多源数据融合的传染病智慧监测模型的应用

孟兰
苏州市吴中区郭巷街道社区卫生服务中心　江苏 苏州　215000

摘　要　目的：探讨多源数据融合的传染病智慧监测模型在基层街道传染病监测中的应用效果。方法：选取2022年1月至2023年12月苏州市吴中区郭巷街道社区卫生服务中心监测某街道传染病数据，将2022年1月至2022年12月采用传统监测方法的数据作为对照组，2023年1月至2023年12月应用智慧监测模型的数据作为观察组。比较两组监测及时性、准确性、预警效率及疫情响应时间。结果：观察组病例发现时间、疫情确认时间、响应启动时间均短于对照组，监测准确率、预警灵敏度、数据完整性均高于对照组（$P<0.05$）。观察组监测病例数显著增加，网络自报病例数从204例提升至2 571例，数据捕获能力明显增强。结论：多源数据融合的传染病智慧监测模型能够显著提升传染病监测的及时性和准确性，提高预警效率和病例发现能力，值得在基层推广应用。

关键词　多源数据融合；传染病；智慧监测；预警系统；大数据

Application of an Intelligent Infectious Disease Surveillance Model Based on Multi-source Data Fusion

Meng Lan
Guoxiang Street Community Health Service Center, Wuzhong District, Suzhou 215000, Jiangsu Province, China

Abstract　Objective: To evaluate the effectiveness of a multi-source data fusion-based intelligent infectious disease surveillance model for infectious disease monitoring at the grassroots community level. Methods: Infectious disease surveillance data from a community, monitored by the Guoxiang Street Community Health Service Center in Wuzhong District, Suzhou City, from January 2022 to December 2023, were selected for this study. Data from January 2022 to December 2022, collected using traditional surveillance methods, served as the control group. Data from January 2023 to December 2023, collected after the implementation of the intelligent surveillance model, served as the observation group. The two groups were compared in terms of surveillance timeliness, accuracy, early warning efficiency, and epidemic response time. Results: The observation group demonstrated significantly shorter times for case detection, outbreak confirmation, and response initiation compared to the control group. Furthermore, the observation group showed higher rates of surveillance accuracy, early warning sensitivity, and data integrity (all $P<0.05$). The number of monitored cases increased substantially in the observation group, with self-reported cases via the network rising from 204 to 2,571, indicating a significant enhancement in data capture capability. Conclusion: The multi-source data fusion-based intelligent infectious disease surveillance model can significantly improve the timeliness and accuracy of infectious disease surveillance, enhance early warning efficiency, and strengthen case detection capabilities. This model is worthy of promotion and application at the grassroots level.

Keywords　Multi-Source Data Fusion; Infectious Diseases; Intelligent Monitoring; Early Warning Systems; Big Data

传染病监测是公共卫生体系的重要组成部分，对于及时发现疫情、控制传染病传播具有关键作用。传统传染病监测主要依赖医疗机构被动报告，存在信息滞后、数据孤岛、预警能力不足等问题[1]。随着大数据、人工智能技术的发展，构建智慧化监测体系成为传染病防控的重要方向[2]。多源数据融合技术通过整合医疗机构报告、药店销售、网络搜索、社交媒体等多维度数据，结合机器学习算法建立预测模型，可实现传染病的早期识别和精准预警[3]。本研究以某区某街道为研究对象，评估多源数据融合的传染病智慧监测模型在基层实际应用中的效果，为提升基层传染病监测能力提供科学依据。

1　资料与方法

1.1 一般资料

选取2022年1月至2023年12月苏州市吴中区郭巷街道社区卫生服务中心监测某街道传染病数据进行前后对照研究。对照组为2022年1月至2022年12月采用传统监测方法监测的数据，共监测传染病病例900例，包括流感270例、手足口病196例、水痘133例、梅毒70例、肺结核43例、新冠39例、淋病23例、腮腺炎15例、乙肝16例、尖锐湿疣13例、其他感染性腹泻71例、丙肝4

例、生殖器疱疹4例、HIV 2例、布鲁氏菌病1例；涉及社区卫生服务中心1个，社区卫生服务站5个，监测覆盖人口16.83万人，全年网络自报病例204例，占总病例数22.7%。观察组为2023年1月至2023年12月应用多源数据融合智慧监测模型监测的数据，共监测传染病病例7 706例，包括流感5 455例、手足口病1 159例、新冠530例、水痘152例、乙肝104例、梅毒90例、肺结核65例、其他感染性腹泻63例、淋病22例、生殖器疱疹21例、尖锐湿疣16例、丙肝8例、腮腺炎8例、HIV 4例、戊肝3例、甲肝2例、百日咳2例、血小板减少综合征1例、结膜炎1例；监测范围同对照组，监测覆盖人口16.97万，全年网络自报病例2 571例，占总病例数33.4%。两个时间段的监测区域、人口基数、医疗机构配置基本一致，具有可比性。本研究已签署知情同意书，并通过伦理委员会审查批准。

纳入标准：①符合国家传染病诊断标准的确诊病例；②病例信息完整，包括发病时间、就诊时间、报告时间等关键数据；③在某街道范围内发病并就诊的病例。排除标准：①诊断信息不明确的疑似病例；②关键时间节点数据缺失的病例；③外地输入但未在本街道就诊的病例；④重复报告的病例。

1.2 方法

对照组：采用传统传染病监测方法。医疗机构通过国家传染病报告信息管理系统进行网络直报，区疾控中心工作人员每日审核报告卡，进行疫情分析和风险评估。监测数据主要来源于社区卫生服务中心和服务站的诊断报告，采用人工审核和简单统计分析方法，通过病例数增长趋势判断疫情风险。

观察组：应用多源数据融合的传染病智慧监测模型。该模型整合五大数据源：①医疗机构传染病报告系统数据，包括社区卫生服务中心和服务站的实时报告；②医院门急诊电子病历数据，捕捉就诊患者的症状和诊断信息；③药店感冒退烧药品销售数据，监测辖区内药店相关药品销售变化趋势；④互联网搜索引擎传染病相关关键词搜索量数据，分析居民健康关注度；⑤社交媒体症状自报数据，收集居民主动报告的疾病症状信息。

数据融合采用基于时空特征的多层级融合架构。首先，对各数据源进行标准化处理和质量控制，建立统一的时空索引体系，确保不同来源数据的可比性和准确性。其次，采用动态加权融合算法整合多源数据，权重系数根据数据源的时效性、准确性和完整性进行动态调整，历史数据分析显示医疗机构报告数据权重为0.35，药店销售数据权重为0.25，网络搜索数据权重为0.20，社交媒体数据权重为0.15，电子病历数据权重为0.05。最后，构建基于长短期记忆网络（LSTM）的深度学习预测模型，结合街道历史疫情规律、季节性因素、人口流动特征进行传染病趋势预测。智慧监测系统设置3级预警阈值：当模型预测发病率超过历史同期均值1.5倍时启动蓝色预警，提示可能出现疫情苗头；超过2倍时启动黄色预警，提示需要加强监测和准备应对措施；超过3倍或出现聚集性疫情时启动红色预警，提示需要立即采取应急响应措施。系统实现24 h自动监测，每2 h更新一次数据，异常信号自动触发预警信息，通过短信、App推送等方式及时通知疾控中心相关人员和社区卫生服务机构负责人。

1.3 观察指标

①病例监测指标：比较两组监测病例总数、各类传染病构成比、网络自报病例数及占比，评估监测覆盖范围和数据捕获能力。②监测及时性指标：病例发现时间，即从患者发病到系统识别该病例的平均时间；疫情确认时间，即从首发病例出现到确认疫情的平均时间；响应启动时间，即从确认疫情到启动应急响应的平均时间。③监测准确性指标：监测准确率，即系统预警的疫情中真实发生疫情的比例；预警灵敏度，即实际发生的疫情中被系统成功预警的比例；数据完整性，即监测数据中关键信息完整的病例比例。④预警效率指标：提前预警时间，即系统发出预警到疫情高峰期的平均时间间隔；假阳性预警率，即发出预警但未发生明显疫情的比例。

1.4 统计学方法

采用SPSS 26.0统计软件进行数据分析。计量资料以（$\bar{x} \pm s$）表示，组间比较采用独立样本t检验；计数资料以例数或百分率表示，组间比较采用χ^2检验。$P < 0.05$为差异有统计学意义。

2 结果

2.1 两组病例监测情况比较

观察组监测病例总数显著高于对照组（7 706例 VS 900例），网络自报病例数大幅提升（2 571例 VS 204例），网络自报占比从22.7%提高至33.4%。主要传染病构成中，流感病例从270例增至5 455例，手足口病从196例增至1 159例，监测发现能力明显增强，见表1。

表1 两组病例监测情况比较

监测指标	对照组（2022年）	观察组（2023年）
监测病例总数（例）	900	7 706
网络自报病例数（例）	204	571
网络自报占比（%）	22.7	33.4
流感病例数（例）	270	5 455
手足口病病例数（例）	196	1 159
水痘病例数（例）	133	152
新冠病例数（例）	39	530
监测覆盖人口（万人）	16.83	16.97

2.2 两组监测及时性指标比较

与对照组相比，观察组病例发现时间、疫情确认时间、响应启动时间均显著缩短，差异有统计学意义（$P<0.05$），见表2。

表2 两组监测及时性指标比较（$\bar{x}\pm s$，h）

组别	监测时段	n	病例发现时间	疫情确认时间	响应启动时间
对照组	2022年	900	52.3 ± 18.6	28.7 ± 9.4	16.5 ± 5.2
观察组	2023年	7 706	18.4 ± 7.3	12.1 ± 4.6	6.8 ± 2.3
t			45.28	38.92	42.15
P			< 0.05	< 0.05	< 0.05

2.3 两组监测准确性和预警效率指标比较

观察组监测准确率、预警灵敏度、数据完整性均高于对照组，提前预警时间延长，假阳性预警率降低，差异均有统计学意义（$P<0.05$），见表3。

表3 两组监测准确性和预警效率指标比较

组别	监测时段	监测准确率（%）	预警灵敏度（%）	数据完整性（%）	提前预警时间（d）	假阳性预警率（%）
对照组	2022年	76.8	68.5	82.3	3.2 ± 1.5	28.6
观察组	2023年	94.2	91.7	96.8	8.7 ± 2.3	8.3
χ^2/t		18.45	16.73	21.38	15.62	19.27
P		< 0.05	< 0.05	< 0.05	< 0.05	< 0.05

3 讨论

传染病监测是疫情防控的首道防线。传统模式依赖被动报告，存在信息滞后、预警不足等局限。多源数据融合的智慧监测模型通过整合多维数据与人工智能算法，为基层监测提供了新路径[4–5]。

研究结果显示，应用智慧监测模型后，基层传染病监测能力全面提升。监测病例总数从900例增至7 706例，增长8.6倍。主要原因包括：系统拓宽了数据采集渠道，网络自报病例占比从22.7%提高至33.4%；多源数据融合提高了病例发现的敏感性，尤其针对症状轻微或未及时就诊的早期病例，可通过药店销售数据与网络搜索行为提前捕捉疫情信号[6]。流感、手足口病等病例数的显著增加，反映了2023年实际疫情变化，也体现了系统对轻症及隐性病例的发现能力显著增强。监测及时性明显改善：病例发现时间从52.3 h缩短至18.4 h，疫情确认时间从28.7 h缩短至12.1 h，响应启动时间从16.5 h缩短至6.8 h，整体响应链条缩短约30 h。这一提升得益于多源数据的协同效应与实时监测能力——药店销售数据、网络搜索数据等前置指标可比医疗诊断报告提前24 ~ 48 h发出异常信号；系统每2 h更新数据，实现准实时监测，一旦多源信号同步异常即触发预警[7]。

监测准确性显著提高：准确率从76.8%提升至94.2%，预警灵敏度从68.5%提高至91.7%。这依托于多源数据的交叉验证机制与深度学习算法（如LSTM模型）的应用，能够结合历史疫情规律、季节性因素与人口流动特征进行精准预测[8]。数据完整性从82.3%提升至96.8%，得益于自动化采集与质量控制体系。预警时间从3.2 d延长至8.7 d，为基层防控争取了关键窗口期，可用于开展健康宣传、物资储备、重点场所预防性消毒等早期措施[9]。假阳性预警率从28.6%降至8.3%，在提升灵敏度的同时保证了预警可信度。系统具备自动化、可视化与分级预警功能，提升了基层应急响应的精准性与可操作性[10]。

4 结论

综上所述，多源数据融合的传染病智慧监测模型在基层的应用显著提升了监测的及时性、准确性和预警效率，监测病例数增加8.6倍，病例发现时间缩短约34 h，提前预警时间延长5.5 d，为基层传染病精准防控提供了有力的技术支撑。

参考文献

[1]郭玉清,连怡遥,吴彦霖,等.传染病多渠道监测预警数据与模型技术研究进展[J].疾病监测,2025,40(1):30-37.

[2]邬安琪,文泽轩,吴强松,等.基于症状等多源数据的急性呼吸道传染病发病趋势预测模型的构建与评价[J].现代预防医学,2025,52(2):220-226.

[3]张怡盾,叶雯婧,林越东,等.基于跨部门协同的福建省厦门市校园传染病监测预警平台设计与应用[J].疾病监测,2025,40(1):38-43.

[4]陶生聪,段彩华,姚雪,等.基于多源数据融合的呼吸道传染病症状预警研究[J].中国卫生信息管理杂志,2025,22(4):612-617,641.

[5]霍大柱,张婷,李中杰,等.传染病监测预警体系智慧化建设的特点、功能与实施策略[J].疾病监测,2025,40(1):16-22.

[6]薛瑜,张继,徐威,等.气候驱动下基于气象大数据的虫媒传染病预测模型优化研究进展[J].中国现代医药杂志,2025,27(8):33-36.

[7]史瑶,王玉,彭琳,等.基于CiteSpace的我国传染病监测预警现状的可视化分析[J].传染病信息,2025,38(3):273-278.

[8]夏思洋,朱学芳.面向老年人的智慧健康信息服务系统研究——基于多源数据融合技术[J].情报科学,2024,42(4):89-97

[9]刘民,梁万年,胡健,等.重大突发传染病智能化主动监测预警系统设计研究[J].中国工程科学,2024,26(6):65-76.

[10]杜明梅,刘运喜.我国传染病监测预警系统的发展与应用[J].中华医院感染学杂志,2022,32(6):801-804.

（本文编辑 王 冬）