

基于 SEAIQR 模型与 Dropout-LSTM 模型的 西安市 COVID-19 疫情趋势预测*

马艺菲¹ 许书君¹ 秦 瑶¹ 李建涛² 雷立健¹ 贺 鹭¹ 余红梅^{1,3Δ} 解 军^{3,4Δ}

【提 要】 目的 基于传染病动力学 SEAIQR (susceptible-exposed-asymptomatic-infected-quarantined-removed) 模型和 Dropout-LSTM (Dropout long short term memory network) 模型预测西安市新型冠状病毒肺炎 (COVID-19) 疫情的发展趋势, 为评估“动态清零”策略防控效果提供科学依据。方法 考虑到西安市本轮疫情存在大量的无症状感染者、依时变化的参数以及采取的管控举措等特点, 构建具有阶段性防控措施的时变 SEAIQR 模型。考虑到 COVID-19 疫情数据的时序性特征及它们之间的非线性关系, 构建深度学习 Dropout-LSTM 模型。选用 2021 年 12 月 9 日-2022 年 1 月 31 日西安市新增确诊病例数据进行拟合, 用 2022 年 2 月 1 日-2022 年 2 月 7 日数据评估预测效果, 计算有效再生数 (R_t) 并评价不同参数对疫情发展的影响。结果 SEAIQR 模型预测的新增确诊病例拐点预计在 2021 年 12 月 26 日出现, 约为 176 例, 疫情将于 2022 年 1 月 24 日实现“动态清零”, 模型 $R^2 = 0.849$ 。Dropout-LSTM 模型能够体现数据的时序性与非线性特征, 预测出的新增确诊病例数与实际情况高度吻合, $R^2 = 0.937$ 。Dropout-LSTM 模型的 MAE 和 RMSE 均较 SEAIQR 模型低, 说明预测结果更为理想。疫情暴发初期, R_0 为 5.63, 自实施全面管控后, R_t 呈逐渐下降趋势, 直到 2021 年 12 月 27 日降至 1.0 以下。随着有效接触率不断缩小、管控措施的提早实施及免疫阈值的提高, 新增确诊病例在到达拐点时的人数将会持续降低。结论 建立的 Dropout-LSTM 模型实现了较准确的疫情预测, 可为 COVID-19 疫情“动态清零”防控决策提供借鉴。

【关键词】 新型冠状病毒肺炎 SEAIQR 模型 Dropout-LSTM 模型 动态清零 预测 对比

【中图分类号】 R181.3 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.02.010

Prediction of COVID-19 Epidemic in Xi'an based on SEAIQR Model and Dropout-LSTM Model

Ma Yifei, Xu Shujun, Qin Yao, et al (School of Public Health, Shanxi Medical University(030001), Taiyuan)

【Abstract】 Objective This study aims to predict the coronavirus disease 2019 (COVID-19) epidemic in Xi'an based on SEAIQR model and Dropout-LSTM model, and to provide a scientific basis for evaluating the effectiveness of the “dynamic zero-COVID policy”. **Methods** Considering a large number of asymptomatic infections, the changing parameters, and control procedures, we developed a time-dependent susceptible-exposed-asymptomatic-infected-quarantined-removed (SEAIQR) model with stage-specific interventions. Considering the time-series characteristics of COVID-19 data and the nonlinear relationship between them, we constructed a deep learning Dropout-LSTM model. The data of newly confirmed cases in Xi'an from December 9th, 2021 to January 31st, 2022 were used to fit the model, and the data from February 1st, 2022 to February 7th, 2022 were used to evaluate the model performance of forecasting. We then calculated the effective reproduction number (R_t) and analyzed the sensitivity of the different measurement scenarios. **Results** The peak of newly confirmed cases predicted by the SEAIQR model would appear on December 26th, 2021, with 176 cases, and the “dynamic zero-COVID policy” may be achieved in January 24th, 2022, with $R^2 = 0.849$. The Dropout-LSTM model can reflect the time-series and nonlinear characteristics of the data, and the predicted newly confirmed cases were highly consistent with the actual situation, with $R^2 = 0.937$. The MAE and RMSE of the Dropout-LSTM model were lower than those of the SEAIQR model, indicating that the predicted results were more ideal. At the beginning of the outbreak, R_0 was 5.63. Since the implementation of comprehensive control, R_t has shown a gradual downward trend, dropping to below 1.0 on December 27th, 2021. With the reduction of effective contact rate, the early implementation of control measures and the improvement of immunity threshold, the peak of newly confirmed cases will continue to decrease. **Conclusion** The proposed Dropout-LSTM model forecasts the epidemic well, which can provide a reference for decision-making of the “dynamic zero-COVID policy.”

【Key words】 COVID-19; SEAIQR model; Dropout-LSTM model; Dynamic zero-COVID Policy; Prediction; Comparison

受全球第四波新型冠状病毒肺炎 (corona virus

disease 2019, COVID-19) 大流行的影响, 我国多地疫情呈现出局部聚集和多点散发态势。2021 年 12 月 9 日, 陕西省西安市出现 1 例本土确诊病例, 根据病毒测序结果, 西安市疫情暴发源头为 2021 年 12 月 4 日巴基斯坦航班输入的 Delta 毒株, 该毒株发病隐匿、传染性强, 加之西安市人口密度高、传播风险大, 原本已得到稳定控制的疫情近期又呈现出二次反弹趋势, 疫情防控形势复杂严峻。

* 基金项目: 国家重点研发计划 (2021YFC2301603); 山西省研究生科研创新项目 (2023KY363)

1. 山西医科大学公共卫生学院 (030001)
2. 山西医科大学管理学院
3. 煤炭环境致病与防治教育部重点实验室
4. 山西医科大学反向病原学中心

Δ 通信作者: 解军, E-mail: junxiesxmu@163.com; 余红梅, E-mail: yu@sxmu.edu.cn

Susceptible-exposed-infected-removed (SEIR) 模型是一种常见的传染病动力学模型,适用于有潜伏期,治愈后会获得终身免疫的传染病。长短期记忆网络(long short term memory network, LSTM)非常擅长处理时序性数据,能够“回忆”过去或未来数据中的模式,且不需要人为添加时间特征,可以在很大程度上探索时序性数据之间的非线性相关关系。目前,SEIR模型和 LSTM 模型已被应用于 COVID-19 预测领域^[1-3],但是由于病毒的不断变异以及防疫举措的优化,原有的新冠肺炎预测模型的结果不可避免地会与实际产生差异。本研究拟以西安市疫情病例为数据,构建适合我国当下疫情发展趋势的时变 SEAIQR 模型和 Dropout-LSTM 模型,为政府管理部门应对后续疫情、优化防疫管控提供参考依据。

资料来源

西安市 2021 年 12 月 9 日-2022 年 2 月 7 日 COVID-19 疫情个案数据来源于陕西省卫生健康委员会官网和“中国疾病预防控制中心信息系统”中的传染病监测系统^[4],包括每日新增确诊病例数、现有确诊病例数、累计确诊病例数、累计治愈病例数、累计死亡病例数、累计无症状感染者数等。同期人口数据来源于西安市统计局官网。

原理与方法

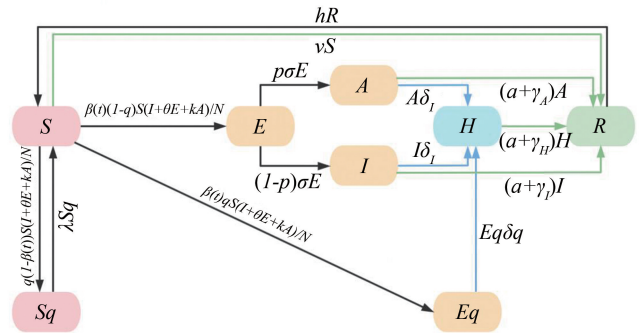
1. 模型设定

(1) 构建 SEAIQR 模型

在经典的 SEIR 模型框架下,考虑到西安市本轮疫情存在大量的无症状感染者、依时变化的参数以及采取的管控举措等特点,本研究进一步新增无症状感染者(asymptomatic infections, A)和隔离人群(quarantined, Q)[隔离易感者(quarantined susceptible, Sq)、隔离潜伏者(quarantined exposed, Eq)、住院患者(hospitalized, H)]仓室,构建具有阶段性防控措施的时变 Susceptible-exposed-asymptomatic-infected-quarantined-removed (SEAIQR) 模型,见图 1。其微分方程组如下:

$$\begin{aligned} \frac{dS}{dt} &= \frac{-[\beta(t)+q(1-\beta(t))]}{N} S(I+\theta E+kA) + \lambda Sq - vS + hR \\ \frac{dSq}{dt} &= \frac{q(1-\beta(t))}{N} S(I+\theta E+kA) - \lambda Sq \\ \frac{dE}{dt} &= \frac{\beta(t)(1-q)}{N} S(I+\theta E+kA) - \sigma E \\ \frac{dEq}{dt} &= \frac{\beta(t)q}{N} S(I+\theta E+kA) - \delta_q Eq \\ \frac{dA}{dt} &= p\sigma E - (\delta_I + \alpha + \gamma_A) A \end{aligned}$$

$$\begin{aligned} \frac{dI}{dt} &= (1-p)\sigma E - (\delta_I + \alpha + \gamma_I) I \\ \frac{dH}{dt} &= \delta_I(I+A) + \delta_q Eq - (\alpha + \gamma_H) H \\ \frac{dR}{dt} &= \gamma_I I + \gamma_A A + \gamma_H H + vS - hR + \alpha(I+A+H) \end{aligned}$$



* : S: 易感者; E: 潜伏者; I: 感染者; R: 康复者; Sq: 隔离易感者; Eq: 隔离潜伏者; A: 无症状感染者; H: 住院患者

图 1 COVID-19 传播动力学模型示意图

① 参数估计

采用马尔科夫链蒙特卡洛 (Markov Chain Monte Carlo, MCMC) 方法中的 Metropolis-Hastings (M-H) 算法进行参数估计。将初步估计出的参数值用作 M-H 算法的先验信息,使用 Python 软件进行 60000 次的迭代,退火后选取后面 55000 次迭代结果计算各参数平均值^[5]。具体参数设置见表 1。

表 1 COVID-19 传播动力学模型参数赋值

参数	含义	取值	来源
S_0	初始易感者人数	12952876	实际疫情
Sq_0	初始隔离易感者人数	12	实际疫情
E_0	初始潜伏者人数	6	实际疫情
Eq_0	初始隔离潜伏者人数	2	实际疫情
A_0	初始无症状感染者人数	1	实际疫情
I_0	初始确诊患者人数	1	实际疫情
H_0	初始住院患者人数	4	实际疫情
R_0	初始康复者人数	0	实际疫情
β_0	$t \in [0, t_1]$ 的有效接触率	0.902	MCMC
r	指数下降率	0.801	MCMC
θ	潜伏者相对于确诊患者的传染率系数	0.997	MCMC
k	无症状感染者相对于确诊患者的传染率系数	0.986	MCMC
q	隔离率	0.01	实际疫情
σ	潜伏者向感染者的转化速率	1/5.2	实际疫情
p	无症状感染者比例	0.3	实际疫情
α	病死率	0.002	实际疫情
λ	隔离解除速率	1/14	实际疫情
δ_I	感染者隔离速率	0.794	MCMC
δ_q	隔离潜伏者向住院患者转化速率	0.772	MCMC
γ_I	确诊患者康复率	0.085	MCMC
γ_A	无症状感染者康复率	0.122	MCMC
γ_H	住院患者康复率	0.088	MCMC
v	免疫阈值(疫苗接种率×疫苗保护率)	0.7	实际疫情
h	抗体滴度水平下降率	0.7	实际疫情

注: MCMC: 马尔科夫链蒙特卡洛

在本模型中 $\beta(t)$ 表示随时间变化的疫情有效接触率，它是关于时间 t 的分段函数。2021 年 12 月 9 日-2021 年 12 月 22 日属于病毒的快速传播期，这段时间防控措施较为欠缺，有效接触率为一个常数，即 $\beta(t) = \beta_0$ ；之后政府意识到疫情的严重性，开始实施全面管控，对感染者进行追踪隔离，此时有效接触率开始下降，假设以速率 r 进行指数式减少，那么本轮疫情内的有效接触率可以用如下的分段函数来表示：

$$\beta(t) = \begin{cases} \beta_0 & t < t_1 \\ \beta_0 \times e^{-r(t-t_1)} & t \geq t_1 \end{cases}$$

②基本再生数与有效再生数

基本再生数(basic reproduction number, R_0)是指一个感染者在具有传染性的这一段时间内平均可以传染多少个人。若 $R_0 > 1$ ，提示病例数会呈指数式增长，疾病将暴发或流行；若 $R_0 = 1$ ，提示该种疾病将在人群中持续存在并保持稳定；若 $R_0 < 1$ ，提示传染病将会逐渐消失^[6]。有效再生数(Effective reproduction number, R_t)则能够实时反映传染病在人群中的传播能力，随着政府管控逐步见效， R_t 会越来越低，当下降到 1 水平以下时，疫情会逐渐消亡。

利用下一代再生矩阵方法估计 R_t ^[7]，记：

$$F = \begin{bmatrix} \frac{\beta(t)(1-q)S(I+\theta E+kA)}{N} \\ \frac{\beta(t)qS(I+\theta E+kA)}{N} \\ 0 \\ 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma_E \\ \delta_q E_q \\ -p\sigma E + (\delta_I + \alpha + \gamma_A)A \\ -(1-p)\sigma E + (\delta_I + \alpha + \gamma_I)I \end{bmatrix}$$

进一步，对 F 和 V 在无病平衡点处求关于 E、Eq、A、I 的偏导，得：

$$F(t) = \begin{bmatrix} \frac{\beta(t)(1-q)S_0\theta}{N} & 0 & \frac{\beta(t)(1-q)S_0k}{N} & \frac{\beta(t)(1-q)S_0}{N} \\ \frac{\beta(t)qS_0\theta}{N} & 0 & \frac{\beta(t)qS_0k}{N} & \frac{\beta(t)qS_0}{N} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$V(t) = \begin{bmatrix} \sigma & 0 & 0 & 0 \\ 0 & \delta_q & 0 & 0 \\ -p\sigma & 0 & \delta_I + \alpha + \gamma_A & 0 \\ -(1-p)\sigma & 0 & 0 & \delta_I + \alpha + \gamma_I \end{bmatrix}$$

下一代再生矩阵定义为：

$$FV^{-1} = \begin{bmatrix} \frac{\beta(t)(1-q)S_0\theta}{N\sigma} + \frac{p\beta(t)(1-q)S_0k}{N(\delta_I + \alpha + \gamma_A)} + \frac{(1-p)\beta(t)(1-q)S_0}{N(\delta_I + \alpha + \gamma_I)} & & & \\ \frac{\beta(t)qS_0\theta}{N\sigma} + \frac{p\beta(t)qS_0k}{N(\delta_I + \alpha + \gamma_A)} + \frac{(1-p)\beta(t)qS_0}{N(\delta_I + \alpha + \gamma_I)} & & & \\ 0 & & & \\ 0 & & & \\ 0 & \frac{\beta(t)(1-q)S_0k}{N(\delta_I + \alpha + \gamma_A)} & \frac{\beta(t)(1-q)S_0}{N(\delta_I + \alpha + \gamma_I)} & \\ 0 & \frac{\beta(t)qS_0k}{N(\delta_I + \alpha + \gamma_A)} & \frac{\beta(t)qS_0}{N(\delta_I + \alpha + \gamma_I)} & \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \end{bmatrix}$$

因此， R_t 定义如下：

$$R(t) = \rho(FV^{-1}) = \beta(t)(1-q)S_0 \left[\frac{\theta}{N\sigma} + \frac{pk}{N(\delta_I + \alpha + \gamma_A)} + \frac{(1-p)}{N(\delta_I + \alpha + \gamma_I)} \right]$$

(2)构建 Dropout-LSTM 模型

长短期记忆网络(long short term memory network, LSTM)是一种时间循环神经网络，由 Hochreiter & Schmidhuber 等人于 1997 年提出。LSTM 广泛应用于深度学习领域，可用于解决循环神经网络(recurrent neural network, RNN)存在的长期依赖问题，能够综合处理长、短期输入并进行预测^[8]。LSTM 模型包括三个门结构[输入门(input gates)、输出门(output gates)、遗忘门(forget gates)]和一个细胞状态(cell)，其中遗忘门通过函数控制之前信息的输入程度，输入门控制当前信息的输入程度，输出门则用来控制最终输出，记忆细胞用于存储细胞状态信息。LSTM 模型整体框架图见图 2，计算公式如下：

$$f_t = \sigma(W_f[h_{t-1}, X_t] + b_f)$$

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, X_t] + b_c)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t$$

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o)$$

$$h_t = O_t \otimes \tanh(C_t)$$

式中： f_t 、 i_t 、 O_t 分别代表遗忘门、输入门、输出门的输出； \tilde{C}_t 为当前输入的记忆； C_{t-1} 为上一时刻的细胞状态； C_t 为当前时刻的细胞状态； h_t 为当前时刻的输出； W 为连接两层的权重矩阵； b 为偏置项； σ 为 sigmoid 激活函数。

本研究使用 Tensorflow 和 Keras 框架搭建 LSTM 网络。为最大程度保留数据特征并提高预测精度，首先对西安市每日新增确诊病例数进行归一化处理：

$$x_{\text{norm}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$

式中， x 代表本研究的原始数据， x_{max} 和 x_{min} 分别

为其最大值和最小值; x_{norm} 为数据归一化后的值。

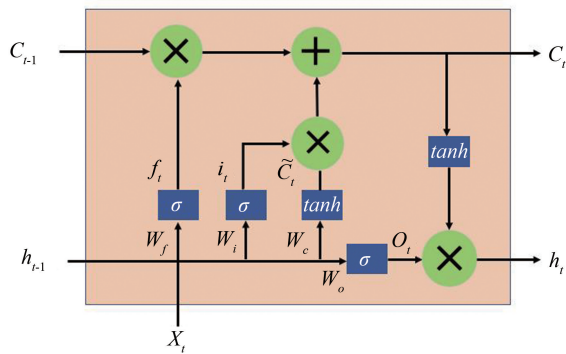


图2 LSTM模型整体框架图

将2021年12月9日-2022年1月31日西安市每日新增确诊病例数作为训练集输入模型,2022年2月1日-2022年2月7日数据作为测试集输出结果。根据模型试验确定窗口长度为7,隐藏层节点数为25,隐藏层层数为1,数据训练轮次 epoch 为100, batch_size 为1,添加一个全连接层,输出维度为1,模型的损失函数 loss 采用均方误差,优化器为 Adam。为有效避免 LSTM 神经网络的过拟合问题,本研究在隐藏层中的非循环部分采用 Dropout 技术对神经元进行随机概率失活,构建 Dropout-LSTM 模型,Dropout 为 0.1。因数据进入模型前进行了归一化处理,因此输出结果范围在 0 到 1 之间,故需将其反归一化以获得原始数据的预测值。

2. 软件实现

采用 Python 的 pandas 软件包构建 SEAIQR 模型, keras 软件包构建 Dropout-LSTM 模型。决定系数 (coefficient of determination, R^2)、平均绝对误差 (mean absolute error, MAE) 和均方根误差 (root mean squared error, RMSE) 用来比较模型的拟合和预测效果。 R^2 、MAE 和 RMSE 的定义分别为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

式中: y_i 为实际值, \hat{y}_i 为预测值, \bar{y}_i 为平均值, n 为预测天数。

结 果

1. SEAIQR 模型与 Dropout-LSTM 模型预测效果对比

选用2021年12月9日-2022年1月31日西安市 COVID-19 疫情个案数据进行拟合,2022年2月1日

-2022年2月7日数据评估预测效果。本研究构建具有阶段性防控措施的时变 SEAIQR 模型和 Dropout-LSTM 模型,分别预测西安市疫情中新增确诊病例的变化过程。由图 3(a)可知,SEAIQR 模型预测的新增确诊病例预计在 2021 年 12 月 26 日到达拐点,约为 176 例,随着大量住院患者被治愈,疫情将于 2022 年 1 月 24 日实现“动态清零”,经相关性检验分析,该模型决定系数 $R^2 = 0.849$,从而验证了所建立的模型对疫情走势的评估是可靠的。由图 3(b)可知,Dropout-LSTM 模型充分挖掘了新冠数据的时序性与非线性关系,预测出的新增确诊病例数与实际情况高度吻合, $R^2 = 0.937$ 。为了进一步衡量两个模型的性能,表 2 对其 MAE 和 RMSE 进行了对比。从评价指标来看,Dropout-LSTM 模型预测结果更理想,这说明 Dropout-LSTM 模型可较精准预测 COVID-19 的传播趋势。

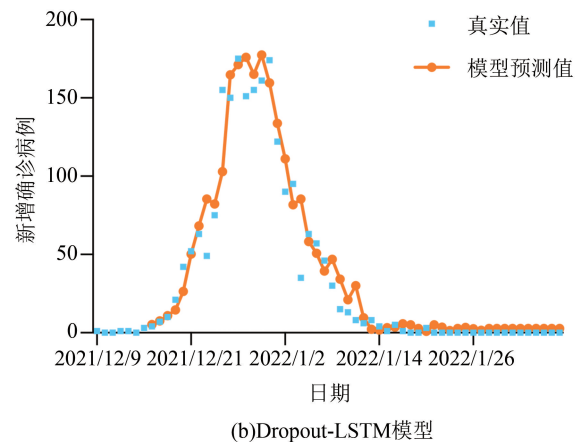
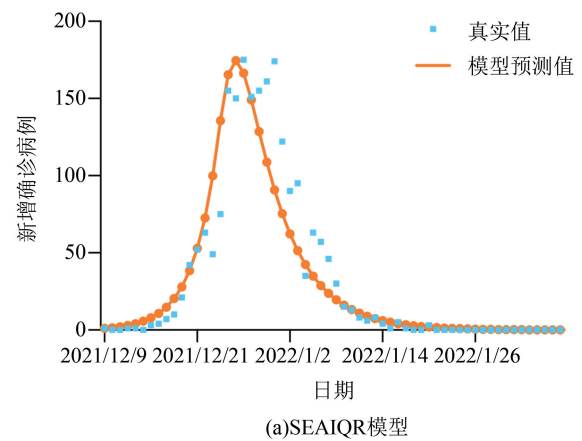


图3 SEAIQR 模型与 Dropout-LSTM 模型预测的西安市新增确诊病例数与实际确诊病例数比较

表2 评价指标对比

模型	MAE	RMSE
SEAIQR	0.111	0.119
Dropout-LSTM	0.051	0.053

2. 有效再生数的变化趋势

疫情暴发初期, R_0 是一个恒定不变的常数,为 5.63,此时若不实施严格管控,疫情则会快速蔓延,甚至失控。在疫情开始后第 15 天,西安市卫健委全

面追踪感染者并采取集中隔离措施, R_t 呈逐渐下降趋势, 直至 2021 年 12 月 27 日 (第 19 天), R_t 降至 1.0 以下, 疫情逐渐稳定。由图 4 可知, 有效再生数与有效接触率存在紧密联系, 二者变化趋势一致, 因此, 通过隔离、人员接触限制和社会疏远等措施控制有效接触率可减小本轮疫情的有效再生数。

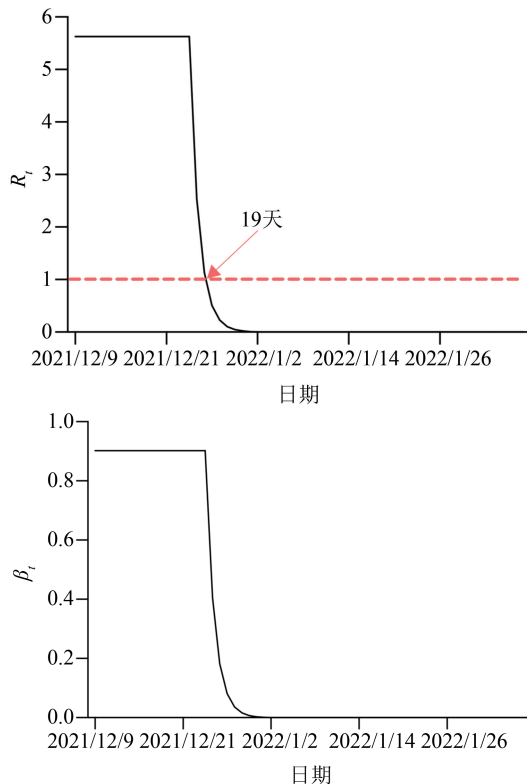


图 4 有效再生数和有效接触率的变化趋势

3. 不同参数对疫情发展的影响

(1) 有效接触率

接触传播是 COVID-19 最主要的传播途径, 降低人员接触数可有效控制病毒蔓延。有效接触率与接触数成正比例关系, 图 5 分析了有效接触率系数降低对疫情发展的影响。若将有效接触率缩小为原来的 0.9 和 0.7 倍时, 新增确诊病例拐点人数将会大幅降低, 尤其当取 $0.7\beta_0$ 时, 拐点人数将降低 11 倍, 仅有 16 人。此外, 随着有效接触率不断缩小, 新增确诊病例在到达拐点前的增幅和到达拐点后的降幅均会越来越

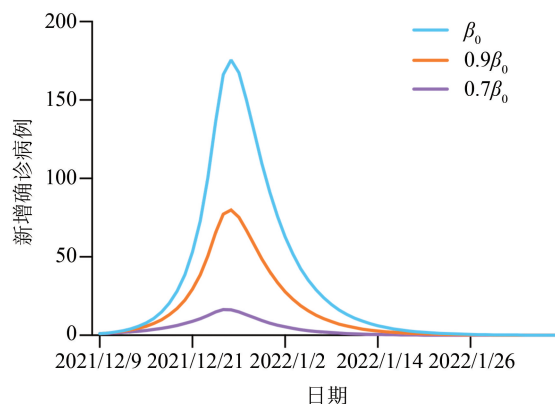


图 5 不同的有效接触率对新增确诊病例的影响

小。具体数值见表 3。

(2) 管控开始实施的时间

图 6 分析了管控开始实施的时间对疫情发展的影响。若政府能够提早 5 天、10 天实施全面管控, 新增确诊病例将于 2021 年 12 月 21 日和 2021 年 12 月 16 日迎来拐点, 拐点人数约降低 80% 和 96%; 若政府推迟 10 天采取措施, 拐点将延后至 2021 年 12 月 31 日出现, 此时感染人数大幅增长。具体数值见表 3。

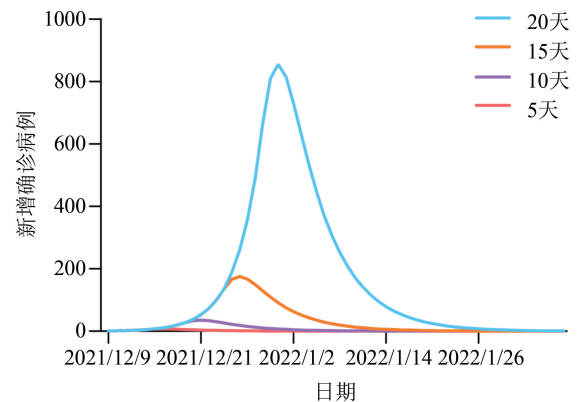


图 6 管控开始实施的时间对新增确诊病例的影响

(3) 免疫阈值

图 7 分析了免疫阈值对疫情发展的影响。随着免疫阈值不断提高, 新增确诊病例在到达拐点时的人数将持续降低, 当 ν 取值 0.75 和 0.8 时, 约降低 23% 和 40%。具体数值见表 3。

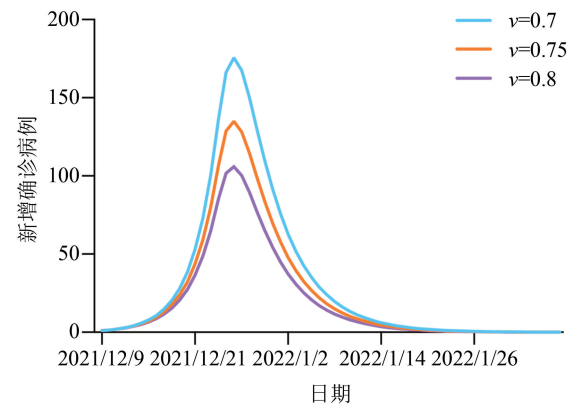


图 7 不同的免疫阈值对新增确诊病例的影响

表 3 不同参数下新增确诊病例到达拐点时间和拐点人数

参数	参数取值	新增确诊病例峰值(人)	到达峰值时间
有效接触率	β_0^*	176	第 18 天
	$0.9\beta_0$	80	第 18 天
	$0.7\beta_0$	16	第 17 天
管控开始实施的时间	第 20 天	855	第 23 天
	第 15 天*	176	第 18 天
	第 10 天	36	第 13 天
	第 5 天	7	第 8 天
免疫阈值	0.7^*	176	第 18 天
	0.75	135	第 18 天
	0.8	106	第 18 天

*: 指定参数值作为参考

讨 论

传染病动力学模型可用来研究疾病的流行规律与动态发展过程,为量化病毒传播风险、寻求防治最优策略、未雨绸缪确定检测及救治资源配置提供循证支持。本研究在经典的 SEIR 模型框架下,增加了对无症状感染者、隔离易感者、隔离潜伏者和住院患者的考虑,构建了具有阶段性防控措施的时变 SEAIQR 模型。西安市新增确诊病例预计在 2021 年 12 月 26 日到达拐点,2022 年 1 月 24 日实现“动态清零”,与疫情的实际情况(2021 年 12 月 27 日到达拐点,2022 年 1 月 21 日实现“动态清零”)基本一致。本研究对比了 SEAIQR 模型和 Dropout-LSTM 模型在新冠肺炎趋势预测上的性能,结果显示 Dropout-LSTM 模型的 MAE 和 RMSE 均较 SEAIQR 模型低,说明误差降低、准确率提高,Dropout-LSTM 模型可较精准预测 COVID-19 的传播趋势。为加快 loss 下降,减少模型训练时间,本研究对新增确诊病例数据进行归一化处理,以期更好地模拟 COVID-19 疫情数据之间的非线性关系,尽可能保留原始数据特征。因此,使用 Dropout-LSTM 模型对 COVID-19 进行预测是切实可行的。

本研究基于再生矩阵方法估计得出的西安市 COVID-19 疫情基本再生数为 5.63,远高于 SARS 和 MERS^[9],表明本轮疫情传播速度快、传染性较强。面对复杂、严峻的疫情形势,坚持“动态清零”总方针是一种正确且有效的防控策略。根据西安市 COVID-19 流行曲线,新增确诊病例数于 2022 年 1 月 21 日实现“动态清零”,防控效果已逐步显现。

本研究还发现,有效再生数与有效接触率联系密切。由于疫情防控措施逐步加强,有效接触率等参数不再是一个固定的常数,而是动态变化的,本研究进而将有效接触率设置为关于时间 t 的分段函数,以便更贴合防控措施下的实际疫情发展情况。通过参数敏感性分析可知,随着有效接触率不断缩小或管控措施的提早实施,新增确诊病例在到达拐点时的人数将会持续降低。此结果得到了相关研究的证实^[10]。

除有效接触率等参数外,本研究还考虑了人群免疫阈值(疫苗接种率 \times 疫苗保护率)与抗体滴度水平下降(疫苗保护时长有限)等因素在病毒传播过程中的作用,通过敏感性分析得出提高免疫阈值的必要性。COVID-19 疫苗有助于控制 SARS-CoV-2 的传播,但由疫苗诱导产生的保护会随着时间的推移而衰减。目

前我国疫苗覆盖率已经超过 85%,若进一步提高群体免疫阈值,则需要尽快接种第三针加强针以提高疫苗保护率。加强免疫方案有两种:一是同源加强免疫接种,二是异源加强免疫接种(也称“序贯加强免疫”)。已有研究表明,无论同源加强还是异源加强,都对变异株有效,而且异源加强对于提高灭活新冠疫苗的保护性免疫应答效果更好^[11]。因此,我国有必要推进序贯加强免疫接种,引导人民群众积极接种、尽早接种,进一步加强对人体的保护。

本研究也存在一些局限性:首先,本研究只考虑了防控措施下有效接触率的变化,而未考虑医疗救治情况改善对康复率的影响以及个人防护意识提高对自我隔离的影响。其次,忽视了人口流入、流出对于病毒传播的重要意义,这些问题在后续研究中需进一步探讨。

参 考 文 献

- [1] Zhao Y, Huang J, Zhang L, et al. Is the Omicron variant of SARS-CoV-2 coming to an end? *Innovation (Camb)*, 2022, 3(3): 100240.
- [2] Cai J, Deng X, Yang J, et al. Modeling transmission of SARS-CoV-2 Omicron in China. *Nat Med*, 2022, 28(7): 1468-1475.
- [3] 王瑞, 闫方, 逮静, 等. 运用 Dropout-LSTM 模型的新冠肺炎趋势预测. *电子科技大学学报*, 2021, 50(3): 414-421.
- [4] 陕西省卫生健康委员会. <http://sxwjw.shaanxi.gov.cn/>.
- [5] 丁中兴, 宋文煜, 方欣玉, 等. 基于 SEIAQR 动力学模型预测湖北省武汉市新型冠状病毒肺炎疫情趋势. *中国卫生统计*, 2020, 37(3): 327-330+334.
- [6] Yuan R, Ma Y, Shen C, et al. Global dynamics of COVID-19 epidemic model with recessive infection and isolation. *Math Biosci Eng*, 2021, 18(2): 1833-1844.
- [7] van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci*, 2002, 180: 29-48.
- [8] ArunKumar KE, Kalaga DV, Kumar CMS, et al. Forecasting of COVID-19 using deep layer Recurrent Neural Networks(RNNs) with Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) cells. *Chaos Solitons Fractals*, 2021, 146: 110861.
- [9] Petrosillo N, Viceconte G, Ergonul O, et al. COVID-19, SARS and MERS: are they closely related? *Clin Microbiol Infect*, 2020, 26(6): 729-734.
- [10] 曹盛力, 冯沛华, 时朋朋. 修正 SEIR 传染病动力学模型应用于湖北省 2019 冠状病毒病(COVID-19)疫情预测和评估. *浙江大学学报(医学版)*, 2020, 49(2): 178-184.
- [11] Ai J, Zhang H, Zhang Y, et al. Omicron variant showed lower neutralizing sensitivity than other SARS-CoV-2 variants to immune sera elicited by vaccines after boost. *Emerg Microbes Infect*, 2022, 11(1): 337-343.

(责任编辑:张悦)