doi:10. 3969/j.issn. 1672-4623.2023.01.009

河南省新冠肺炎疫情时空聚集性与影响要素研究

任利敏[1]，李　娜[1]，胡世雄[1,2]，洪小瑞[3]，吕玮璇[1]
（1. 黄河交通学院 交通工程学院，河南 焦作 454950；
2. Department of History and Geography, East Stroudsburg University of Pennsylvania, PA 18360；
3. 河南大学规划设计有限公司，河南 开封 475000）

摘　要：以河南省为例，对2020年1月21日—2020年3月29日的新冠肺炎确诊病例进行性别、年龄统计，以日发病数为基本单位对其进行时序与特征分析；同时以病例患者所在区、县地理位置为基本单元，采用回顾性时空扫描统计量方法对该时期河南省新冠肺炎疫情进行时空聚集性分析，并利用地理探测器探究了影响疫情的主要自然环境和社会经济要素。结果表明，河南省新冠肺炎疫情时空聚集的重点区域是1月底至2月上旬的南阳、信阳、驻马店部分县区，主要受输入型病例、人口密度因素的影响。因此，建议河南省新冠肺炎疫情防控应当重点关注输入型病例较多、人口密度较大的地区。

关键词：新冠肺炎；时空扫描统计量方法；地理探测器；时空聚集性；疫情影响因素
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）01-0046-04

Research on the Spatio-temporal Aggregation and Influential Factors of COVID-19 in Henan Province

REN Limin[1], LI Na[1], HU Shixiong[1,2], HONG Xiaorui[3], LYU Weixuan[1]
(1. Traffic Engineering College, Huanghe Jiaotong University, Jiaozuo 454950, China;
2. Department of History and Geography, East Stroudsburg University of Pennsylvania, PA 18360, USA;
3. Henan University Planning and Design Co. Ltd., Kaifeng 475000, China)

Abstract: Taking Henan Province for example, we performed the gender and age statistics of confirmed COVID-19 cases from January 21, 2020 to March 29, 2020, and analyzed the time sequence and characteristics of disease with the daily incidence as the basic unit. At the same time, taking the geographical location of confirmed case's district and county as the basic unit, we used the retrospective spatio-temporal scanning statistical method to analyze the spatio-temporal aggregation of COVID-19 epidemic in Henan Province during this period, and used Geodetector to explore the main natural environment and social and economic factors influencing the epidemic. The results indicate that in Henan Province, the key areas for spatio-temporal aggregation of COVID-19 epidemic were part of the counties in Nanyang, Xinyang and Zhumadian from late January to early February. It is mainly affected by imported cases and population density. Therefore, it is suggested that the prevention and control of COVID-19 in Henan Province should focus on imported cases and population density, especially effective prevention and control interventions in areas with more imported cases and high population density.

Key words: COVID-19, spatio-temporal scanning statistical method, Geodetector, spatio-temporal aggregation, influential factors of epidemic

2019年12月，湖北省武汉市出现新型冠状病毒肺炎（以下简称新冠肺炎）。随着疫情的迅速蔓延，我国各地相继出现了新冠肺炎病例。截至2020年3月29日，全国已累计确诊新冠肺炎病例84 250例，河南省与湖北省相邻，累计报告新冠肺炎确诊病例1 276例，位列全国第三位。河南省采取早期防控措施，取得了显著成效，因此研究河南省新冠肺炎疫情对全国疫情防控具有重要的指导意义。

目前，国内外关于新冠肺炎的研究多集中于医学层面和时空特征分析，如蔺茂文[1]等分析了荆州市新冠肺炎疫情数据，比较了本地病例和输入病例的时空分布特征；张态[2]等利用探索性数据分析方法研究了新冠肺炎疫情的时空扩散特征，并进行了人口流动风险评估。针对新冠肺炎疫情时空聚集性及其影响因素的研究较少，本文利用地理信息制图功能和时空分析模型对河南省新冠肺炎数据进行了分析，以探测其时空聚集性，并利用地理探测器研究了疫情的影响因素，为河南省新冠肺炎疫情防疫与控制提供有价值的信息[3]。

收稿日期：2021-07-02。
项目来源：国家自然科学基金资助项目（41674027）；河南省高等学校重点科研资助项目（21A570004）。

1　基本特征统计与分析

1.1　数据来源与属性统计

本文分析了2020年1月21日—2020年3月29日河南省新冠肺炎确诊病例的属性特征，根据病例患者的家庭住址属性，获取病例个案的经纬度信息，每条信息包含患者的姓名、性别、年龄、家庭住址、确诊日期等。新冠肺炎的传播途径主要包括直接传播、气溶胶传播和接触传播。根据参考文献[4]的研究结果，温度、湿度等环境因素的增加可能降低新冠肺炎病毒在人群中的传播速率，因此本文选取与疫情传播相关的输入型病例、人口密度、气温、湿度、降雨量作为河南省新冠肺炎的潜在影响因子。人口密度为各政府网站提供的总人口与总面积的比值，气候数据来源于中国气象数据网。

据统计，自2020年1月21日发现首例新冠肺炎病例至2020年3月29日，河南省共确诊新冠肺炎病例1 276例，除去缺失关键信息的54例，本文最终获取有效病例1 222例。其中，男性病例655例，女性病例567例。本文按年龄段童年（0~6岁）、少年（7~17岁）、青年（18~40岁）、中年（41~65岁）和老年（≥66岁）统计了患者人数[4]，结果如图1所示，可以看出，河南省确诊病例最多的年龄段为中年，占比为48.6%；其次是青年，占比为38.3%；老年的占比为9.4%，童年和少年的占比均低于3%，这主要是由于学生和老年人与外界接触较少，减少了传播感染新冠肺炎的概率，而作为家庭主力的青年与中年人，为了家庭生计需外出工作或购物，与外界长时间接触增加了其感染新冠肺炎的概率。

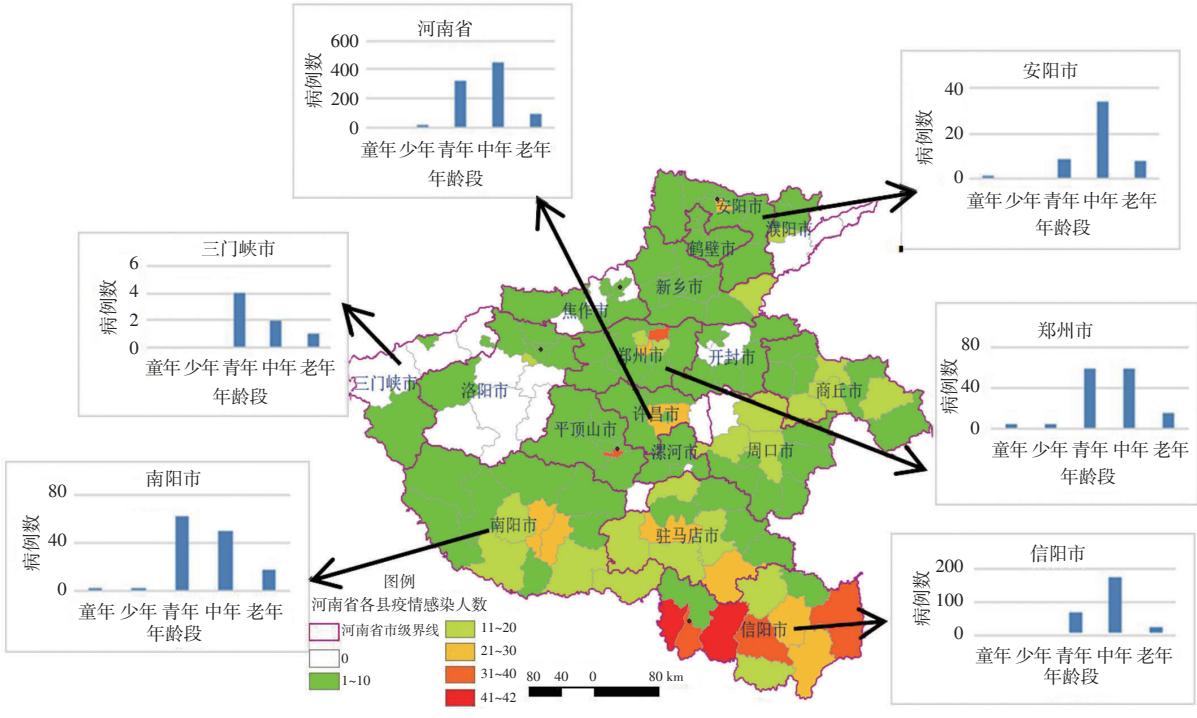


图1　河南省新冠肺炎患者各年龄段数据

（审图号：GS(2020)4814）

1.2　病例时序分析

为分析新冠肺炎疫情的发展趋势，本文统计了河南省2020年1月21日—3月29日每日的确诊病例，结果如图2所示，可以看出，河南省于1月21日发现首例新冠肺炎确诊病例，在2月3日达到最高值109例，随后河南省采取严格疫情防控措施，日新增病例开始波动性下降。随着政府疫情防控措施的实施以及个人防疫意识的提高，新冠肺炎疫情逐渐下降，最终得到有效控制。

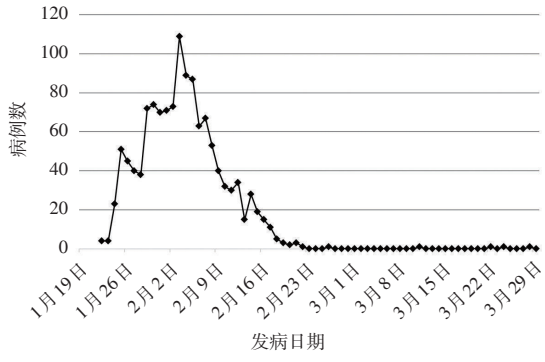


图2　河南省新冠肺炎病例日分布曲线

2 研究方法

根据专家估计，河南省新冠肺炎病例在空间上存在显著的聚集性。早期的流行病聚集性探测主要是从时间上进行，而这种单纯的时间聚集性探测方法存在很大的局限性，即城市某个小区域病例数急剧增加不足以使整个城市的病例数曲线上升，则无法判别出该聚集性。因此，本文采用时空扫描统计量方法从时间和空间两个方面进行聚集性分析，为流行病学研究提供论证基础。

2.1 时空扫描统计量方法

该方法利用扫描窗口内外发病数差异对数似然比来衡量发病的异常程度，计算公式为：

$$\mathrm{LLR}=\log\left\{\left(\frac{c}{n}\right)^{c}\left[\frac{C-c}{C-n}\right]^{(C-c)}\right\} \quad (1)$$

式中，c 为该扫描窗口实际包含的新冠肺炎病例数；n 为该扫描窗口中理论上预期的病例数；C 为研究区域内新冠肺炎的总病例数[5]。

利用蒙特卡罗方法对LLR值进行检验，得到 P 值，若 P <0.05，则该扫描窗口对应的时空范围称为时空聚集区[6-8]。

2.2 地理探测器

地理探测器的核心思想是假设环境因子对于疾病有影响，则其与疾病的空间分布也存在相似性，常被用于研究疾病与环境因子之间的关联性[9]。本文采用地理探测器定量探测不同影响因素及其交互作用对新冠肺炎疫情的影响程度（q），计算公式为[10]：

$$q=1-\frac{1}{N\sigma^{2}}\sum_{h=1}^{L}N_h\sigma_h^{2} \quad (2)$$

式中，h 为类别数；N 和 σ 分别为整体对应的县区数量和新冠肺炎发病率方差；N_h 和 σ_h 分别为类别 h 对应的县区数量和新冠肺炎发病率方差；$0\leqslant q\leqslant 1$[11-12]。

3 数据处理与结果分析

3.1 河南省新冠肺炎疫情时空聚集性

本文对每条病例记录进行地理空间定位，为其赋予空间属性；以河南省区、县地理位置为基本单元，以行政区域的中心点为扫描中心。地理空间定位时，存在21条家庭住址不在河南省境内的记录（均为湖北省人员），对于该类数据，需对患者所在河南省暂时居住地进行地理空间定位。本文统计得到1 222条病例记录（包括14条湖北省有效人员信息），将这些数据或表格导入行政区划矢量文件（shp）的属性表，然后在ArcMap中制作河南省新冠肺炎病例空间分布图（图3）。

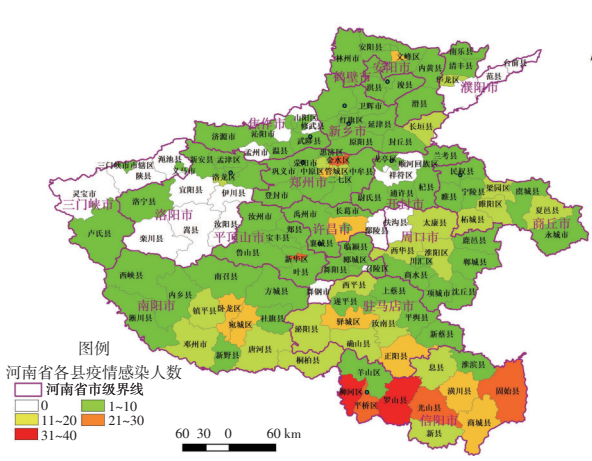


图3 河南省新冠肺炎病例空间分布图
（审图号：GS(2020)4814）

研究发现，新冠肺炎潜伏期的0.95分位点为15.05 d[13]，因此选取15 d作为最大聚集长度；空间扫描半径上限为所覆盖总人数的50%；时间步距为1 d。扫描结果如表1所示，前4个聚集区均发生在1月底至2月初，这与统计的病例日分布曲线基本一致，聚集区1包括南阳市桐柏县，信阳市浉河区、平桥区、罗山县、光山县、新县、潢川县、息县，驻马店驿城区、正阳县、确山县、泌阳县；聚集区2包括焦作市

表1 河南省新冠肺炎疫情回顾性时空扫描结果

聚集区域	圆柱位置（ZIP ID）	覆盖区域代码	聚集半径/km	时间范围	LLR	P
1	411502	411502,411503,411521,411724,411330, 411528,411529, 411725, 411522, 411523, 411526, 411711,411726, 411702	98.92	2020-01-20—2020-02-16	49.67	<0.000 1
2	410804	410804, 410811, 410802, 410871, 410823, 410724, 410882, 410825, 410721, 410773, 410782, 410703, 410702, 410106, 410108	51.21	2020-02-03—2020-02-16	8.16	0.072
3	410402	410402, 470471, 410411, 410403	4.75	2020-02-03—2020-02-09	8.39	0.094
4	410302	410302, 410304	4.28	2020-02-03—2020-02-16	5.18	0.679
5	410502	410502, 410503, 410506, 410505	5.83	2020-01-21—2020-02-23	3.55	0.992

解放区、马村区、山阳区、武陟县、温县，新乡市红旗区、卫滨区、新乡县、获嘉县、辉县以及郑州市上街区、惠济区；聚集区3包括平顶山市新华区、卫东区、湛河区和新城区；聚集区4包括洛阳市老城区、瀍河区；聚集区5包括安阳市文峰区、北关区、殷都区和龙安区。

同时，本文以扫描得到的疫情聚集点为中心，以对数似然比最大窗口半径为缓冲区半径，生成新冠肺炎疫情聚集区图层，并与疫情病例空间分布图进行叠加显示，生成河南省新冠肺炎疫情时空聚集性分布图（图4），图中红色圆圈即为聚集区。

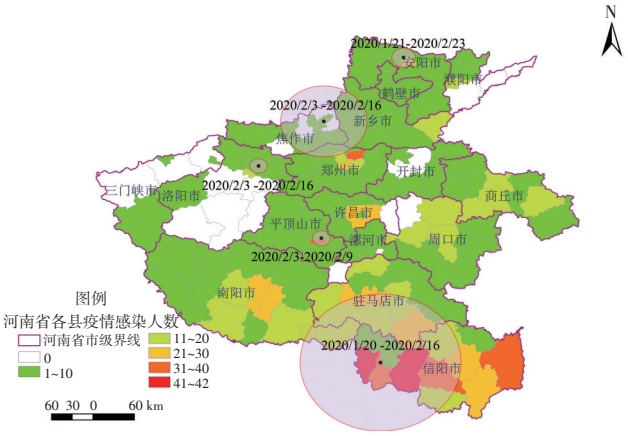


图4　河南省新冠肺炎时空聚集性分布图

（审图号：GS(2020)4814）

3.2　新冠肺炎疫情影响因素

本文利用地理探测器对2020年1月21日—3月29日新冠肺炎病例的发生（Y）进行影响因子分析，影响因子或其代理变量（X）包括人口流动率、输入型病例、人口密度、气温和降雨量。输入数据以及软件运行输出结果如图5所示。因子探测器结果如表2所示，可以看出，输入型病例具有最高值0.719 3，说明输入性病例与河南省新冠肺炎疫情空间聚集性关系最密切，其次是人口密度（0.709 4）、降雨量（0.700 9）和气温（0.694 2）。

交互探测器结果如表3所示，可以看出，影响因素两两交互作用均会增强对河南省新冠肺炎疫情的解释力，输入型病例分别与人口密度、温度、降水量等因素的交互作用共同驱动了河南省新冠肺炎疫情的空间分异，是新冠肺炎疫情的显著控制因子。输入型病例与人口密度、输入型病例与降雨量、气温与人口密度的交互作用 q 值较高，分别能够解释该时期约99%、98%、97%的河南省新冠肺炎疫情。因此，河南省新冠肺炎疫情空间分异主要受输入型病例和人口密度的影响。

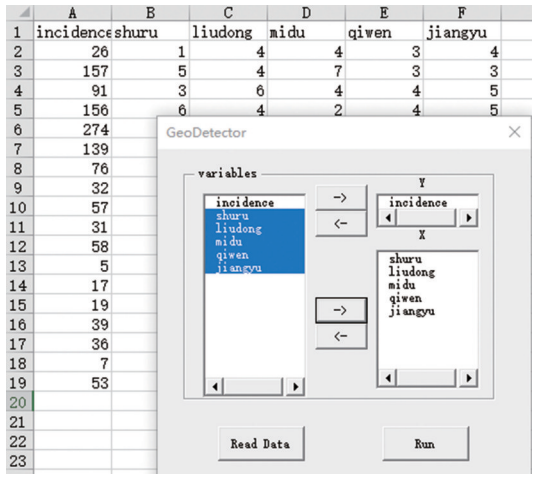


图5　数据输入格式与运行界面

表2　河南省不同变量的风险因子探测

指标	输入型病例	人口流动率	人口密度	气温	降雨量
q 值	0.719 3	0.309 2	0.709 4	0.694 2	0.700 9
p 值	0.280 6	0.642 6	0.290 2	0.305 8	0.236 4

表3　河南省不同变量的交互作用探测

交互作用	q 值
输入型病例∩人口密度	0.992
输入型病例∩降雨量	0.989
气温∩人口密度	0.970

4　结　语

本文统计分析了河南省2020年1月21日—3月29日新冠肺炎病例数据的基本特征，并利用时空扫描统计量方法对其进行了时空聚集性分析；在此基础上，探究了河南省新冠肺炎疫情的主导影响因素。

疫情在性别特征分布上未呈现明显的聚集性，年龄分布特征主要集中在流动性较大的中青年年龄段；疫情前期病例数呈级数增长，随后日新增病例开始波动性下降，这主要得益于政府反应迅速，疫情得到及时控制。因此，在防控期间要严格控制人员流动，并做好各种隔离防护措施。

河南省病例分布存在比较典型的时空聚集性。河南省新冠肺炎疫情空间分异主要受输入型病例和人口密度的影响。与湖北省接壤的南部地区（如南阳、信阳、驻马店部分县区）是最大的疫情集聚区，这与输入型病例较多（如南阳、信阳）密切相关。其他聚集区均与输入型病例（如安阳、郑州）以及人口密度（如郑州、新乡、焦作、平顶山）等密切相关。河南南部地区是高发病聚集县区，是河南省新冠肺炎疫情流行的关键区域，建议河南省新冠肺炎疫情防控重点关注输入型病例较多、人口密度较大的区域并进行有

（下转第100页）