·调查报告与分析·

非药物干预措施对手足口病传播及预测模型阶段影响：基于 2014—2023 年安徽省凤阳县手足口病监测数据分析

常宗楠，方翼飞，杨圣，梁戈玉

东南大学公共卫生学院，南京 210009
通信作者：梁戈玉，E-mail：lianggeyu@163.com

【摘　要】目的　探究新型冠状病毒感染（COVID-19，简称新冠）疫情期间非药物干预措施（NPIs）对手足口病传播趋势的阶段性影响，评估不同预测模型在疫情干预背景下的适用性和动态响应能力。**方法**　以 2014—2023 年中国安徽省凤阳县手足口病月度发病率数据为基础，采用 Theil–Sen 稳健回归估计疫情前后年度百分比变化（APC）与平均年度百分比变化（AAPC）；基于“基线方案（反事实压力测试）”与“干预方案（适应性预测）”进行样本外预测比较。基线方案：以 2014 年 1 月—2019 年 12 月数据为训练集，以 2020 年 1 月—2023 年 12 月数据为验证集，不纳入疫情干预变量；干预方案：以 2014 年 1 月—2022 年 11 月资料作为训练集并纳入阶段性干预哑变量，以 2022 年 12 月—2023 年 12 月资料作为验证集。比较含外生回归量的季节性差分自回归滑动平均模型（SARIMAX）、加性分解时间序列模型（Prophet）、长短期记忆网络（LSTM）的平均绝对百分比误差（MAPE）、均方根误差（RMSE）、平均绝对误差（MAE）。研究创新性提出阶段敏感度指数（PSI）与恢复弹性系数（REC），从动态响应性与抗干扰能力角度评估模型的适应性。**结果**　疫情前（2014—2019）手足口病发病率呈下降趋势（APC=−15.97%，95%*CI*= −22.06% ~ −6.80%，*P*=0.004）；疫情后（2020—2023）转为上升（APC=36.17%，95%*CI*=8.84% ~ 54.72%，*P*=0.002）；全期（2014—2023）总体仍下降（AAPC=−12.28%，95%*CI*= −16.49% ~ −7.37%，*P*<0.001）。在开放期（2022.12—2023.12）样本外预测中，干预方案优于基线方案：SARIMAX 的 MAPE 由 73.77% 降至 44.54%（下降 29.23 个百分点），LSTM 由 46.18% 降至 41.85%，Prophet 则仍偏高（69.73%）。PSI/REC 显示：SARIMAX 与 Prophet 干预方案对常态化防控→放开过渡较敏感（PSI>1），且 REC=1（1 个月内恢复）；LSTM 的 PSI 接近 1，但 REC=0.33（约 3 个月），提示恢复较慢。**结论**　本研究分析了非药物干预对手足口病流行趋势的阶段性作用与反弹风险，验证了统计模型与深度学习模型在传染病监测中的互补性，前者对外生冲击敏感且恢复快，后者短期拟合优但在放开后需加强泛化与自适应更新。未来应结合多模型策略优化传染病预警体系，以应对疫情干预背景下的复杂流行态势。

【关键词】 手足口病；非药物干预措施；时间序列预测；动态响应性；恢复弹性

Non-pharmaceutical intervention measures and their stage-specific effects on hand, foot and mouth disease transmission and forecasting models: an analysis based on HFMD surveillance data from Fengyang county, Anhui province, 2014–2023

CHANG Zongnan, FANG Yifei, YANG Sheng, LIANG Geyu (*School of Public Health*, *Southeast University*, *Nanjing* 210009, *China*)
Corresponding author: LIANG Geyu, E-mail: lianggeyu@163.com

【Abstract】 Objective To investigate the phase-specific impacts of non-pharmaceutical interventions (NPIs) during the coronavirus disease 2019 (COVID-19) pandemic on the transmission of hand, foot and mouth disease (HFMD) and evaluate the applicability and dynamic responsiveness of different prediction models in an intervention context. **Methods** Monthly HFMD incidence data from 2014 to 2023 in Fengyang county, Anhui province, China were analyzed. Theil–Sen robust regression was employed to estimate the annual percent change (APC) and the average annual percent change (AAPC) before and after COVID-19. Baseline scheme (counterfactual stress test): the model was trained on the data from January 2014 to December 2019 and validated on the data from January 2020 to December 2023, without incorporating any pandemic intervention variables. Intervention scheme (adaptive prediction): the model was trained on the data from January 2014 to November 2022, incorporating phase-specific intervention dummy variables, and validated on the data from December 2022 to December 2023. The prediction performance of seasonal autoregressive integrated moving average model with exogenous regressors (SARIMAX), Prophet (a time-series decomposable additive model), and long short-term memory (LSTM) was compared based on mean absolute percentage error (MAPE), root mean square error (RMSE), and mean absolute error (MAE). Two novel indices, the phase sensitivity index (PSI) and the resilience coefficient (REC), were proposed to assess model adaptability in terms of dynamic responsiveness and resistance to perturbations. **Results** In the pre-pandemic period from 2014 to 2019, HFMD incidence declined (APC = −15.97%, 95%*CI*: −22.06% to −6.80%, *P* = 0.004). In the post-pandemic period from 2020 to 2023, it increased (APC = 36.17%,

开放获取：CC BY-NC-ND 4.0　DOI：10.11847/zgggws1147448
第一作者：常宗楠（1997—），医师，硕士，研究方向：传染病预防与控制。
收稿日期：2025-03-20　修回日期：2025-10-14　录用日期：2025-12-03
利益冲突：不存在　伦理审查：免除审查　出版授权：全体作者已与编辑部签署作者贡献声明及版权转让协议

95%*CI*: 8.84% to 54.72%, $P = 0.002$). Over the full period from 2014 to 2023, the overall trend still decreased (AAPC = −12.28%, 95%*CI*: −16.49% to −7.37%, $P < 0.001$). During the relaxation period from December 2022 to December 2023, the intervention scheme outperformed the baseline scheme in out-of-sample prediction: the MAPE of SARIMAX decreased from 73.77% to 44.54% (a reduction of 29.23 percentage points); the MAPE of LSTM decreased from 46.18% to 41.85%; whereas the MAPE of Prophet remained high at 69.73%. The PSI/REC indicated that the SARIMAX and Prophet intervention models were more sensitive to the transition from routine control to relaxation, with PSI greater than 1 and recovery within 1 month (REC equal to 1); LSTM showed a PSI close to 1 but slower recovery (REC equal to 0.33) in about 3 months. **Conclusions** Non-pharmaceutical interventions exert phase-specific effects on HFMD transmission with a risk of rebound. Statistical models and deep learning models are complementary for surveillance under intervention scenarios. Statistical models are sensitive to exogenous shocks and recover quickly, whereas deep learning models excel in short-term fitting but require enhanced generalization and adaptive updating after relaxation of control measures. Integrating multiple models is recommended to optimize infectious disease early warning systems under complex dynamics in the presence of interventions.

【Keywords】 hand, foot, and mouth disease; non-pharmaceutical interventions; time-series prediction; dynamic responsiveness; resilience

全球传染病监测体系正经历着前所未有的挑战。2019 年底暴发的新型冠状病毒感染（the coronavirus disease 2019, COVID-19）引发了空前的全球公共卫生响应，各国实施了居家隔离、社交距离限制、公共场所关闭、加强手卫生、环境消毒、健康监测等在内的综合性非药物干预措施（non-pharmaceutical interventions, NPIs），显著改变了多种传染病的传播环境和流行趋势[1−3]。手足口病（hand, foot and mouth disease, HFMD）是由肠道病毒引起的常见传染病，主要通过接触传播，好发于 5 岁以下儿童，在中国呈现明显的季节性流行特征[4−6]。作为一种与社会接触密切相关的疾病，手足口病的传播受 COVID-19 疫情期间 NPIs 影响尤为明显。虽然 NPIs 主要针对 COVID-19 这一呼吸道传染病，但其中的居家隔离减少了儿童聚集活动，学校停课降低了密切接触频率，加强手卫生和环境消毒直接切断了手足口病的接触传播途径，因此，对手足口病等肠道传染病的传播也产生了显著的“溢出效应”[7]。随着防控措施的逐步解除，手足口病等传染病的发病趋势及预测模型适用性面临新的挑战。传染病预测模型主要包括基于统计学原理的时间序列模型如 SARIMAX、基于贝叶斯框架的分解模型如 Prophet，以及深度学习方法如 LSTM 网络[8−11]，这些模型在常规环境下各有优势，但在面对 COVID-19 等重大公共卫生事件干预背景下的预测能力尚未得到充分验证，特别是当防控政策从严格管控到全面放开转变时，传染病发病模式的变化及其对预测模型的影响更需深入研究。已有研究发现，多种传染病在 COVID-19 疫情期间发病率显著下降。Feng L Z 等[12] 报告中国和美国流感活动在实施 NPIs 后大幅减少。有研究分析不同程度的 NPIs 对广州市 HFMD 发病率的影响，发现疫情期间 HFMD 发病率显著降低，但在放松防控后出现反弹[7]。也有研究提示，NPIs 解除后可能出现的“免疫债务”现象，可能导致流感、水痘等传染病的反弹[13]。这种动态变化对预测模型提出新要求，不仅需准确捕捉干预下的趋势变化，还需具备应对政策调整的敏感性和恢复能力。目前，关于不同预测模型在 COVID-19 疫情各阶段的适用性比较研究相对有限，特别是针对中国县域级别的手足口病等传染病，在从严格防控到全面开放过渡期间的预测表现缺乏系统评估。现有研究多关注模型的整体预测精度，较少从动态响应能力和恢复弹性的角度评估模型对防控政策变化的适应能力。为此，本研究以 2014—2023 年安徽省凤阳县手足口病月度发病率数据为基础，探究 COVID-19 疫情前后手足口病的发病趋势变化，并系统比较 SARIMAX、Prophet 和 LSTM 这 3 种代表性预测模型在不纳入疫情因素干预下的预测表现，引入阶段敏感度指数（phase sensitivity index, PSI）和恢复弹性系数（resilience coefficient, REC）两项评价指标，从动态响应和恢复能力角度评估各模型在疫情防控政策转变期的适应性。研究结果将为 COVID-19 后疫情时代传染病监测预警系统的优化提供科学依据，并为公共卫生决策和资源配置提供参考。结果报告如下。

1 资料与方法

1.1 *数据来源与预处理* 利用中国疾病预防控制信息系统中的传染病监测系统获取 2014 年 1 月—2023 年 12 月安徽省滁州市凤阳县手足口病的月度发病率数据。凤阳县为安徽省手足口病高发地区，根据第七次全国人口普查公报，该县 2020 年常住人口 63.19 万人。提高的人口基数和较高的发病水平为本研究提供了坚实的样本基础。选择此数据是基于以下两点考量：手足口病发病具有极强的季节性周期，长达 120 个月的序列足以让模型捕捉此稳定规律；此行政单元内，COVID-19 期间 NPIs 的执行在县域内部更具同质性，便于用阶段性变量捕捉综合强度变化；该层级与公共卫生决策的资源配置与预警触发机制高度匹配，研究流程便于在疾病控制中心日常工作中直接复用。为消除噪声效应，本研究对极端值进行了

Winsorize 处理，即将高于 99 分位数的值截断至该分位数，并对低于阈值的异常值进行下限截断。

1.2 *疫情阶段划分*　鉴于 COVID-19 疫情期间防控政策呈现阶段性特征，根据凤阳县的实际情况，将 2014 年 1 月—2019 年 12 月视为疫情前阶段，用于评估长期平稳期的流行水平；2020 年 1 月—2023 年 12 月则代表疫情及其后续时段，并进一步细分为 4 个阶段：第一波疫情暴发期（2020 年 1 月—2020 年 3 月）、常态化防控期（2020 年 4 月—2022 年 11 月）、开放过渡期（2022 年 12 月—2023 年 3 月）和开放期（2023 年 4 月—2023 年 12 月）。但在后续模型构建部分，考虑到部分阶段样本量过少可能导致模型估计不稳，本研究将疫情期间合并为两个阶段：疫情防控期间（2020 年 1 月—2022 年 11 月）和开放期（2022 年 12 月—2023 年 12 月）。

1.3 *趋势分析方法*　为探讨传染病在疫情前后的长期演变趋势，本研究采用 Theil-Sen 稳健回归模型计算年度百分比变化（annual percent change, APC）与平均年度百分比变化（average annual percent change, AAPC）。先分别计算逐年发病率的年化变化率，再对多年的变化率取平均以得到 AAPC。与传统回归方法相比，Theil-Sen 回归对异常值与非正态分布更具鲁棒性，能在疫情冲击下更准确反映发病率的整体波动[14]。

1.4 *模型比较与干预设计*　为评估疫情干预对传染病发病率预测的影响，从统计时序建模、机器学习分解和深度学习等多角度切入，选取了 SARIMAX、Prophet 及 LSTM 这 3 种具有代表性的建模方法。考虑到 NPIs 的实施是多方面且难以精确量化的，本研究采用阶段性哑变量代表不同时期综合干预措施的整体强度。将不同管控政策时期设为分类哑变量纳入模型，用以捕捉这些措施在特定时间段内的综合效应，并以此预测开放期发病率。为严格评估模型在不同情境下的泛化能力，设计了两种基于训练集—验证集框架的样本外预测方案：（1）基线方案（反事实压力测试）。此方案旨在模拟无疫情信息下的模型表现，并量化疫情冲击对原有疾病传播模式的颠覆程度。本研究将 2014—2019 年的数据作为训练集进行建模，并将此后的 2020 年 1 月—2023 年 12 月的数据作为验证集。需要强调的是，该方案并非旨在模拟真实的滚动预测实践，而是一个“反事实”设定[15–16]，其目的是通过观察模型在未知疫情冲击下的预测表现，量化疫情对此前稳定传播模式的“结构性断裂”影响。因此，该方案产生的预测误差本身即是关键的衡量指标，并为后续评估干预方案的改进效果提供了一个必要的比较基准。（2）干预方案（适应性预测评估）。此方案旨在评估模型在学习了干预信息后的预测能力，更贴近实际应用场景。本研究使用更大范围的 2014 年 1 月—2022 年 11 月的数据（包含疫情干预变量，干预变量为哑变量，分别表示第一波疫情暴发期、常态化防控期）作为训练集，然后对“开放期”（2022 年 12 月—2023 年 12 月）这一独立的验证集开展样本外预测。通过比较两种方案在相应验证集上的误差表现，旨在科学评估纳入疫情干预信息对模型远期预测能力的真实影响。鉴于县级层面难以获得各项 NPIs 的量化执行数据，且中国防控实践中各措施多为组合实施，本研究采用阶段性哑变量代表不同时期的综合干预强度，这一方法在国际同类研究中被广泛采用[7,17]。通过比较两种方案在各阶段的误差表现，评估疫情干预对模型外推能力的影响。

1.4.1 SARIMAX 模型　SARIMAX 模型（seasonal auto-regressive integrated moving average with exogenous regressors）基于 Box-Jenkins 方法的季节性差分自回归滑动平均模型。对于一个具有季节周期 S 的时间序列 $\{y_t\}$，其形式可用下式表征：

$$(1-\sum_{i=1}^{p}\phi_i B^i)(1-\sum_{j=1}^{P}\Phi_j B^{jS})(1-B)^d(1-B^S)^D y_t = c + (1+\sum_{k=1}^{q}\theta_k B^k)(1+\sum_{l=1}^{Q}\Theta_l B^{lS})\varepsilon_t + X_t^{\top}\beta. \quad (1)$$

其中 p，d，q 与 P，D，Q 分别表示非季节性与季节性的自回归、差分和滑动平均阶数，S 为季节周期，B 为滞后算子，ε_t 为零均值白噪声，Xt 为外生变量向量，β 为其系数；c 为常数项。通过网格搜索（grid search）在可行参数空间（p，d，q）×（P，D，Q，S）中确定最优组合。该模型具有清晰的统计理论基础，能够在季节性和趋势基础上同时容纳干预变量，从而对突发事件或政策变化予以显式建模[18]，也便于检验残差自相关性与模型适配度，解释性强。

1.4.2 Prophet 模型　Prophet 模型以加性分解思想将时间序列拆解为趋势项 $g(t)$、季节项 $s(t)$ 和干预项 $h(t)$：

$$y(t) = g(t)+s(t)+h(t)+\varepsilon_t \quad (2)$$

在构建 Prophet 模型时，将 changepoint prior scale 初始设为 0.05，并通过网格搜索调优该参数，同时利用“假日/特殊事件”机制或自定义回归器将疫情防控期间的干预措施作为特殊事件引入模型。Prophet 模型对异常值、缺失数据具有较高稳健性，可通过其机制或外生变量灵活捕捉突发干预效应，且参数调试简单易行，适于公共卫生环境多变、离散数据的快速验证[19]。

1.4.3 LSTM 模型 在深度学习层面，构建 3 层 LSTM 模型（long short-term memory）网络来捕捉时间序列的长期依赖和非线性特征。输入层包含过去 12 个月的发病率及疫情干预变量，隐藏层含 64 个神经元（激活函数为 tanh），并设置 Dropout（rate=0.2）防止过拟合。为增强对突发事件的鲁棒性，在输入端添加高斯噪声层（σ=0.1），通过贝叶斯优化搜索关键超参数（如隐藏单元数量、学习率等）。干预变量可与历史时间并行输入，从而捕捉疫情冲击下发病率的复杂动态。LSTM 模型作为循环神经网络的特殊结构，能捕捉时间序列的长期依赖与非线性特征，且在面对疫情等极端干预时有潜在优势；通过增加干预变量与噪声模拟，可更充分挖掘复杂时序中的隐含模式；对大规模数据具有强表达能力，适合挖掘潜在模式[20]。

1.5 评估指标 为量化模型性能，从两个维度评价：（1）整体预测偏差。使用平均绝对百分比误差（mean absolute percentage error, MAPE）、均方根误差（root mean square error, RMSE）以及平均绝对误差（mean absolute error, MAE）衡量模型对真实发病率的偏离度，具体计算公式见参考文献 [21]；对于阶段性评估，分别计算在第一波疫情暴发期、常态化防控期及全面放开过渡期各阶段内的平均误差。（2）为了评估模型对疫情防控策略调整的响应特性，本研究引入阶段敏感度指数（phase sensitivity index, PSI），用于衡量模型在关键防控措施转换期（如疫情初期→第一波疫情、常态化防控→全面开放）对发病率变化的响应程度。设计思想借鉴了多个学科中基于斜率比值来衡量系统响应与恢复能力的方法。例如，在社区行为学中，“斜率比值”（slope ratio）被用来量化社区韧性[22]；在医学研究中，恢复曲线的斜率也是评估身体韧性的关键参数[23]。这些研究为采用斜率比值作为响应性指标提供了方法学支持。PSI 的计算基于标准化斜率比值法：首先，对真实发病率和模型预测值在转换期内分别进行简单线性回归，得到各自的趋势斜率；其次，将斜率除以该时期的均值予以标准化，得到相对变化率；最后，计算模型相对变化率与真实相对变化率的比值。这一设计保留了传统线性回归中 β 系数对干预响应的解释效应，同时解决了不同模型结构下参数不可比的问题，具有更强的模型间适用性与横向比较能力。其数学表达式为：

$$\mathrm{PSI}=\left|\frac{\Delta\hat{Y}/\overline{\hat{Y}}}{\Delta Y/\overline{Y}}\right| \qquad (3)$$

其中，$\Delta\hat{Y}$ 与 $\overline{\hat{Y}}$ 分别表示模型预测值在该阶段的斜率与均值；ΔY 与 $\overline{Y}$ 则分别表示真实值的斜率与均值。PSI 的核心是量化模型在应对外部环境突变时的行为模式，PSI 越接近 1，说明模型对真实变化的响应越准确；若远离 1，则分别对应过度敏感（>1）或响应滞后（<1）的情况。

采用恢复弹性系数（resilience coefficient，REC）评价模型在全面放开后的误差回归稳定性。该指标的设计与灾害风险科学领域中基于恢复速率量化韧性的指标（如 $R10$，$R11$）逻辑高度一致[24]，旨在评估模型在遭遇冲击后的系统韧性，即误差收敛的速度。针对全面放开后的恢复情况，首先定义恢复时间 T_r 为预测误差（MAPE）降至基准期误差（baseline MAPE）1.5 倍以内所需的月份数，随后将其取倒数得到 REC，以此反应速度，计算公式如下：

$$\mathrm{REC}=\frac{1}{T_r} \qquad (4)$$

其中 T_r=min{t|MAPE(t)≤1.5×Baseline MAPE}。若模型能够在更短时间内恢复到基准误差水平，则 T_r 较小，REC 值较大，表示模型的恢复弹性更强；反之则说明模型在新政策环境下的误差波动时间较长，恢复能力较弱。PSI 与 REC 并非旨在替代 MAPE 等传统精度指标，而是作为其重要的互补评价维度，分别从“瞬时响应”和“后期恢复”两个动态维度提供额外洞见。这种多维度评估框架在突发公共卫生事件的非平稳环境下具有重要应用价值。为验证指标的有效性，本研究对窗口长度（±1 个月）和恢复阈值（1.2、1.5、2.0 倍）进行了敏感性测试，并分析了与传统精度指标的相关性。

1.6 统计学分析 所有统计学分析与可视化操作均在 Python 3.11.5 环境下完成：SARIMAX 通过 statsmodels 库实现，Prophet 使用 fbprophet 库，LSTM 则利用 TensorFlow 建模与验证。统计学检验采用双侧检验，检验水准 α=0.05。以 P<0.05 为差异有统计学意义。

2 结 果

2.1 描述性统计与基本趋势（图 1） 2014—2023 年凤阳县手足口病的月度发病率呈现明显的季节性特征，发病高峰多出现在 4—7 月。2020 年初 COVID-19 疫情暴发后，手足口病的发病率出现了一定程度的下降。进一步分析各阶段的发病率均值显示，疫情前为 25.909/10 万，第一波疫情暴发期（2020 年 1 月—2020 年 3 月）降至 4.789/10 万，常态化防控期（2020 年 4 月—2022 年 11 月）回升至 12.085/10 万，全面放开过渡期（2022 年 12 月—2023 年 3 月）为 12.815/10 万，疫情后（2023 年 4 月及之后）进一步升至 18.019/10 万。参照常住人口（63.19 万人）将阶段平均月发病率换算为平均每月病例数，疫情前约 164 例/月，初期暴发期约

30 例/月，常态化防控约 76 例/月，过渡期约 81 例/月，全面开放期约 114 例/月。Kruskal-Wallis *H* 检验结果显示，不同阶段的发病率差异有统计学意义（*H*=24.411，*P*<0.001）。



图 1　2014—2023 年手足口病月度发病率趋势
Fig. 1　Trend of HFMD incidence during 2014–2023

2.2　趋势分析结果　基于 Theil-Sen 稳健回归模型评估 2014—2023 年手足口病发病率的长期趋势，并计算了年度百分比变化（APC）及平均年度百分比变化（AAPC）。在疫情前（2014—2019 年），手足口病的发病率呈下降趋势（APC= −15.97%，95%*CI*=−22.06%～−6.80%，*P*=0.004）。然而，在疫情后（2020—2023 年）发病率趋势发生逆转，呈现上升趋势（APC=36.17%，95%*CI*=8.84%～54.72%，*P*=0.002）。综合全研究期（2014—2023 年），尽管疫情后期发病率有所回升，但整体仍呈下降趋势（AAPC=−12.28%，95%*CI*= −16.49%～−7.37%，*P*<0.001）。

2.3　预测模型比较结果

2.3.1　基线方案（表 1、图 2）　各模型在疫情前（2014—2019 年）的拟合精度尚可，但在疫情后预测误差有较大增长，以基线方案的 MAPE 为例，SARIMAX 模型为 34.69%，Prophet 方案为 53.31%，LSTM 模型为 37.81%；然而，在疫情期间（2020—2022 年），基线方案的预测误差大幅上升，SARIMAX 模型增至 38.85%，Prophet 模型飙升至 86.39%，LSTM 模型虽有所下降（30.80%），但仍表现出较大误差。在开放期（2022-12—2023-12），基线方案的预测误差进一步扩大，SARIMAX 模型（73.77%）、Prophet 方案（89.08%）和 LSTM 模型（46.18%）基线模型的 MAPE 表明，未考虑疫情干预因素的模型在疫情暴发后难以准确预测发病率的变化趋势。

表 1　不同时间序列模型在基线方案下的拟合与预测性能评估
Table 1　Evaluation of fitting and prediction performance of different time series models under the baseline scheme

模型/方案	时间范围	拟合/预测	MAPE/%	RMSE	MAE
SARIMAX	2014.01—2019.12	拟合	34.69	14.33	9.31
	2020.01—2022.11	预测	38.85	12.61	7.69
	2022.12—2023.12	预测	73.77	18.28	13.85
Prophet	2014.01—2019.12	拟合	53.31	14.22	10.65
	2020.01—2022.11	预测	86.39	15.65	11.26
	2022.12—2023.12	预测	89.08	19.49	15.49
LSTM	2014.01—2019.12	拟合	37.81	14.31	9.06
	2020.01—2022.11	预测	30.80	9.25	5.67
	2022.12—2023.12	预测	46.18	12.35	8.78

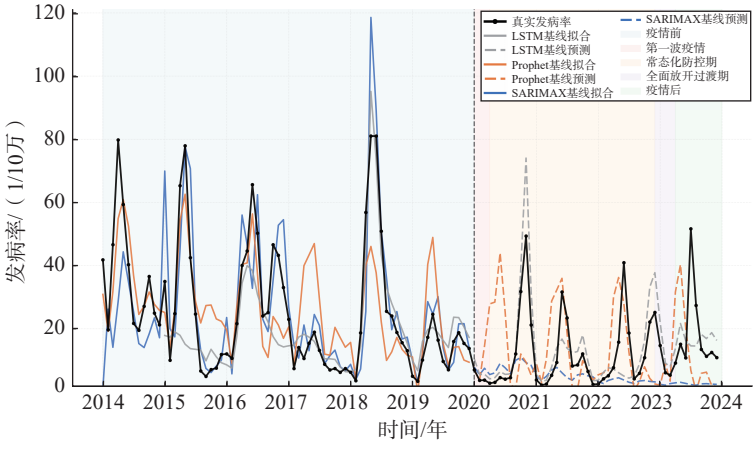


图 2　2014—2023 年手足口病发病率基线方案拟合与预测结果
Fig. 2　Fitting and prediction results of models for hand, foot, and mouth disease incidence under the baseline scheme during 2014−2023

2.3.2 干预方案(表 2、图 3) 干预方案的预测精度优于基线方案。以 MAPE 为例，在疫情前(2014—2019 年)，干预方案的 MAPE，SARIMAX 模型为 32.67%，Prophet 方案为 56.70%，LSTM 模型为 24.80%。在疫情期间(2020—2022 年)，干预方案的拟合误差普遍降低，其中 LSTM 模型的 MAPE 最低(6.56%)，SARIMAX 为模型 30.46%，而 Prophet 方案仍然较高(70.14%)。在开放期(2022.12—2023.12)，SARIMAX 模型的 MAPE 降至 44.54%，相比基线方案减少 29.23 个百分点，LSTM 降至 41.85%，但 Prophet 方案的误差仍然较大(69.73%)，说明该模型对疫情防控放开后的趋势适应能力有限。整体来看，干预方案的预测误差小于基线方案，表明纳入疫情防控因素的模型能够更准确地展现疫情后传染病的传播模式。SARIMAX 模型在各阶段的预测能力较为稳定，LSTM 模型在疫情期间的拟合误差最低，但在开放期的误差相对较大，而 Prophet 模型的预测能力在疫情期间和开放期均表现较差。这些结果表明，考虑疫情干预的统计模型(如 SARIMAX)在长期趋势预测中更具优势，而深度学习模型(如 LSTM)在短期趋势拟合方面表现突出，但在疫情后期的预测能力仍有待优化。

表 2 不同时间序列模型在干预方案下的拟合与预测性能评估

Table 2 Evaluation of fitting and prediction performance of different time series models under the intervention scheme

模型/方案	时间范围	拟合/预测	MAPE/%	RMSE	MAE
SARIMAX	2014.01—2019.12	拟合	32.67	13.04	8.55
	2020.01—2022.11	拟合	30.46	7.70	5.02
	2022.12—2023.12	预测	44.54	13.04	8.61
Prophet	2014.01—2019.12	拟合	56.70	14.87	11.47
	2020.01—2022.11	拟合	70.14	12.34	9.31
	2022.12—2023.12	预测	69.73	17.79	13.18
LSTM	2014.01—2019.12	拟合	24.80	12.08	6.72
	2020.01—2022.11	拟合	6.56	2.02	1.11
	2022.12—2023.12	预测	41.85	11.61	8.21

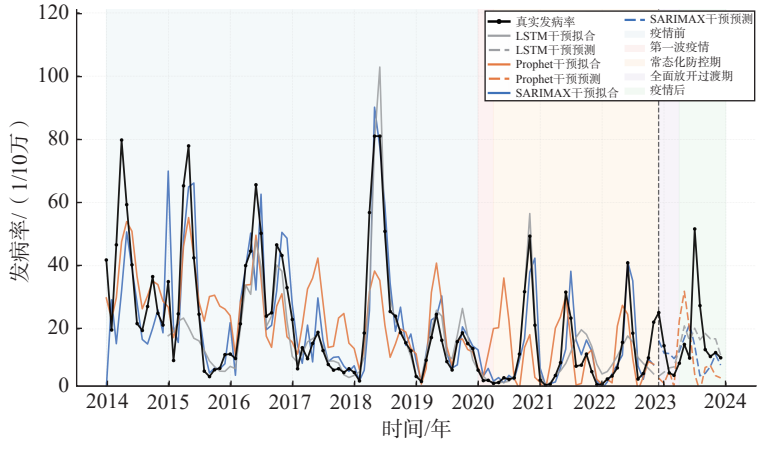


图 3 2014—2023 年手足口病发病率干预方案拟合与预测结果

Fig. 3 Fitting and prediction results of models for hand, foot, and mouth disease incidence under the intervention scheme during 2014−2023

2.4 阶段敏感性与恢复弹性分析(图 4) 不同模型在疫情防控转换阶段的 PSI 存在明显差异。其中，Prophet 干预方案在常态化防控至过渡期(normal→transition)阶段的 PSI 值最高(PSI=5.19)，其次是 SARIMAX 干预方案(PSI=4.85)，表明这两种模型在该阶段对疫情变化极为敏感，可能存在较强的短期波动性。相比之下，LSTM 干预方案在所有阶段的 PSI 均接近 1 或更低(最低值为 0.28)。SARIMAX 基线方案在常态化防控至过渡期(normal→transition)的 PSI 仅为 0.38，显示出较强的稳定性。SARIMAX 干预方案和 Prophet 干预方案的 REC 值均为 1，恢复时间最短，仅需 1 个月。相比之下，LSTM 干预方案的 REC 值为 0.33，恢复时间为 3 个月。基线模型中，LSTM 基线方案的恢复能力最强，REC 值为 1，恢复时间仅为 1 个月，而 SARIMAX 和 Prophet 基线方案的 REC 值均为 0.33，恢复时间为 3 个月。为验证 PSI 和 REC 指标的有效性，本研究进行了 3 项补充分析。(1)独立性验证：Spearman 相关分析显示，PSI 与阶段 MAPE 几乎无相关(ρ=0.098，P=0.854)，REC 与基线 MAPE 呈弱正相关(ρ=0.490，P=0.106)。(2)稳健性验证：在 9 种不同参数组合(3 种窗口偏移×3 种恢复阈值)下，大部分模型的 PSI 和 REC 排名标准差均小于 2.0，仅 Prophet 基线方案的 PSI 排名波动较大(标准差=2.29)。(3)构念效度：指标结果与已知模型特性一致，PSI 与 REC 具有较好的独立性与稳健性。

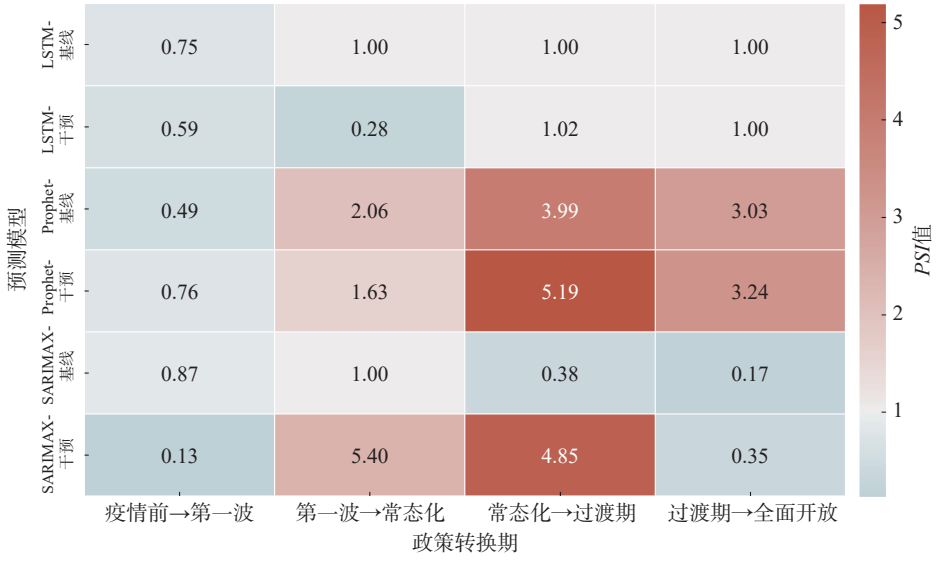


图 4　不同模型阶段敏感度指数(PSI)分布

Fig. 4　PSI distribution of different models across transition periods

3 讨　论

本研究基于中国县域数据，揭示了 COVID-19 疫情期间的非药物干预措施(NPIs)对手足口病传播的阶段性影响，并系统评估了 3 种主流预测模型在外部强干预背景下的适应性。研究发现，疫情防控措施实施期间手足口病发病率降低，但在措施放开后又迅速反弹，提示非药物干预措施对疾病传播存在阶段依赖性和后期风险反弹效应。这一现象在其他研究中也得到证实，表明非药物干预措施对疾病传播存在阶段依赖性和后期风险反弹效应[7,13,25]。

本研究观察到的长期趋势变化，需结合中国手足口病流行病学的动态背景进行解释。疫苗接种和病原谱的变迁是影响手足口病长期趋势的两个核心驱动力。首先，本研究观察到的疫情前(2014—2019 年)发病率呈下降趋势，与 2016 年起在中国获批并广泛使用的 EV-A71 疫苗密切相关，该疫苗的推广有效控制了主要的重症病原体，降低了总体发病水平[26]。EV-A71 的有效控制为其他血清型病毒(如 CV-A6 和 CV-A10)腾出了“生态位”，导致病原谱发生根本性转变。到本研究期末(2022—2023 年)，CV-A6 已成为中国手足口病的优势病原体[27]。因此，2023 年观察到的发病率强劲反弹，可能是 NPIs 解除后“免疫债务”的释放[28]，以及优势毒株 CV-A6 自身流行周期(如“隔年高发”模式)[29] 共同作用的结果。本研究表明，非药物干预(NPIs)作为社会层面的干预措施，可在一定时间内有效改变传染病的传播轨迹。因此，本研究的核心贡献，正是在于量化了这一重大公共卫生干预的阶段性影响，并评估了不同预测模型在这种复杂背景下捕捉突发波动的能力，对于后疫情时代建立更具韧性的传染病预警系统具有现实意义。

本研究结果表明，若模型训练未包含 2020—2022 年疫情期间的数据与干预信息，基线 SARIMAX 模型对 2023 年发病率的外推误差高于干预方案。SARIMAX 模型在纳入疫情干预变量后表现出优于基线方案的预测精度和稳定性，尤其是在疫情防控措施实施与解除的阶段，预测误差降低。这种表现与以往研究的结论相一致，即 SARIMAX 模型能够较好地同时捕获数据的长期趋势、周期性和外源干预因素的影响，表现出良好的稳健性和可解释性，适合公共卫生事件中的长期动态监测[12,30–31]。在部分传染病预测研究中，Prophet 的加性分解与自定义“节假日”回归器机制被证实可灵活应对多种突发干预情形[18,32]。但在本研究中 Prophet 模型则在预测过程中存在明显不足，特别是在疫情政策变动剧烈阶段预测误差增加，这可能与该模型内在的加性分解结构有关[18]。Prophet 模型对趋势的变化点虽能进行分段拟合，但对于疫情所导致的急剧变化点和异常波动，导致预测误差的放大。与上述统计模型相比，LSTM 等深度学习方法通过捕捉数据的非线性与长期依赖关系，在应对极端波动方面展现出一定优势，特别是在 COVID-19 及其他疫情预测中表现优于传统统计模型[33–35]。本研究结果显示，LSTM 模型在 2020—2022 年疫情防控阶段的拟合与预测表现尤为出色，误差低于其他模型，这主要归功于其门控单元机制(gate mechanism)能够有效捕捉短期异常事件及剧烈变化带来的复杂动态特征[20]。但在疫情防控措施放开后的长期预测阶段，LSTM 模型的误差明显上升，可能是由于其对数据规模和分布稳定性要求较高，在小样本、高

变异性或阶段性剧烈波动的数据中，易出现过度拟合或泛化能力不足的问题，从而影响预测精度。

为进一步探究模型的动态特性，本研究构建的 PSI 与 REC 指标从响应敏感度和恢复弹性两个新维度提供了洞见。本研究补充的验证分析证实了这两个指标的有效性：二种指标与传统精度指标（MAPE）无统计学相关性，表明其提供了关于模型误差的动态结构而非仅仅是误差大小的独特信息。这些指标揭示了不同模型在应对外部冲击时的内在特性与实践中的权衡。这一路径与跨学科韧性研究中采用“扰动—恢复”过程及恢复斜率/速率来度量系统韧性的思路一致[23–24]。这些指标比较了不同模型在应对外部冲击时的内在特性与实践中的权衡。例如，Prophet 和 SARIMAX 等统计模型表现出较高的 PSI 值，反映了对外部干预变量和结构变化点的高敏感性。这种特性在需要快速捕捉政策转向信号时是优势，但也可能导致预测过度波动，产生“误报”风险。这与 Prophet 的变化点机制（对趋势断点施加稀疏先验以允许趋势突变）及 ARIMA/干预分析在检测结构突变/外生冲击方面的经典属性相一致[19]。相比之下，LSTM 模型的 PSI 值普遍更接近 1，响应更为平稳，但在冲击后的恢复弹性上表现不佳。这一现象与近年来关于概念漂移/分布迁移下深度时序模型泛化受限的认识相吻合，在训练数据未覆盖的突发情形下（如政策骤变导致的接触模式改变），深度模型可能出现适应迟滞，需额外的自适应或在线更新策略来提升韧性[36]。这些评估结果表明，在传染病预警实践中不存在单一“最优”模型，应根据预警（高敏感性）或稳健预测（高恢复力）等不同目标，策略性地选择或整合多模型。

但本研究也存在一定局限性。首先，虽然手足口病作为症状明显的儿童常见病，其报告相对完整，但不能完全排除疫情期间因就诊意愿下降等因素导致的漏报。然而，本研究关注的是模型对干预措施的相对响应而非绝对预测，这种设计在一定程度上减少了漏报对结果的影响。未来研究可考虑结合医院就诊量、药品销售等多源数据予以交叉验证。其次，本研究对干预仅做了阶段性划分，并未进一步量化防控措施的强度差异（如居家隔离、社交距离限制、公共场所关闭、加强手卫生、环境消毒、健康监测等）。NPIs 的量化仅采用阶段性哑变量，无法识别具体措施的独立贡献。未来研究可考虑构建综合强度指数或利用大数据技术获得更细颗粒度的干预信息，更细化的干预指数或社会活动指标将有助于提高模型解释力。最后，研究基于安徽省凤阳县单一县域的数据，鉴于中国不同地区在经济发展水平、人口密度、气候条件、医疗资源可及性等方面存在显著异质性，本研究的具体数值结果不宜直接外推至其他地区。然而，本研究的核心价值在于提供了一套可复制的方法学框架，这些方法论贡献具有普适性，可为其他地区开展本地化研究提供参考。未来研究应当在多个地理单元开展验证，通过 meta 分析等方法综合不同地区的证据，以获得更具普遍性的结论。

数据可用性声明 本研究使用的手足口病月度发病率数据来源于中国疾病预防控制信息系统（传染病监测系统），属于第三方业务数据库，受数据共享协议与隐私保护要求限制，未对公众开放（该系统为内部业务平台，无公开下载入口）。依据数据提供方管理规定，原始监测数据不可公开。

参考文献

[1] Geng M J, Zhang H Y, Yu L J, et al. Changes in notifiable infectious disease incidence in China during the COVID-19 pandemic[J]. Nature Communications, 2021, 12(1): 6923.

[2] Shen L, Sun M, Song S, et al. The impact of anti–COVID-19 non-pharmaceutical interventions on hand, foot, and mouth disease–A spatiotemporal perspective in Xi′an, northwestern China[J]. Journal of Medical Virology, 2022, 94(7): 3121-3132.

[3] Hu C Y, Tang Y W, Su Q M, et al. Public health measures during the COVID-19 pandemic reduce the spread of other respiratory infectious diseases[J]. Frontiers in Public Health, 2021, 9: 771638.

[4] Wang J, Hu T, Sun D P, et al. Epidemiological characteristics of hand, foot, and mouth disease in Shandong, China, 2009–2016[J]. Scientific Reports, 2017, 7(1): 8900.

[5] Huang Z H, Wang M Z, Qiu L X, et al. Seasonality of the transmissibility of hand, foot and mouth disease: a modelling study in Xiamen city, China[J]. Epidemiology and Infection, 2019, 147: e327.

[6] Ding Z Y, Lu Q B, Wu H C, et al. Trend of hand, foot and mouth disease before, during, and after China′s COVID control policies in Zhejiang, China[J]. Frontiers in Public Health, 2024, 12: 1472944.

[7] Wu K Y, Ma X W, Liu H M, et al. Effects of different levels of non-pharmaceutical interventions on hand, foot and mouth disease in Guangzhou, China[J]. BMC Public Health, 2022, 22(1): 2398.

[8] Etu E E, Monplaisir L, Masoud S, et al. A comparison of univariate and multivariate forecasting models predicting emergency department patient arrivals during the COVID-19 pandemic[J]. Healthcare, 2022, 10(6): 1120.

[9] Mann S, Yadav D, Muthusamy S, et al. A novel method for prediction and analysis of COVID-19 transmission using machine learning based time series models[J]. Wireless Personal Communications, 2023, 133(3): 1935-1961.

[10] Dias E, Diniz A M A, Souto G R, et al. Predicting COVID-19 cases in Belo Horizonte–Brazil taking into account mobility and vaccination issues[J]. PLoS One, 2024, 19(2): e0269515.

[11] Islam M S, Shahrear P, Saha G, et al. Mathematical analysis and prediction of future outbreak of dengue on time-varying contact rate using machine learning approach[J]. Computers in Biology and Medicine, 2024, 178: 108707.

[12] Feng L Z, Zhang T, Wang Q, et al. Impact of COVID-19 outbreaks and interventions on influenza in China and the United States[J].

Nature Communications, 2021, 12(1): 3249.

[13] Oh K B, Doherty T M, Vetter V, et al. Lifting non-pharmaceutical interventions following the COVID-19 pandemic–the quiet before the storm?[J]. Expert Review of Vaccines, 2022, 21(11): 1541-1553.

[14] Ohlson J A, Kim S. Linear valuation without OLS: the Theil-Sen estimation approach[J]. Review of Accounting Studies, 2015, 20(1): 395-435.

[15] Bonander C, Humphreys D, Degli Esposti M. Synthetic control methods for the evaluation of single-unit interventions in epidemiology: a tutorial[J]. American Journal of Epidemiology, 2021, 190(12): 2700-2711.

[16] Mills M C, Rüttenauer T. The effect of mandatory COVID-19 certificates on vaccine uptake: synthetic-control modelling of six countries[J]. The Lancet Public Health, 2022, 7(1): E15-E22.

[17] Haug N, Geyrhofer L, Londei A, et al. Ranking the effectiveness of worldwide COVID-19 government interventions[J]. Nature Human Behaviour, 2020, 4(12): 1303-1312.

[18] Hyndman R J, Athanasopoulos G. Forecasting: principles and practice[M]. 2nd ed. Melbourne: OTexts, 2018: 221-296.

[19] Taylor S J, Letham B. Forecasting at scale[J]. The American Statistician, 2018, 72(1): 37-45.

[20] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. Cambridge: MIT Press, 2016: 404-407.

[21] Zhou L L, Zhao P, Wu D D, et al. Time series model for forecasting the number of new admission inpatients[J]. BMC Medical Informatics and Decision Making, 2018, 18(1): 39.

[22] Podesta C, Coleman N, Esmalian A, et al. Quantifying community resilience based on fluctuations in visits to points-of-interest derived from digital trace data[J]. Journal of the Royal Society Interface, 2021, 18(177): 20210158.

[23] Colón-Emeric C, Pieper C F, Schmader K E, et al. Two approaches to classifying and quantifying physical resilience in longitudinal data[J]. The Journals of Gerontology: Series A, 2020, 75(4): 731-738.

[24] Tang J Q, Han S, Wang J, et al. A comparative analysis of performance-based resilience metrics via a quantitative-qualitative combined approach: are we measuring the same thing?[J]. International Journal of Disaster Risk Science, 2023, 14(5): 736-750.

[25] Haque S, Lambert S B, Mengersen K, et al. Assessing the impact of non-pharmaceutical interventions against COVID-19 on 64 notifiable infectious diseases in Australia: a Bayesian Structural Time Series model[J]. Journal of Infection and Public Health, 2025, 18(3): 102679.

[26] Xiao J P, Zhu Q, Yang F, et al. The impact of enterovirus A71 vaccination program on hand, foot, and mouth disease in Guangdong, China: a longitudinal surveillance study[J]. Journal of Infection, 2022, 85(4): 428-435.

[27] Chen Y, Chen S H, Shen Y F, et al. Molecular evolutionary dynamics of coxsackievirus A6 causing hand, foot, and mouth disease from 2021 to 2023 in China: genomic epidemiology study[J]. JMIR Public Health and Surveillance, 2024, 10: e59604.

[28] Cohen R, Ashman M, Taha M K, et al. Pediatric infectious disease group (GPIP) position paper on the immune debt of the COVID-19 pandemic in childhood, how can we fill the immunity gap?[J]. Infectious Diseases Now, 2021, 51(5): 418-423.

[29] Li F J, Wang D, Su F, et al. Epidemiological, etiological, and serological characteristics of hand, foot, and mouth disease in Guizhou Province, Southwest China, from 2008 to 2023[J]. PLoS Neglected Tropical Diseases, 2025, 19(8): e0013394.

[30] Mulla S, Pande C B, Singh S K. Times series forecasting of monthly rainfall using seasonal auto regressive integrated moving average with EXogenous variables (SARIMAX) model[J]. Water Resources Management, 2024, 38(6): 1825-1846.

[31] Elamin N, Fukushige M. Modeling and forecasting hourly electricity demand by SARIMAX with interactions[J]. Energy, 2018, 165: 257-268.

[32] Chen Z, Zhao Y L, Pan X Y, et al. An overview of prophet[C]//9th International Conference on Algorithms and Architectures for Parallel Processing. Taipei, China: Springer, 2009: 396-407.

[33] Chandra R, Jain A, Singh Chauhan D. Deep learning via LSTM models for COVID-19 infection forecasting in India[J]. PLoS One, 2022, 17(1): e0262708.

[34] Chimmula V K R, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks[J]. Chaos, Solitons and Fractals, 2020, 135: 109864.

[35] Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM[J]. Chaos, Solitons and Fractals, 2020, 140: 110212.

[36] Yuan L H, Li H, Xia B H, et al. Recent advances in concept drift adaptation methods for deep learning[C]//Proceedings of the 31st International Joint Conference on Artificial Intelligence. Vienna: International Joint Conferences on Artificial Intelligence Organization, 2022: 5654-5661.