



# 基于 Web 数据挖掘的 COVID-19 流行病学特征分析

郑文<sup>1\*</sup>, 赵 僊<sup>1</sup>, 李泽堃<sup>1</sup>, 武啸泽<sup>1</sup>, 胡 涛<sup>2</sup>

(1. 太原理工大学大数据学院公共安全大数据研究所 太原 030060; 2. 华中科技大学同济医学院附属武汉中心医院肿瘤科 武汉 430014)

**【摘要】**基于 Selenium 数据挖掘技术, 通过对 2020 年 2 月 4 日-22 日新浪微博“肺炎患者求助超话”中 690 例有效病例的分析, 获得了新浪微博中真实求助病例的流行病学特征。研究发现, 求助患者 97.6% 来自于武汉, 重点集中在武昌、硚口、汉阳等中心城区, 与当地的医疗资源和人口密度成正比。微博求助病例主要分布在 2020 年 2 月 4 日-7 日, 随着医疗资源紧张程度的缓解, 通过微博求助的病例明显减少。求助患者确诊日期主要分布在 2020 年 1 月 16 日-2 月 6 日, 与中国疾控中心发布的病例分布情况基本一致。求助患者年龄分布中位数为 60 岁, 明显高于中国疾控中心发布的数据, 但与武汉市中心医院的数据基本吻合。该文研究结果说明, 针对重大突发性传染病, 微博等社交媒体除了在舆论传播上发挥作用, 在流行病学分析上也具有重要意义。基于社交媒体的实时性和广泛性, 结合数据挖掘和大数据分析等方法, 有助于决策层快速掌握一线真实情况。

**关键词** 新冠肺炎; 数据挖掘; 流行病学; 社交媒体

**中图分类号** TP391 **文献标志码** A **doi**:10.12178/1001-0548.2020079

## Epidemiological Characteristics of Novel Coronavirus COVID-19 Based on Web Data Mining

ZHENG Wen<sup>1\*</sup>, ZHAO Cai<sup>1</sup>, LI Ze-kun<sup>1</sup>, WU Xiao-ze<sup>1</sup>, and HU Tao<sup>2</sup>

(1. Institute of Public-Safety and Big Data, College of Data Science, Taiyuan University of Technology Taiyuan 030060;

2. The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology Wuhan 430014)

**Abstract** Based on the Selenium data mining technology, the epidemiological characteristics of real help cases in Sina Weibo were obtained by the analysis of 690 valid cases posted in the Sina Weibo “Pneumonia Patients Asking for Help” topic from February 4 to February 22, 2020. The research showed that 97.6% of the patients seeking for help came from Wuhan, mainly centralized in Wuchang, Tongkou, Hanyang etc. urban areas, and the proportion is directly proportional to the local medical resources and population density. The cases of Weibo help were mainly distributed from February 4 to February 7, 2020. With the relief of medical resources, the number of cases seeking help through social media decreased significantly. The distribution of patients, whose diagnosed date was mainly from January 16 to February 6, 2020, was basically consistent with the case information released by the Chinese Center for Disease Control and Prevention (CCDC). The median age of patients seeking for help was 60 years old, which was much higher than the data released by the CCDC but was roughly coincident with the data of the central hospital of Wuhan. The results of this study indicate that when dealing with major outbreaks of infectious diseases, social media are equally important in epidemiological analysis as well as the role in the dissemination of public opinion. Based on the wide adoption and timeliness nature of social media, it will be helpful for decision-makers to quickly grasp the real-world situation as it is combined with data mining or big data analysis.

**Key words** COVID-19; data mining; epidemiological; social media

自 2019 年 12 月以来, 新型冠状病毒肺炎 (COVID-19) 感染者迅速增加<sup>[1]</sup>, 截至 2020 年 2 月 20 日全国发现感染病例 7 万余例<sup>[2]</sup>。目前, 已有相

关学者和科研机构对新冠肺炎病毒展开研究和分析。我国对新冠肺炎病毒的病原学、流行特征和致病机制等方面已基本掌握<sup>[3-8]</sup>, 这不仅为制定疫情防

收稿日期: 2020-02-24; 修回日期: 2020-03-23; 网络首发日期: 2020-04-15

基金项目: 国家自然科学基金(11702289); 冲击与安全工程教育部重点实验室开放基金(cj201902)

作者简介: 郑文(1985-), 男, 博士, 研究员, 主要从事公共安全大数据、数字政府建设等方面的研究. E-mail: zhengwen@tyut.edu.cn

控策略和措施提供科学依据, 也对全球社区了解新冠病毒提供了重要的参考。在遗传学、病毒学等领域, 研究人员对病毒基因组序列展开研究<sup>[9-12]</sup>, 加速了对追溯病毒源头、预防疾病、研究疫苗等方面的工作进展。在疫情预测和评估方面, 通过建立 SEIR 模型, 加入不同参数进行仿真分析<sup>[13-16]</sup>, 不仅证明了模型分析与疫情发展的真实表现基本吻合, 进一步肯定了对 COVID-19 疫情防控措施的有效性, 对接下来做好疫情防控具有较好的指导价值。令人欣喜的是, 我国在新冠病毒肺炎药物研发方面取得了重大进展<sup>[17-18]</sup>, 多种药物对治疗新型冠状病毒肺炎方面的患者起着积极有效的作用, 加速推进了我国战胜疫情的步伐。

值得注意的是, 以上相关研究大多侧重于新型冠状病毒肺炎的传播模型、流行病学特征、病因病理和治疗护理的某一个方面, 而且大部分数据来源于国家或者地方卫健委官网, 数据来源方式相对单一。最重要的是在疫情初期, 政府无法快速获取一线真实有效数据, 不利于开展疫情防控工作, 也不利于有效措施的实施。而新一代信息技术作为国务院在“十二五”规划中确定的 7 个战略性新兴产业之一, 在传统流行病学研究中的重要意义并未展现。

另一方面, 近年来中国社交媒体的发展引人注目<sup>[19]</sup>, 相继出现了微信、微博、抖音等社交软件。社交媒体软件借助互联网的平台涵盖了以人类社交为核心的所有网络服务形式, 助力互联网从研究部门、学校、政府、商业应用等平台扩展到每一个人, 同时也造成社交数据的爆炸式增长。社交媒体数据之间存在大量信息和知识, 而且可以广泛用于各种应用场景, 包括商务管理、生产控制、工程设计、市场分析和科学探索等。但是如何有效地提取并利用这些信息成为一个巨大的挑战。为了解决这一问题, 定向抓取相关网页资源的网络数据挖掘技术应运而生。网络数据挖掘可以通过程序或者脚本, 按照一定的规则, 自动地抓取万维网信息, 实现对相关网页的数据资源分析。目前数据挖掘技术广泛应用在电力、经济、通信、民生等领域<sup>[20]</sup>。

本文通过 Selenium 数据挖掘技术, 从社交媒体(新浪微博)上获得有效新型冠状病毒肺炎求助病例信息 690 例。随后利用该数据集, 分别从求助患者的地理空间、求助人数、确诊日期和患病人员年龄 4 个方面, 对新型冠状病毒肺炎的流行病学特征进行分析, 最后结合社交媒体的广泛性、实时性特征, 对发生重大传染性疾病期间, 管理部门的各

项政策效果进行了讨论和评估。

## 1 资料与方法

本文主要是基于开源的 Web 应用程序 Selenium 测试工具, 结合 Python 编写的采集程序, 在模拟操作浏览器的情况下, 实现社交媒体工具上新型冠状病毒肺炎病例数据的自动采集。

新型冠状病毒肺炎疫情期间, 新浪微博迅速成为民众了解疫情动态和走向的重要平台。平均每天超过 2 亿网友通过新浪微博关注疫情最新信息, 获取疫情防治服务, 参与公益募捐。微博上的疫情话题数量不断增长, 截至 2020 年 2 月 22 日, 累计 88 万名个人认证用户发布了 1 688 万条微博, 内容涉及内容包括医疗、科普等多个领域。

针对新型冠状病毒肺炎疫情, 新浪微博于 2020 年 2 月 4 日官方发布了“肺炎患者求助超话”, 相关政府部门也设置专门的通道与求助者进行核实和对接。截至 2020 年 2 月 22 日, 该超级话题收集帖子 1 222 个, 关注粉丝 57.5 万, 阅读量超过 29 亿。基于社交媒体数据挖掘方法, 本文选取新浪微博“肺炎患者求助超话”求助专区上发布的求助信息为对象, 研究新型冠状病毒肺炎疫情期间, 社交媒体上求助病例的流行病学特征。其中, 该求助超话包含的求助人员信息字段如表 1 所示。

表 1 求助人员信息字段

序号	字段
1	姓名
2	年龄
3	所在城市
4	所在小区、社区
5	患病时间
6	病情描述
7	联系方式

截至 2020 年 2 月 22 日, 本文通过数据挖掘方法在微博“肺炎患者求助超话”上一共获得有效求助病例数据 690 例, 部分求助人员信息数据如表 2 所示。鉴于本文研究的重点在于新型冠状病毒肺炎病例的流行病学特征分析, 所以求助人员信息主要侧重于“年龄, 所在城市, 所在小区、社区, 患病时间, 求助日期”共 6 个数据段。需要说明的是, 微博超话中的患病时间, 具体指的是患者的确诊时间, 求助患者中绝大部分都给出了详细的确诊证明材料。另外出于保护个人隐私的需要, 本文研究的最终数据集, 隐去了求助者姓名, 具体居住地址、联系方式和详细病情介绍等信息。

表 2 求助人员信息表

姓名	年龄	所在城市	所在小区、社区	患病时间	求助日期
程×	25	武汉市	武昌区苑路梅苑	1/24	2/5
鲍××	85	武汉市	青山区江南春城	1/20	2/5
毕××	57	武汉市	青山区现代花园	1/18	2/5
卜××	57	武汉市	硚口区赛达国际	1/27	2/5
⋮	⋮	⋮	⋮	⋮	⋮
张×	35	武汉市	洪山区北洋桥鑫园	1/23	2/5
李×	37	武汉市	东湖区江南家园	2/1	2/5
蔡××	54	武汉市	武昌区杨园欧景苑	1/30	2/5

## 2 结 果

### 2.1 求助患者区域分布

基于社交媒体数据挖掘获得的 690 例有效求助患者的详细区域统计数据如表 3 所示。从表 3 可以看出,虽然“肺炎患者求助超话”是面向所有公众开放的,但最终的统计数据表明,来源于湖北省以外的仅 4 例,来源于湖北省内,但不属于武汉市的仅 12 例。绝大部分 (97.6%) 的求助患者来源于武汉市。另一方面,社交媒体求助患者的数量在不同区域具有明显的差异。武昌、江汉、汉阳、硚口、江岸、洪山是求助患者的密集区,而新洲,江夏等区求助患者人数较少。

表 3 武汉各地区求助数量

编号	区域名称	求助人数	面积/km <sup>2</sup>	常住人口/万人	人口密度
1	省外	4	/	/	/
2	省内	12	/	/	/
3	江汉	55	28.29	68.00	2.403 6
4	江岸	118	70.25	100.00	1.423 4
5	硚口	101	41.46	62.00	1.495 4
6	武昌	120	107.76	126.37	1.172 7
7	洪山	106	220.50	117.16	0.531 3
8	青山	44	80.47	54.00	0.671 0
9	汉阳	86	111.54	65.27	0.585 1
10	黄陂	8	2 256.70	113.32	0.050 2
11	东西湖	15	499.71	58.48	0.117 0
12	蔡甸	4	1 093.57	46.66	0.042 6
13	新洲	7	1 500.66	105.00	0.069 9
14	江夏	10	2 018.30	91.37	0.045 2

为了进一步地定量分析求助患者的地理分布信息,表 3 还记录了武汉市各区域的地理面积和常驻人口两项数据,并且通过定义常驻人口和地理面积的比表示人口密度。通过对比武汉市各区域人口密度的分布情况可以发现,除了江汉区因为人口密度特别大,数据表现异常之外,其他各区域求助患者人数和该区域人口密度表现出明显的正相关关系。

在中心城区等人口密集区域,是求助患者分布

的主要来源。同时,这个结果也从侧面说明,在重大传染性疫情疫情期间,人口密集区域,医疗资源、救助力量各方面都比较紧张,满足不了求助患者数量的需求,容易造成应急处置效率低等情况。而在人口相对稀少区域,医疗资源可满足大部分求助人群的需求,因此求助患者人数少。所以,医疗人员和收治床位等医疗资源的增加会减少求助患者的数量。

### 2.2 求助患者时序分析

基于社交媒体数据挖掘获得的 690 例有效求助患者的数据,每日患者求助数量随时间变化的分布如图 1 所示。

由图 1 可知,患者求助主要分布在 2020 年 2 月 4 日-7 日这个时间段,这段时间内平均每日的患者求助数量都超过 100 次,其中数量最高峰出现在 2 月 5 日,数量接近 200 次。另一方面,从 2020 年 2 月 8 日开始,肺炎患者求助数量急剧下降,平均每日不超过 20 次。

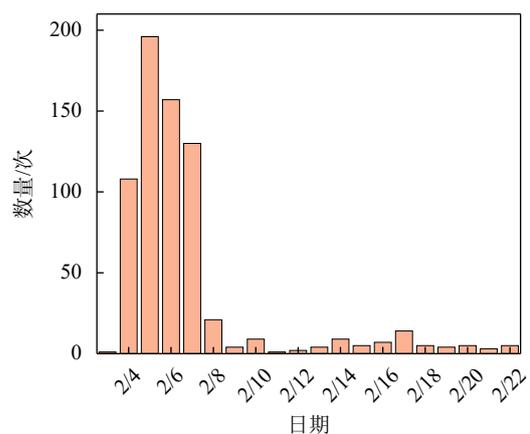


图 1 求助患者数量随时间分布

为了进一步说明求助患者的分布规律,根据武汉市卫生健康委员会公布的《全市定点医院病床使用情况》<sup>[21]</sup>分析表,获取了武汉市 2020 年 2 月 1 日-22 日期间,全市各定点医院总的开放床位数

和空床位数如图 2 和图 3 所示。

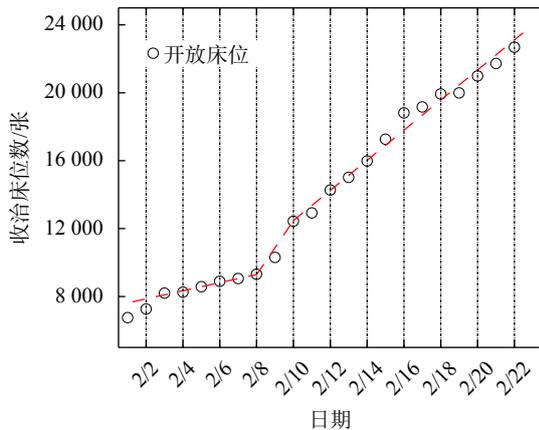


图 2 武汉市收治床位数随时间分布

由图 2 可知, 全市开放床位数在 2020 年 2 月 8 日出现了第一个大的拐点, 这恰好跟图 1 中求助患者急剧减少的日期相对应。进一步的数据分析可以发现, 2 月 8 日, 武汉火神山医院增设床位 200 个, 雷神山医院开始投入运行, 其他各定点医院一共增加床位 255 个。2020 年 2 月 9 日, 火神山、雷神山等定点医院开放床位数进一步增加, 同时方舱医院启用, 全国各地救援医疗队陆续到达武汉加入到疫情防控一线。正是由于医疗资源和医疗队伍得到保障, 社交媒体上的求助患者才出现大幅度的下降, 这也证明火神山、雷神山和方舱医院等定点医院的建设, 在应对突发性传染病方面, 具有非常重要的作用。

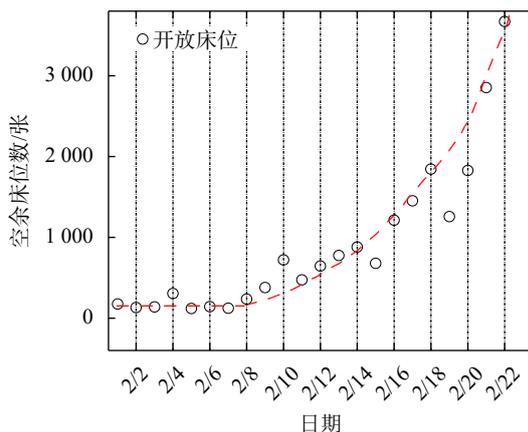


图 3 武汉市空余床位数随时间分布

另一方面, 医疗资源是否满足当前防疫的需要, 最直观的指标就是空床位数, 由图 3 可知, 从 2020 年 2 月 8 日开始, 武汉市定点医院的空床位数才开始增加。虽然在 2 月 8 日以前, 全市空床位数一直保持在 170 个左右, 但这些数字均是由于

出院、转院等因素造成的影响, 真实情况下, 2020 年 2 月 8 日以前, 武汉市各定点医院一直处于超饱和状态运行, 这也是社交媒体上出现大量求助患者的重要原因。

### 2.3 求助患者确诊日期分布

图 4 为本文 690 例有效求助患者的患病确诊时间随日期的分布图, 其中也包括了中国疾控中心发布的确诊人数随日期变化的数据。整体趋势上来看, 通过社交媒体获取的病例的流行病学历史与中国疾控中心发布的数据<sup>[5]</sup>大致吻合, 两者的 Pearson 相关系数达到了 0.925。两个数据均是从 2020 年 1 月 15 日开始, 新型冠状病毒肺炎确诊患者显著增加, 2020 年 1 月 20 日开始出现了一个小跳跃式的增加, 随后在 2020 年 1 月 23 日-28 日达到第一个流行峰, 然后缓慢下降。最后在 2 月 1 日出现一个异常高峰值, 后逐渐下降。进一步定量地分析发现, 以武汉“封城”的 2020 年 1 月 23 日为界限, 在 1 月 23 日以前, 中国疾控中心发布的确认人数分布数据明显高于社交媒体求助患者中的确诊人数分布, 而在 1 月 23 日开始“封城”到 1 月 28 日, 武汉确诊病例明显偏高。一方面是由于“封城”之后, 阻止了感染人群的外流; 另一方面, 可能还是因为人员聚集, 以及前期医疗资源紧张, 居家隔离导致的大面积感染。

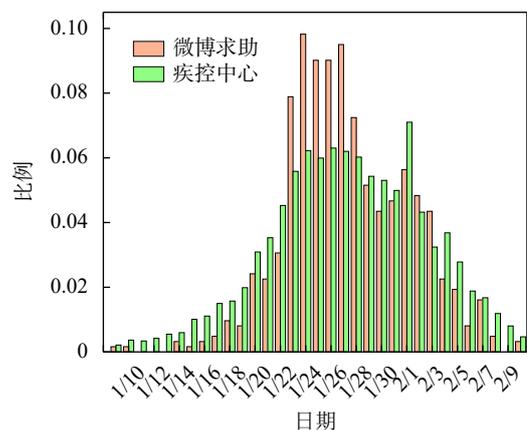


图 4 求助患者确诊日期分布

图 5 是每日求助患者确诊时间分布, 由图可知, 求助患者比较集中的 2020 年 2 月 4 日-7 日, 其确诊时间大多分布在 2020 年 1 月中旬至求助当天日期。由图中蓝色虚线可知, 随着时间的推移, 求助患者中早期确诊的人数越来越少, 这基本符合国家“应收尽收, 刻不容缓”的政策要求, 也说明绝大多数患者都得到有效地救助和安排。

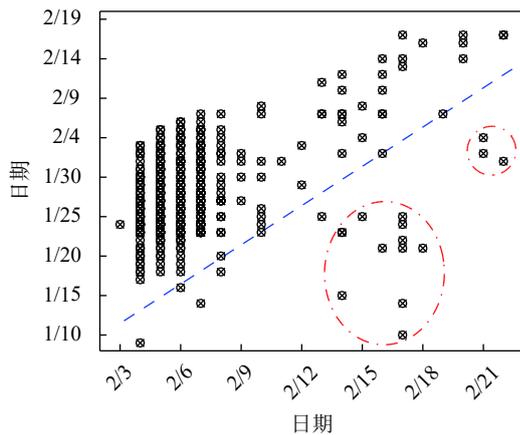


图 5 每日求助患者日期分布

但是,从2020年2月13日开始,由图5中红色椭圆区域可知,出现了一批早期确诊患者的求助信息。通过对红色区域的14例患者信息做进一步分析发现,这部分患者主要可分为两组:一组是已经正常住院求助康复者血浆,另一组是存在其他基础性疾病求助治疗。抗击新型冠状病毒肺炎期间,除了确诊收治病人,另外一个重要的工作就是针对新冠肺炎的新药、新治疗方法的研究。从社交媒体肺炎求助患者的数据可以发现,“血浆疗法”在医疗一线具有一定的影响力。另一方面,重大传染病疫情期间,本身具有其他基础性疾病的人群往往因为免疫力低下而容易被感染。尤其是接受化疗的患者,更是成为新型冠状病毒的易感人群,患者们只能居家监护,用药治疗。加上疫情期间,武汉多家肿瘤医院科被征用抗疫,造成一些患者的化疗时间被耽误,所以这些患者通过社交媒体寻求帮助。

#### 2.4 求助患者年龄分布

图6为社交媒体求助的690例有效患者的年龄分布图。从图中分析可得,求助患者大多数集中在50~80岁(71.88%)年龄段,年龄分布的中位数为60岁。此年龄分布与中国疾控中心<sup>[5]</sup>发布的新型冠状病毒肺炎确诊病例分布特征基本吻合,由图可知,老年患者更容易被新型冠状病毒感染。

图7为社交媒体求助患者、武汉中心医院收治患者和中国疾控中心分别公布的年龄分布图。由图可知,社交媒体求助患者的年龄分布与武汉中心医院的数据曲线走势基本吻合,两者的Pearson相关系数达到了0.914,尤其在60~80岁的老年区间高度吻合,该结果从侧面进一步说明了社交媒体数据的实时性。

另一方面,社交媒体求助患者与中国疾控中心公布曲线有一定的差距,两者的Pearson相关系数仅仅只有0.693。中国疾控中心的数据年龄分布中

位数为41岁,高龄患者整体偏少,这主要是因为中国疾控中心的数据统计人群为全国患者,全国相比于疫情中心武汉,整体医疗资源相对宽松,导致高龄易感人群偏少。

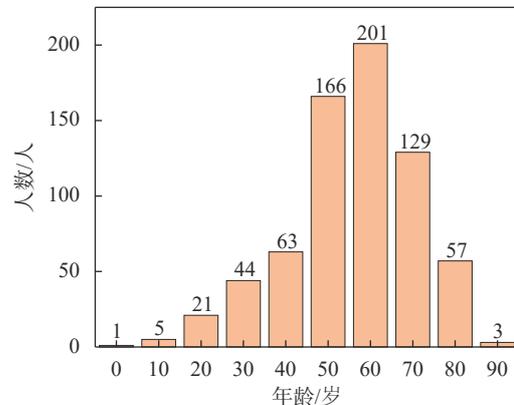


图 6 求助患者年龄分布

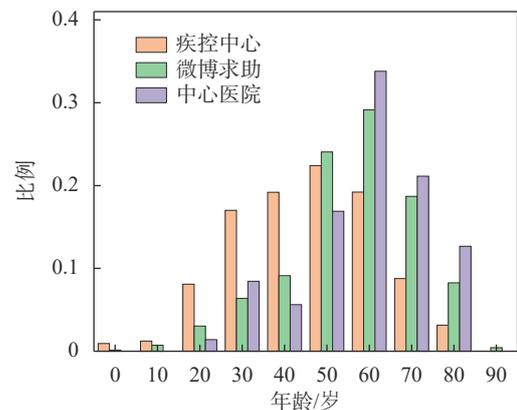


图 7 不同数据来源年龄分布比较

### 3 结束语

本文发现:社交媒体上获得的真实有效病例分析结果说明,新型冠状病毒肺炎的迅速蔓延最主要的原因是医疗资源的紧缺,造成应急处置工作运作低效,引发恐慌。在医疗队、收治床位等医疗资源满足需求以后,求助患者病例明显减少,疫情也得到了有效的遏止。另一方面发现:通过社交媒体对疫情期间的求助患者信息进行可视化呈现和统计分析,能更加有效、及时地获得其流行病学特征。下一步工作中,可以在此基础上引入传染性疾病的传播模型,有效分析疫情的时空演变和扩散特征,为进一步支援武汉、狙击疫情传播提供重要的数据参考。

同时,本文的研究结果说明,在中央有关部门的正确领导下,从2020年1月24日开始,武汉以“中国速度”相继建立的火神山医院、雷神山医院

以及若干方舱医院,增设数千张床位,有效地解决了大量患者无院可住的问题,让广大患者得到妥善的治疗和照顾,这在重大传染性疫情防控方面,具有非常重要的意义。湖北省政府、武汉市政府贯彻落实习近平总书记重要指示精神,不折不扣落实“四类人员”分类集中管理措施,真正做到应收尽收、不漏一人,这在切断传染源、降低传染率和缩短传播时间等方面发挥了重要的作用。

综上所述,针对重大突发性传染疾病,社交媒体不但在舆论宣传上发挥作用,在流行病学分析上同样具有重要的意义。相关技术部门可以充分利用社交媒体的广泛性和及时性,通过社交媒体获得有效的数据病例,然后结合数据挖掘和大数据分析等方法,帮助国家决策部门快速掌握一线的真实情况,有助于政府部门迅速展开疫情防控工作。

感谢太原市大数据应用局乔熙,太原理工大学公共安全大数据研究所廉涛、徐震寰的交流与讨论。

### 参 考 文 献

- [1] 武汉市卫健委. 武汉市卫健委关于当前我市肺炎疫情的情况通报[EB/OL]. (2020-01-18). <http://wjw.wuhan.gov.cn/front/web/showDetail/2020011809064>. opens in new tab. Wuhan Municipal Health Commission. Wuhan municipal health and health commission's briefing on the current pneumonia epidemic situation in our city[EB/OL]. (2020-01-18).<http://wjw.wuhan.gov.cn/front/web/showDetail/2020011809064>. opens in new tab.
- [2] 周娟,李丹,龙云铸. 新型冠状病毒(2019-nCoV)相关研究进展[J]. 中国感染控制杂志, 2020, 19(3): 1-5. ZHOU Juan, LI Dan, LONG Yun-zhu. Advances in related research on novel coronavirus(2019-nCoV)[J]. Chinese Journal of Infection Control, 2020, 19(3): 1-5.
- [3] 韩明慧,方虹霖,杨东见,等. 国外新型冠状病毒肺炎发病的现状和趋势分析[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/31.1635.R.20200225.0951.002.html>. HAN Ming-hui, FANG Hong-ji, YANG Dong-jian, et al. Current situation and trend of coronavirus diseases 2019(COVID-2019) in foreign countries[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/31.1635.R.20200225.0951.002.html>.
- [4] 庄英杰,陈竹,李进,等. 26例新型冠状病毒肺炎确诊病例临床和流行病学特征[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/11.3456.R.20200219.1312.006.html>. ZHUANG Ying-jie, CHEN Zhu, LI Jin, et al. Clinical and epidemiological characteristics of 26 patients diagnosed With novel coronavirus pneumonia[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/11.3456.R.20200219.1312.006.html>.
- [5] 中国疾病预防控制中心新型冠状病毒肺炎应急响应机制流行病学组. 新型冠状病毒肺炎流行病学特征分析[J]. 中国流行病学杂志, 2020, 41(2): 145-151. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China[J]. Chinese Journal of Epidemiology, 2020, 41(2): 145-151.
- [6] HUANG C, WANG Y, LI X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China[J]. *The Lancet*, 2020, 395(10223): 497-506.
- [7] CHEN N, ZHOU M, DONG X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study[J]. *The Lancet*, 2020, 395(10223): 507-513.
- [8] LI Q, GUAN X, WU P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia[J]. *New England Journal of Medicine*, 2020, 382: 1199-1207.
- [9] 李鑫,段广有,张伟,等. 2019新型冠状病毒S蛋白可能存在Furin蛋白酶切位点[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/23.1513.Q.20200224.1809.002.html>. LI Xin, DUAN Guang-you, ZHANG Wei, et al. A furin cleavage site was discovered in the S protein of the 2019 novel coronavirus[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/23.1513.Q.20200224.1809.002.html>.
- [10] 熊子军,张喆,王雅琦,等. 新型冠状病毒基因组变异与诊断[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/61.1399.R.20200224.0933.008.html>. XIONG Zi-jun, ZHANG Zhe, WANG Ya-qi, et al. Genomic variations of SARS-COV-2 and detection [EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/61.1399.R.20200224.0933.008.html>.
- [11] ZHOU P, YANG X L, WANG X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. *Nature*, 2020, 579: 270-273.
- [12] WU F, ZHAO S, YU B, et al. A new coronavirus associated with human respiratory disease in China[J]. *Nature*, 2020, DOI: 10.1038/s41586-020-2008-3.
- [13] 范如国,王奕博,罗明,等. 基于SEIR的新型肺炎传播模型及拐点预测分析[J]. 电子科技大学学报, 2020, DOI: 10.12178/1001-0548-2020029. FAN Ru-guo, WANG Yi-bo, LUO Ming, et al. SEIR-based novel pneumonia transmission model and inflection point prediction analysis[J]. Journal of University of Electronic Science and Technology of China, 2020, DOI: 10.12178/1001-0548-2020029.
- [14] 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[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/33.1248.R.20200303.1722.004.html>. CAO Sheng-li, FENG Pei-hua, SHI Peng-peng. Study on the epidemic development of corona virus disease-19(COVID-19) in Hubei province by a modified SEIR model. [EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/33.1248.R.20200303.1722.004.html>.
- [15] 周涛,刘权辉,杨紫陌,等. 新型冠状病毒感染肺炎基本再生数的初步预测[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/33.1248.R.20200303.1722.004.html>.

- [net/kcms/detail/51.1656.r.20200204.1640.002.html](http://kns.cnki.net/kcms/detail/51.1656.r.20200204.1640.002.html).  
ZHOU Tao, LIU Quan-hui, YANG Zi-mo, et al. Preliminary prediction of the basic reproduction number of the novel coronavirus 2019-nCoV[EB/OL]. [2020-02-21]. <http://kns.cnki.net/kcms/detail/51.1656.r.20200204.1640.002.html>.
- [16] ZHOU T, LIU Q, YANG Z, et al. Preliminary prediction of the basic reproduction number of the novel coronavirus 2019-nCoV[J]. *Journal of Evidence-Based Medicine*, 2020, 13(1): 3-7.
- [17] CHEN L, XIONG J, BAO L, et al. Convalescent plasma as a potential therapy for COVID-19[J]. *The Lancet Infectious Diseases*, 2020, DOI: [10.1016/S1473-3099\(20\)30141-9](https://doi.org/10.1016/S1473-3099(20)30141-9).
- [18] 田成, 向明. 瑞德西韦与干扰素  $\alpha$  联用治疗新型冠状病毒肺炎的可行性分析[J]. *医药导报*, 2020(32): 1-13.  
TIAN Cheng, XIANG Ming. Feasibility analysis of remdesivir combined with interferon- $\alpha$  in the treatment of COVID-19[J]. *Herald of Medicine*, 2020(32): 1-13.
- [19] 郑睿颖, 涂珍, 孟令坤. 浅析社交网络的现状及发展趋势——以微博为例[J]. *电脑迷*, 2018(8): 224.  
ZHENG Rui-ying, TU Zhen, MENG Ling-kun. A brief analysis of the current situation and development trend of social network—taking microblog as an example[J]. *PC Fan*, 2018(8): 224.
- [20] 程增辉, 夏林旭, 刘茂福. 基于 Python 的健康数据数据挖掘设计与实现[J]. *软件导刊*, 2019, 18(2): 66-69.  
CHENG Zeng-hui, XIA Lin-xu, LIU Mao-fu. The Design and implementation of health data crawler using python[J]. *Software Guide*, 2019, 18(2): 66-69.
- [21] 武汉市卫健委. 武汉市卫健委关于全市定点医院病床使用情况[EB/OL]. [2020-02-22]. <http://wjw.wuhan.gov.cn/front/web/list3rd/yes/803>.  
Wuhan Municipal Health Committee. Wuhan municipal health committee on the use of hospital beds in designated hospitals[EB/OL]. [2020-02-22]. <http://wjw.wuhan.gov.cn/front/web/list3rd/yes/803>.

编辑 税红