARIMA-BPNN模型是预测太原市丙肝发病趋势的较优模型,预测性能高于单一模型,在传染病发病趋势预测中有重大前景。

【关键词】 丙型肝炎; 差分自回归移动平均模型; 反向传播神经网络; 组合模型; 发病趋势

基金项目:山西省青年科学研究基金(201901D211326,202103021223216);山西省高等教育“百亿工程”科技引导专项(BYBLD002);山西省太原市医学重点学科(2024-2046)

Research on the prediction of Hepatitis C incidence trend in Taiyuan City based on combination model

Guo Siyao^{1,2,3,4}, Zhao Qiyu^{1,2,3,4,5}, Zhang Yue^{2,3,4}, Zhang Ping^{2,3,4}, Che Xiaowen⁶, Zheng Jing⁶, Wang Lei⁶

¹School of Medical Sciences, Shanxi Medical University, Taiyuan 030001, China; ²Key Laboratory of Coal Environmental Pathogenesis and Prevention, Ministry of Education, Taiyuan 030001, China; ³Department of Epidemiology, School of Public Health, Shanxi Medical University, Taiyuan 030001, China; ⁴Research Center of Environmental Pollution and Major Chronic Diseases Epidemiology, Shanxi Medical University, Taiyuan 030001, China; ⁵Department of Radiation Health, Taiyuan Center for Disease Control and Prevention, Taiyuan 030012, China; ⁶Department of AIDS Protection and Control, Taiyuan Center for Disease Control and Prevention, Taiyuan 030012, China

Corresponding author: Zhao Qiyu, Email: maradona164@163.com

【Abstract】Objective Based on the autoregressive integrated moving average (ARIMA) model, back propagation neural network (BPNN), and ARIMA-BPNN model, select the optimal model suitable for predicting the incidence trend of hepatitis C in Taiyuan City according to the

DOI:10.3760/cma.j.cn112338-20240814-00502

收稿日期 2024-08-14 本文编辑 张婧

引用格式:郭思瑶,赵启玉,张越,等.基于组合模型的太原市丙型肝炎发病趋势预测研究[J].中华流行病学杂志,2025,46(2):204-209. DOI:10.3760/cma.j.cn112338-20240814-00502.

Guo SY, Zhao QY, Zhang Y, et al. Research on the prediction of Hepatitis C incidence trend in Taiyuan City based on combination model[J]. Chin J Epidemiol, 2025, 46(2):204-209. DOI: 10.3760/cma.j.cn112338-20240814-00502.



characteristics of the data. **Methods** The data of reported cases of hepatitis C in Taiyuan from 2008 to 2021 were selected, and the seasonal trend decomposition chart was used to analyze the seasonal characteristics of the monthly incidence rate of hepatitis C in Taiyuan during the period, and the ARIMA model, BPNN model, and ARIMA-BPNN model were established to predict. The performance of the model was measured using four indicators: mean absolute error (MAE), mean squared error (MSE), root mean square error (RMSE), and mean absolute percentage error (MAPE). **Results** A total of 20 025 cases of hepatitis C were reported, and the overall incidence trend was stable. The BPNN model performed well on MSE, MAE, and RMSE indicators, the ARIMA-BPNN model performed well on MAPE indicators, and the ARIMA model performed relatively averagely. **Conclusions** The ARIMA-BPNN model is a better model for predicting the trend of hepatitis C in Taiyuan City, with a higher predictive performance than a single model. It has significant prospects in predicting the trend of infectious diseases.

【Key words】 Hepatitis C; Autoregressive integrated moving average model; Back propagation neural network; Combination model; Incidence trend

Fund programs: Youth Science Research Fund of Shanxi Province (201901D211326, 202103021223216); Shanxi Province Higher Education "Billion Project" Science and Technology Guidance Special Project (BYBLD002); Medical Key Discipline Funding Project of Taiyuan City, Shanxi Province (2024-2046)

丙型肝炎(丙肝)是一种慢性传染性疾病,在世界范围内流行,不同性别、年龄、种族人群均易感^[1]。我国人口基数庞大,是全球丙肝感染人数较多的国家,目前仍是我国重大公共卫生问题^[2]。山西省是我国丙肝发病热点区域之一,其中太原市报告病例数在山西省位居第二^[3]。丙肝隐匿性较强,医疗机构病例报告尚无法充分反映疾病的实际流行负担。面对这一严峻的公共卫生挑战,准确预测丙肝发病趋势对于制定预防和控制措施至关重要^[4]。丙肝发病率数据呈现线性趋势和非线性波动的特征,差分自回归移动平均(ARIMA)模型能够有效处理数据中的线性趋势和季节性变化,反向传播神经网络(BPNN)模型可以捕捉复杂非线性动态,ARIMA-BPNN模型综合两者的优势,实现丙肝传染的及时预警,推动太原市丙肝消除目标的实现。

资料与方法

1. 数据来源:来源于全民健康保障疾控信息系统报告的2008年1月1日至2021年12月31日现住址为太原市的丙肝病例数据,删除重报病例,筛选保留“实验室确诊病例”和“临床诊断病例”。病例的诊断依据《丙型肝炎病毒学诊断标准(WS213-2008)》^[5]和《丙型肝炎病毒学诊断标准(WS213-2018)》^[6]进行。报告数据经过相关政府部门的审查和质量控制,内容可靠。

2. 原理与方法:

(1)ARIMA模型:表示为 $ARIMA(p,d,q)(P,D,Q)_s$,

其中, d,D 为差分的阶数和步长, p 和 P 为自回归的阶数, q 和 Q 为移动平均的阶数^[7]。将季节性差分与自回归(AR)、差分(I)和移动平均(MA)组件结合在一起,通过调整模型参数以适应各种类型的时间序列数据,有效地预测未来值。

①数据预处理:将2008年1月至2020年12月的丙肝月发病率数据作为训练集建立模型,2021年数据作为验证集,比较实际值与预测值相差大小。在模型预测中,首先通过绘制时序图以及Augmented Dickey-Fuller检验判断原始数据是否为平稳时间序列($P<0.05$ 为平稳序列),若为不平稳序列需通过差分转化为平稳时间序列^[8]。

②模型识别与定阶:通过观察自相关函数(ACF)图和偏自相关函数(PACF)图选择合适的 p 、 q 以及 P 、 Q 值,根据差分次数确定差分的阶数和步长 d,D 。当自相关系数显著不为0时,意味着序列的当前值与过去值之间存在相关性。在ACF图和PACF图中添加基于95%CI的上下限线,若某滞后阶数的自相关系数或偏自相关系数超出该区间,表明在该滞后阶数下存在显著的相关性。当自相关滞后阶数 q 阶拖尾、偏自相关滞后阶数 p 阶截尾时,应选择 $AR(p)$ 模型;当 q 阶截尾 p 阶拖尾时,应选择 $MA(q)$ 模型;当 p 阶、 q 阶均为截尾时,应选择 $ARMA(p,q)$ 模型。赤池信息准则(AIC)、贝叶斯信息准则(BIC)均可用于平衡模型的拟合优度和复杂度,AIC适用于预测精度,BIC适用于模型解释问题^[9]。本研究利用AIC最小值来估计模型参数,并选择相对最优模型,AIC越小,模型拟合数据的能力相对

越好^[7]。

③参数估计与模型诊断:使用 Ljung-Box 检验来评估模型拟合的充分性,根据建立的模型进行未来值的预测。

(2)BPNN 模型:为典型的反向前馈神经网络,结构包括输入层、隐藏层和输出层,其中每层包含多个神经元,隐藏层位于输入层和输出层之间,起到连接和转换的作用^[10]。通过某种形式将输出误差反向传播到隐藏层和输入层,调整隐藏层中每个神经元的权值逐渐减小误差,直到实际输出与期望输出之间的误差满足精度要求或达到最大学习系数^[11]。

①数据处理:在建立模型之前将所有输入特征在训练前进行了 Min-Max 归一化,原始发病率数据范围可能会差异太大,导致模型的灵敏性下降,降低训练速率,其次数据中异常值可能会导致网络不能收敛,数据做归一化变化,对原始数据进行线性变换,使结果值映射到 0~1 之间,可以避免因不同特征的数值范围差异过大而引起的训练不稳定问题,进而提升模型的稳定性和收敛速度。公式:

$$\min\max_x(x) = \frac{x - \min(x)}{\max(x) - \min(x)}$$

进行网络结构设计时采用了季节性分解方法,这种方法假设时间序列具有周期性,通过历史数据来预测未来同一时期的值,使用前几年同一时期的数据作为输入,以预测之后同一时期的数据。适用于处理本研究以年为周期的时间序列数据,输入样本选取前 12 个月的发病率数据,输出样本选取下一个月的发病率数据,可以得知输入节点数为 12,输出节点数为 1。例如输入 2008 年 1-12 月的发病率数据,用其预测 2009 年 1 月的发病率,输入 2008 年 2 月至 2009 年 1 月的发病率数据,用其预测 2009 年 2 月发病率,以此类推。

②网络结构设计:确定隐藏层节点数,对输出层通常使用线性激活函数,并确定学习速率、训练精度、迭代次数,根据学习速率的特点,其取值范围为 0.01~0.80;隐藏层计算公式: $m = \sqrt{N + M} + a$ 。式中, N 为输入节点数, M 为输出节点数, a 为取值范围 1~10 的自然数。

③网络训练:使用确定的网络结构进行模型建立,最终选择均方误差(MSE)最小的网络结构为最优模型。

(3)ARIMA-BPNN 模型:为一种将 ARIMA 模型和 BPNN 模型进行结合的预测技术。首先利用实

际值进行 ARIMA 建模,并对 2008-2021 年数据进行拟合,将 ARIMA 模型的预测值作为 BPNN 模型的输入值,原始数据作为输出值。把 2008 年 1 月至 2020 年 12 月拟合值、实际值数据作为训练集,2021 年数据作为测试集。建立模型之前,为了避免数据过度收敛或过度拟合,同样先将数据进行归一化处理。隐含层节点数、学习速率、训练精度、迭代次数与 BPNN 建模方法相同。

3. 统计学分析:采用 Excel 2021 软件建立数据库并进行原始数据预处理,使用 R 4.3.2 软件进行 3 种模型建立。通过评估 3 种模型的平均绝对误差(MAE)、MSE、均方根误差(RMSE)以及平均绝对百分比误差(MAPE),全面评价模型预测能力。结果数值越低,表明模型的预测准确度越高。双侧检验,检验水准 $\alpha=0.05$ 。

结 果

1. 描述性分析:2008-2021 年太原市累计报告 20 025 例丙肝病例,其间每月均有病例散发,无明显明显的季节性。年平均增长速度为 -0.31%,总体呈略微下降趋势($P=0.023$)。见图 1。

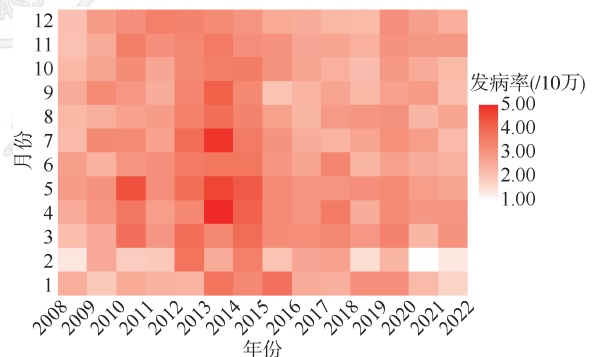


图 1 2008-2021 年太原市丙型肝炎月发病率热力分布图

2. ARIMA 模型:以原始数据建立时间序列图可知该序列呈长期增长趋势,通过季节性趋势分解将时间序列分解为长期趋势、季节趋势和随机效应。季节性变化以年为周期,属于非平稳序列,进行 1 阶 12 步差分后该序列类似平稳,且 Augmented Dickey-Fuller 检验结果为差分平稳,不具有白噪声序列的特征(Dickey-Fuller=-7.55, Lag order=5, $P = 0.010$),延迟 6 阶和 12 阶的 Ljung-Box 检验结果显示差异均有统计学意义($\chi^2=36.95, df=6, P<0.001$; $\chi^2=76.87, df=12, P<0.001$)。

通过绘制 ACF 图和 PACF 图(图 2),可观察到 ACF 图呈截尾状态, PACF 图呈现拖尾状态,可以确定模型的 $p=0, d=1, q=1$, 初步确定模型为 ARIMA $(0, 1, 1)(P, 1, Q)_{12}$ 。有研究表明,参数 P, Q 的取值范围通常不超过 $2^{[12]}$ 。结合以上信息确定候选模型进行拟合优度检验,根据 AIC 最小原则确定 ARIMA $(0, 1, 1)(0, 1, 1)_{12}$ 为相对最优模型(表 1), Ljung-Box 检验 $P=0.639$, 表明数据特征被完全提取,此模型为高斯白噪声序列。利用最优模型对 2021 年太原市丙肝发病率进行预测,将预测值与真实值进行比较,发现存在一定的误差,但每个月预测值均在 95%CI 内。见表 2,图 3。

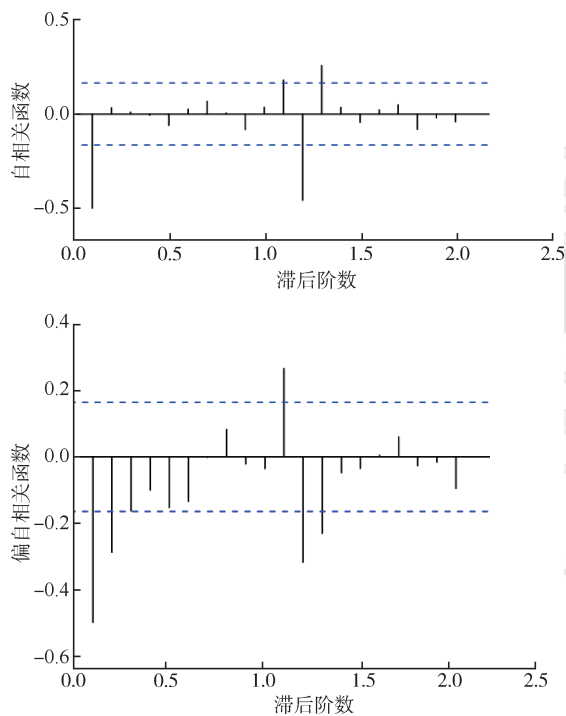


图 2 差分序列自相关图与偏自相关图

表 1 候选模型评估

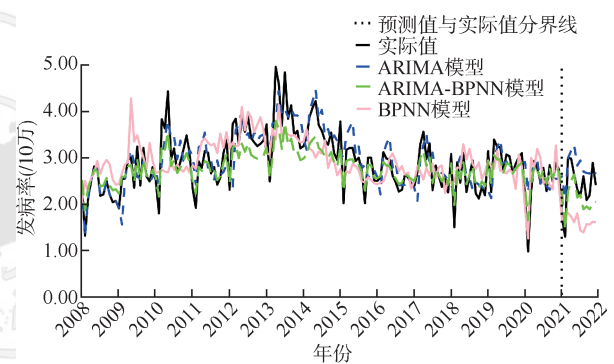
备选模型	赤池信息准则	贝叶斯信息准则
$(0, 1, 1)(0, 1, 0)_{12}$	274.82	280.74
$(0, 1, 1)(0, 1, 1)_{12}$	200.50	209.39
$(0, 1, 1)(0, 1, 2)_{12}$	201.89	213.74
$(0, 1, 1)(1, 1, 0)_{12}$	241.07	249.96
$(0, 1, 1)(1, 1, 1)_{12}$	201.98	213.83
$(0, 1, 1)(1, 1, 2)_{12}$	203.60	218.41
$(0, 1, 1)(2, 1, 0)_{12}$	222.65	234.50
$(0, 1, 1)(2, 1, 1)_{12}$	203.25	218.06
$(0, 1, 1)(2, 1, 2)_{12}$	201.50	219.33

3. BPNN 模型:本研究的隐含层节点范围在 5~14,学习速率、训练精度、迭代次数分别设置为

表 2 2021 年太原市丙型肝炎发病率实际值与模型预测值比较

月份	实际值	ARIMA 模型预测值	BPNN 模型预测值	ARIMA-BPNN 模型预测值
1	1.68	2.58	1.64	2.22
2	1.28	2.10	1.71	1.45
3	2.98	2.97	1.79	2.25
4	2.96	3.13	1.71	2.49
5	2.62	3.26	1.59	2.62
6	2.30	2.81	1.78	2.12
7	2.15	2.92	1.41	2.18
8	2.58	2.68	1.37	1.87
9	2.07	2.65	1.55	1.93
10	2.17	2.64	1.54	1.87
11	2.87	2.74	1.60	1.99
12	2.39	2.63	1.59	2.03

注:ARIMA:差分自回归移动平均;BPNN:反向传播神经网络



注:ARIMA:差分自回归移动平均;BPNN:反向传播神经网络

图 3 3 种模型预测报告发病率的比较

0.01、0.1、100 000。不同隐藏层节点数对应的 MSE 见表 3。

4. ARIMA-BPNN 模型:将 ARIMA $(0, 1, 1)(0, 1, 1)_{12}$ 模型的预测值作为混合模型的输入变量,以丙肝的原始值作为输出变量。模型隐含层节点数范围在 2~12,学习速率、训练精度、迭代次数分别设置为 0.01、0.01、100 000。通过训练可知 BPNN 模型选择隐含层节点数为 5 时,预测值 MSE 为 0.013 1,小于其他结构,因此确定预测效果最好的网络结构为 12-5-1。ARIMA-BPNN 模型选择隐含层节点数为 5 时, MSE 为 0.098 2,小于其他结构,因此确定预测效果最好的网络结构为 1-5-1。见表 3。对训练出的数据进行反归一化处理获得最终预测值。

5. 模型性能评估:从预测指标对比结果来看,混合模型在 4 个预测性能指标上均优于单一模型,预测精度较高。从预测值与实际值对比图来看

表3 BPNN模型和ARIMA-BPNN模型中不同隐藏层节点数对应的MSE

隐含层节点数	MSE	隐含层节点数	MSE
BPNN		ARIMA-BPNN	
5	0.013 1	2	0.100 2
6	0.015 2	3	0.099 1
7	0.016 1	4	0.099 4
8	0.016 1	5	0.098 2
9	0.015 8	6	0.098 6
10	0.013 3	7	0.100 1
11	0.015 7	8	0.098 5
12	0.014 3	9	0.100 2
13	0.014 6	10	0.099 2
14	0.025 3	11	0.099 2

注:BPNN:反向传播神经网络;ARIMA:差分自回归移动平均;MSE:均方误差

(图3),ARIMA模型在MSE、MAE和RMSE指标上表现较一般,但在MAPE指标上表现相对较好(22.87%)。BPNN模型在MSE、MAE和RMSE指标上均表现优于ARIMA模型。然而,BPNN模型的MAPE为25.49%,高于ARIMA模型。ARIMA-BPNN模型在所有指标上均表现较好,特别是MAPE仅为15.79%,显著低于单一模型。见表4。

表4 3种模型预测性能的比较

模型	预测性能			
	MSE	MAE	RMSE	MAPE(%)
ARIMA	0.29	0.45	0.53	22.87
BPNN	0.21	0.38	0.46	25.49
ARIMA-BPNN	0.22	0.38	0.47	15.79

注:MSE:均方误差;MAE:平均绝对误差;RMSE:均方根误差;MAPE:平均绝对百分比误差;ARIMA:差分自回归移动平均;BPNN:反向传播神经网络

讨 论

本研究结果显示,太原市丙肝报告发病率在2013年前逐步上升,与中国31个省(自治区、直辖市)以及山西省丙肝整体变化趋势走向一致^[13],可能由于近年来我国不断提高对丙肝的重视程度,加之传染病网络报告系统逐步完善,丙肝患者检出率提升,漏报率逐步下降^[14]。太原市自2014年起丙肝发病率呈下降趋势,可能由于中国CDC性病艾滋病预防控制中心于2013-2015年连续3年开展了全国范围内丙肝病例报告抽查、核查,太原市积极响应,核查结果连续性和可比较性较好,病例报告质量不断规范^[15]。

ARIMA模型^[16-17]、BPNN模型^[18-19]在传染病发

病趋势预测中均被广泛使用,同时近年来有许多学者使用将二者相结合的混合模型对传染病流行情况进行预测。本研究应用ARIMA模型、BPNN模型、ARIMA-BPNN模型分别对太原市丙肝发病趋势进行预测,结果显示,3种模型均可反映未来发病趋势,但预测特征仍存在差异。其中,ARIMA模型主要用于预测呈现线性趋势和季节性模式的时间序列数据^[20]。本研究中时间序列数据经过模型识别、估计和检验,最终确定相对最优的模型为ARIMA(0,1,1)(0,1,1)₁₂,预测趋势与实际大致相符,但仍存在一定误差,可能是由于差分导致部分数据丢失,长期预测中需要不断纳入新的监测数据进行更新,在高发病率的传染病预测中可能无法达到最佳效果^[21],预测效果有限。BPNN模型具备建立复杂的非线性映射关系能力,其网络结构能够适应各种数据模式,通过学习输入和输出之间的复杂关系来实现模式识别和预测^[22],具有强大的学习和泛化能力,与ARIMA模型相比,BPNN模型预测精度更高,容错性更强^[23]。BPNN模型的MSE、MAE、RMSE均低于ARIMA模型,表明该模型在整体误差控制方面表现较好,能够有效减少预测误差。MSE和RMSE对较大的误差点更加敏感,BPNN模型在这些指标上的优越性反映出其在控制较大偏差上的能力更强。此外,MAE的降低也表明BPNN模型在所有数据点上的平均误差较小,整体预测效果更加精确。然而,MAPE在BPNN模型中高于ARIMA模型,MAPE作为无量纲的相对误差指标,能够反映模型在不同发病率水平下的预测误差控制情况。ARIMA模型存在潜在过拟合现象,在预测未来趋势时的可靠性有所下降。

本研究将2种单一模型进行混合,结果显示,ARIMA-BPNN模型的预测指标均优于单一模型,尤其是MAPE大幅降低,说明混合模型能够更好地处理不同数据尺度下的相对误差,从而在实际应用中提供了更加可靠的预测结果。通过结合ARIMA模型对线性趋势的准确捕捉能力和BPNN模型在非线性和波动上的处理优势,表明混合模型在丙肝发病趋势预测中表现更为优越,对数据特征进行了充分提取,与其他学者的研究结果一致^[24]。ARIMA-BPNN模型可充分捕捉时间序列数据的趋势和周期性,能更好、更稳定地适应事件变化。此外,单一模型在数据特征处理时可能会产生不同误差,使用混合模型可以相互抵偿误差,减少单一模型的偏差、方差,提高混合模型的鲁棒性、稳定性以及预测

精度。通过模型评价对比结果可知,组合模型预测能力强于单一模型,更能够接近丙肝真实发病情况。不同发病率水平下的误差控制更加稳定和精准,这为公共卫生部门提供了一个更为可靠的预测工具。在实际防控工作中,精准的发病率预测有助于更早发现潜在的发病高峰,提前部署预防措施,最终推动实现丙肝消除的目标。这些成果对丙肝防控工作具有重要的现实意义,不仅有助于减少疾病的传播和社会负担,还能为长期的疾病管理和防控提供数据支持与指导。

本研究存在局限性。首先,丙肝病例报告数据的质量可能受到医疗资源的影响,预测结果的准确性仍受限于输入数据的质量和范围;其次,丙肝传播还会受到社会经济因素和个人健康指标等因素的影响。未来研究将考虑更广泛的地区数据,并进一步优化模型结构和参数选择,以提高预测的准确性和泛化能力。通过全面理解丙肝在更广泛地区的流行趋势和特点,为未来丙肝流行预测和相关卫生部门的预防控制措施提供科学依据,也为类似地区的传染病趋势预测提供参考。

利益冲突 所有作者声明无利益冲突

作者贡献声明 郭思瑶:数据收集、统计分析、文章撰写;赵启玉:研究指导、文章审阅、经费支持;张越、郑金鹤、王蕾:数据整理;张萍:文章审阅;车晓文:数据收集、文献检索

参 考 文 献

- [1] Yue TT, Zhang QQ, Cai T, et al. Trends in the disease burden of HBV and HCV infection in China from 1990-2019[J]. *Int J Infect Dis*, 2022, 122:476-485. DOI:10.1016/j.ijid.2022.06.017.
- [2] Greenaway C, Ma AT, Kloda LA, et al. The seroprevalence of hepatitis c antibodies in immigrants and refugees from intermediate and high endemic countries: a systematic review and meta-analysis[J]. *PLoS One*, 2015, 10(11):e0141715. DOI:10.1371/journal.pone.0141715.
- [3] 王蕾, 刘海波, 韩红, 等. 2015-2019 年太原市丙肝重复报告分析[J]. *预防医学论坛*, 2021, 27(4):296-299. DOI:10.16406/j.pmt.issn.1672-9153.2021.04.018.
- [4] Wang L, Liu HB, Han H, et al. Analysis on duplicated reporting of hepatitis C, Taiyuan City, 2015-2019[J]. *Prev Med Trib*, 2021, 27(4):296-299. DOI:10.16406/j.pmt.issn.1672-9153.2021.04.018.
- [5] Taha G, Ezra L, Abu-Freha N. Hepatitis c elimination: opportunities and challenges in 2023[J]. *Viruses*, 2023, 15(7):1413. DOI:10.3390/v15071413.
- [6] 中华人民共和国卫生部. 丙型肝炎病毒学诊断标准: WS213-2008[S]. 北京:中国标准出版社, 2008. Ministry of Health of the People's Republic of China. Diagnosis for hepatitis C: WS213-2008[S]. Beijing: Standards Press of China, 2008.
- [7] 中华人民共和国国家卫生和计划委员会. 丙型肝炎病毒学诊断标准:WS213-2018[S]. 北京:中国标准出版社, 2018. National Health and Family Planning Commission of the People's Republic of China. Diagnosis for hepatitis C: WS213-2018[S]. Beijing:Standards Press of China, 2018.
- [8] Shoko C, Njuho P. ARIMA model in predicting of COVID-19 epidemic for the southern africa region[J]. *Afr J Infect Dis*, 2023, 17(1):1-9. DOI:10.21010/Ajidv17i1.1.
- [9] Lou HR, Wang X, Gao Y, et al. Comparison of ARIMA model, DNN model and LSTM model in predicting disease burden of occupational pneumoconiosis in Tianjin, China [J]. *BMC Public Health*, 2022, 22(1):2167. DOI:10.1186/s12889-022-14642-3.
- [10] Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection[J]. *Sociol*

Methods Res, 2024, 33(2): 261-304. DOI: 10.1177/0049124104268644.

- [10] An QY, Wu J, Meng J, et al. Using the hybrid EMD-BPNN model to predict the incidence of HIV in Dalian, Liaoning Province, China, 2004-2018[J]. *BMC Infect Dis*, 2022, 22(1):102. DOI:10.1186/s12879-022-07061-7.
- [11] Chen Q, Zhao H, Qiu HF, et al. Time series analysis of rubella incidence in Chongqing, China using SARIMA and BPNN mathematical models[J]. *J Infect Dev Ctries*, 2022, 16(8):1343-1350. DOI:10.3855/jidc.16475.
- [12] 寇晓杰, 张泽旭, 倪进东, 等. 2005-2019 年湛江市主要血源及性传播疾病流行特征分析与模型预测[J]. *中华疾病控制杂志*, 2023, 27(7): 763-768, 806. DOI: 10.16462/j.cnki.zhjbkz.2023.07.004.
- [13] Kou XJ, Zhang ZX, Ni JD, et al. Analysis and model prediction of epidemiological characteristics of main blood-borne and sexually transmitted diseases in Zhanjiang City from 2005 to 2019[J]. *Chin J Dis Control Prev*, 2023, 27(7): 763-768, 806. DOI: 10.16462/j.cnki.zhjbkz.2023.07.004.
- [14] 孙海泉, 肖革新, 鄂莹, 等. 中国 2008-2012 年丙肝流行规律及空间聚集性分析[J]. *中国公共卫生*, 2014, 30(3): 286-289. DOI:10.11847/zgggws2014-30-03-10.
- [15] Sun HQ, Xiao GX, Guo Y, et al. Spatial clustering and changing trend of hepatitis C prevalence during 2008-2012 in China[J]. *Chin J Public Health*, 2014, 30(3): 286-289. DOI:10.11847/zgggws2014-30-03-10.
- [16] 聂晓勇, 张飞, 解雁茹, 等. 2010-2019 年山西省丙型肝炎病毒学流行病学特征分析[J]. *山西医药杂志*, 2020, 49(16): 2117-2118. DOI:10.3969/j.issn.0253-9926.2020.16.010.
- [17] Nie XY, Zhang F, Xie YR, et al. Epidemiological characteristics analysis of hepatitis C in Shanxi Province from 2010 to 2019[J]. *Shanxi Med J*, 2020, 49(16): 2117-2118. DOI:10.3969/j.issn.0253-9926.2020.16.010.
- [18] 黑发欣, 庞琳, 王晓春, 等. 2013-2015 年我国丙型肝炎病毒学肝炎病例报告数据质量核查分析[J]. *疾病监测*, 2017, 32(6): 513-515. DOI:10.3784/j.issn.1003-9961.2017.06.017.
- [19] Hei FX, Pang L, Wang XC, et al. Reporting quality of hepatitis C cases in China, 2013-2015[J]. *Dis Surveill*, 2017, 32(6):513-515. DOI:10.3784/j.issn.1003-9961.2017.06.017.
- [20] 刘双, 柳晓琳. ARIMA 模型在丙型肝炎月发病数预测中的应用[J]. *锦州医科大学学报*, 2021, 42(6):79-83. DOI:10.13847/j.cnki.lnmu.2021.06.016.
- [21] Liu S, Liu XL. Application of ARIMA model in predicting the monthly incidence of hepatitis C[J]. *J Jinzhou Med Univ*, 2021, 42(6):79-83. DOI:10.13847/j.cnki.lnmu.2021.06.016.
- [22] 马晶, 王梅, 谯小伟, 等. ARIMA 季节性模型在预测兰州市丙肝发病人数中的应用[J]. *中国卫生统计*, 2022, 39(1): 98-100, 105. DOI:10.3969/j.issn.1002-3674.2022.01.022.
- [23] Ma J, Wang M, Qiao XW, et al. Application of ARIMA seasonal model in predicting the number of hepatitis C cases in Lanzhou City[J]. *Chin J Health Stat*, 2022, 39(1): 98-100, 105. DOI:10.3969/j.issn.1002-3674.2022.01.022.
- [24] 刘今, 孙进, 孙明丽, 等. 基于组合模型的手足口病发病趋势预测[J]. *中国卫生统计*, 2023, 40(6):892-896. DOI:10.11783/j.issn.1002-3674.2023.06.021.
- [25] Liu J, Sun J, Sun ML, et al. Prediction of the incidence trend of hand, foot and mouth disease based on a combination model[J]. *Chin J Health Stat*, 2023, 40(6): 892-896. DOI:10.11783/j.issn.1002-3674.2023.06.021.
- [26] 刘天, 姚梦雷, 黄继贵, 等. BP 神经网络在传染病时间序列预测中的应用及其 MATLAB 实现[J]. *预防医学情报杂志*, 2019, 35(8):812-816, 821.
- [27] Liu T, Yao ML, Huang JG, et al. Application of back propagation neural network in prediction of infectious disease time series and its MATLAB implementation[J]. *J Prev Med Inf*, 2019, 35(8):812-816, 821.
- [28] Li ZM, Li YN. A comparative study on the prediction of the BP artificial neural network model and the ARIMA model in the incidence of AIDS[J]. *BMC Med Inform Decis Mak*, 2020, 20(1):143. DOI:10.1186/s12911-020-01157-3.
- [29] Swaraj A, Verma K, Kaur A, et al. Implementation of stacking based ARIMA model for prediction of Covid-19 cases in India[J]. *J Biomed Inform*, 2021, 121: 103887. DOI:10.1016/j.jbi.2021.103887.
- [30] Fang K, Cao L, Fu ZW, et al. Prediction of reported monthly incidence of hepatitis B in Hainan Province of China based on SARIMA-BPNN model[J]. *Medicine (Baltimore)*, 2023, 102(41): e35054. DOI: 10.1097/MD.00000000000035054.
- [31] Wang KW, Deng C, Li JP, et al. Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network[J]. *Epidemiol Infect*, 2017, 145(6):1118-1129. DOI:10.1017/S0950268816003216.
- [32] Zhai MM, Li WH, Tie P, et al. Research on the predictive effect of a combined model of ARIMA and neural networks on human brucellosis in Shanxi Province, China: a time series predictive analysis[J]. *BMC Infect Dis*, 2021, 21(1):280. DOI:10.1186/s12879-021-05973-4.