

本期“新型冠状病毒肺炎疫情防控”专题评述

数学建模和统计分析在流行病趋势预测和精准防控中的作用

周涛^{1*}, 蒋晓²

(1. 电子科技大学大数据研究中心 成都 611731; 2. 电子科技大学学报编辑部 成都 610054)

这次新型冠状病毒肺炎 (COVID-19) 疫情, 是新中国成立以来我国遭遇的传播速度最快、感染范围最广、防控难度最大的公共卫生事件。为了发挥科研工作者在疫情防控中的知识贡献, 《电子科技大学学报》于 2020 年 2 月 1 日针对“新型冠状病毒肺炎疫情防控”进行了专题征稿, 并开辟了快速评审和优先发表的专门通道。本专题短期内收到了 200 余篇投稿, 最终接收发表了其中的 15 篇。这些论文的研究内容覆盖了初期感染人数的估计、早期流行病关键参数的估计、流行病传播建模与趋势预测、防控措施效果的事先预估和事后评估、与疫情相关的社会经济问题分析等方面。

一、初期感染人数的估计。在疾病爆发的初期, 特别是针对未知流行病, 往往缺少成熟有效的检测和确诊的手段, 因此准确估计感染人数对于判断疫情严重程度并做出相应重大决策有关键性的作用。一种通常的方法是根据某区域外流人口中感染的比例来反向估计当地已经感染的人数, 例如早期广受关注的帝国理工大学^[1]和美国东北大学^[2]的报告。当然, 出境感染人员数量非常有限, 以此为依据的感染人数估计误差极大。事实上, 很多研究人员都注意到了, 疫情传播早期我国各地确诊人数和武汉流入当地的人数线性相关^[3-5], 因此可以用中国多个城市早期感染的人数来反推武汉初期的感染人数, 从而估计武汉实际感染人数和已经确诊人数之间的差距^[6]。

二、早期流行病关键参数的估计。基本再生数是指没有干预的情况下, 在全部是易感人群的环境中, 平均一个患者可以传染的人数。这是衡量流行病传播能力最关键的参数之一。我们假设 COVID-19 早期传播可以用一个“易感-潜伏-传染-移除 (SEIR)”动力学模型近似刻画, 且自由传播时呈指数趋势, 估计基本再生数为 2.8-3.9, 后根据最新的流行病学特征参数修正为 2.2-3.0^[7](更多对基本再生数的估计结果请见文献^[8])。随着防控措施的实施, 传染病传播不再是自由传播阶段, 为了评价防控措施的有效性, 通常采用有效再生数对传染病的传播能力进行实时评估。有效再生数 R_t 定义为在某时刻 t , 一个感染者平均而言可以感染其他个体的数量。以前的估计有效再生数的方法需要知道患者出现症状的时间, 陈端兵等人^[9]基于蒙特卡洛方法提出了一种依据少量患者的症状出现时间对有效再生数进行估计的方法, 为及时对防控措施的效果进行定量评估提供了手段。

三、流行病传播建模与趋势预测。在本专题 15 篇论文中, 有近一半 (7 篇) 是关于 COVID-19 的建模和预测。张琳^[10]和梁凯豪等人^[11]直接从数据入手进行拟合, 没有考虑 COVID-19 的传播机制。一般而言, 通过直接的数据拟合进行趋势预测, 工作水平都不高, 预测结果也不太靠谱, 但是张琳^[10]和梁凯豪等人^[11]的结果却和真实情况符合度很高, 而且能够处理一大类相近的流行病。其共同的思路是假设了感染人数的增长率会随着累计感染人数变化, 然后再把这个变化的函数关系带入感染人数增长的微分方程中, 并拟和少量参数。这类方法^[10-11]给我们提出了两个开放性的问题: 一是是否可以通过分析给定感染人数或者感染人口占比后的再生数 R_p (p 是感染人口占比) 得到更好的刻画流行病传播的方式, 二是如果把感染率随着感染人数变化的规律看成刻画流行病传播的某种核函数, 是否不同类型的传播机制可以用若干对应的核函数进行分类和刻画。张琳^[10]和梁凯豪等人^[11]的工作也可以看作是假设了一些关键参数随时间变化的规律, 然后进行拟合。喻孜等人^[12]和梅文娟等人^[13]的工作就是在动力学机制模型中引入了时变参数。这种方法的出发点是合理的, 因为随着疫情防控政策的变化和外界环境的变化 (例如温度等), 流行病的关键动力学

参数可能发生变化。引入时变参数，短期内进行预测的精度也明显高于固定参数或者无机制模型的拟合方法，但同时也作出了巨大牺牲，因为很难用这种方法做长期的预测。一种可能的办法是建立多阶段模型^[14]，在不同传播阶段用不同的参数进行拟合，但不是每天一变——这样得到的模型解释力更强，长期预测能力更强，也许短期预测精度也差不多。研究人员还通过在通用流行病传播动力学模型中引入和 COVID-19 相关的细节，提高模型对真实传播过程的刻画和预测能力，例如范如国等人^[15]考虑了潜伏期对于传播趋势和拐点的影响，林俊峰^[16]考虑了隐性传播者的影响，傅家旗等人^[17]考虑了人口流动的影响，等等。有意思的是，这些基于极早期少量数据的预测结果是比较准确的，例如范如国等人^[15]预测2月19日是武汉现存确诊人数的拐点，结果中国恰好在2月19日出现现在确诊人数第一次下降，只是武汉还要略晚一些；又如林俊峰认为4月底国内传播就基本结束了，累计确诊人数在十万左右，现在看来也是比较准确的。

四、防控措施效果的事先预估和事后评估。通过机制模型，可以模拟一个政策实施与否对于传播范围的影响，进而预估特定政策可能的效果。例如喻孜等人^[12]的时变参数模型显示政府的防控措施带来了有效传染率的大幅度下降，与政府没有出台防控措施（参数保持不变）相比，截至2月1日，实际感染人数下降超过了50%。汪剑眉和李钢^[18]的研究显示，降低接触是抑制疫情所有措施中最关键的，其作用显著高于其他任何措施。结合机制模型和真实数据，还可以对微观具体措施的效果进行预估，例如模拟显示在流感季关闭学校将是的的学生接触人数下降一半以上并整体减少流感感染人数^[19]。孙皓宸等人^[20]分析了中小学学生佩戴RFID后的接触记录，发现绝大部分接触都发生在同班学生间，进一步的仿真结果显示，当仅出现一例确诊学生时，隔离一个班对疫情防控的影响与隔离全校几乎是一致的。数学模型和统计分析工具还可以在事后对防控政策的整体效果进行评估，例如李冀鹏等人^[5]通过消除输入病例的影响，给出了比较不同地区防控效果的定量方法，结果显示了在武汉“封城”前两周流入人口最多的50个城市中，防疫效果最佳和最差的城市。陈端兵等人^[9, 21]通过计算有效生成数，证实了中国防疫的巨大成就——除湖北外所有省份有效生成数都在实施防控措施一周内下降到了1以下。郑文等人^[22]分析了微博中关于COVID-19感染肺炎的求助信息，可以从数据中清楚看到哪些时间和哪些地方医疗资源缺口最大，并通过求助的数量反向评估医疗资源供给的充分程度。

五、与疫情相关的社会经济问题分析。在本次专题的15篇论文中，有两篇是我们特别感兴趣的“非典型论文”，讨论的不是疫情防控本身，而是与疫情有关的社会经济问题。张欣等人^[23]基于2018年版的“中国31省区市区域间投入产出表”，构建了我国区域产业关联网络，并引入网络级联失效模型，模拟了受新冠疫情影响，湖北省产业供给中断和需求下滑，以及全国各省市因疫情防控产能受限和需求下降情况下，停产和流动性不足风险的网络传导效应。有趣的是，四川虽然靠近湖北（人员交流频繁），但因为和湖北产业关联弱，竟然是全国受湖北供给中断和需求下滑冲击影响最小的省份。廖敬仪等人^[24]分析了深圳82个家庭聚集性感染案例，共涉及212例确诊患者。基于统计显著性检验和零模型，廖敬仪等人^[24]证实了女性被家人感染的概率显著高于男性，量化支持了“女性在家庭环境中承担更多护理角色”的假设，并建议要在公共卫生事件中加强性别视角。

本专题收录的工作都是疫情发展早期甚至初期的来稿，部分模型和分析还比较粗糙，但是我们依然能够从这些早期的研究中看出数学模型和基于真实数据的统计分析对于深入理解 COVID-19 本身以及与 COVID-19 爆发相关的社会经济问题的重要作用。未来的相关研究应该更加精细化，具有更强的解释力和预测精度。首先要考虑更广泛和深入的数据，包括多个国家和地区的数据以及记录了具体人与人之间的接触行为的微观数据^[21, 25-26]。特别需要强调的是，了解人口结构、人口流动、家庭规模分布、学校规模分布和企业规模分布等等数据，都是为了再现人与人之间真实的接触模式^[27]——中国的健康码稍加改造就能记录这些数据。国外有些学者对此谈虎色变，认为对个人隐私是一种侵犯，实际上有很多办法可以既保护隐私，又能记录接触行为（这样的例子和最近的研究很多，例如 github 上的 DP-3T 项目）。其次是需要考虑 COVID-19 传播动力学特殊的复杂性，例如可能存在大量隐形的传播^[16, 28]，又如最近的研究显示 SARS 治愈者可能拥有对 COVID-19 的抗体^[29]，因此有必要考虑多种冠状病毒传播之间的交互作用^[30]，再如医疗资源的丰度和基于经济社会综合考量的政策选择也会对 COVID-19 后期的流行程度带来影响^[30-33]。充分考虑这些复杂因素，才能得到更加准确且具有参考价值的结论。

参 考 文 献

- [1] IMAI N, DORIGATTI I, CORI A, et al. Series reports entitled “ Estimating the potential total number of novel coronavirus in Wuhan City, China” [R]. London: Imperial College London, 2020.
- [2] CHINAZZI M, DAVIS J T, GIOANNINI C, et al. Series reports entitled “ Preliminary assessment of the international spreading risk associated with the 2019 Novel Coronavirus (2019-nCoV)” [R]. Boston: Northeastern University, 2020.
- [3] 许小可, 文成, 张光耀, 等. 新冠肺炎爆发前期武汉外流人口的地理去向分布及影响[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020033.
- [4] DU Z, WANG L, CAUCHEMEZ S, et al. Risk for transportation of 2019 Novel Coronavirus Disease from Wuhan to other cities in China[J]. *Emerging Infectious Diseases*, 2020, 26(5): 1049-1052.
- [5] 李冀鹏, 洪峰, 白薇, 等. 评估新型冠状病毒地区防控效果的一种近似方法[J]. *物理学报*, 2020, 69(10): 100201.
- [6] 杨政, 原子霞, 贾祖瑶. 基于迁徙数据估计武汉感染新型冠状病毒的人员数量[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020030.
- [7] ZHOU Tao, LIU Quan-hui, YANG Zi-mo, et al. Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV[J]. *Journal of Evidence-Based Medicine*, 2020, 13: 3-7.
- [8] 栾荣生, 王新, 孙鑫, 等. 新型冠状病毒肺炎的流行病学、临床治疗与疫情防控[J]. *四川大学学报(医学版)*, 2020, 51(2): 131-138.
- [9] 陈端兵, 白薇, 王岩, 等. 新型冠状病毒防控效果的定量评估[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020144.
- [10] 张琳. 新冠肺炎疫情传播的一般增长模型拟合与预测[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020037.
- [11] 梁凯豪, 张文峰, 张小花, 等. 冠状病毒 SARS-CoV-2、SARS-CoV 和 MERS-CoV 的传染动力学分析[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020067.
- [12] 喻孜, 张贵清, 刘庆珍, 等. 基于时变参数-SIR 模型的 2019-nCoV 疫情评估和预测[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020027.
- [13] 梅文娟, 刘震, 朱静怡, 等. 新冠肺炎疫情极限 IR 实时预测模型[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020063.
- [14] ZHANG J, LITVINOVA M, WANG W, et al. Evolving epidemiology of novel coronavirus diseases 2019 and possible interruption of local transmission outside Hubei Province in China: A descriptive and modeling study[J]. *The Lancet Infectious Diseases*, 2020, DOI: 10.1016/S1473-3099(20)30230-9.
- [15] 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.
- [16] 林俊锋. 基于引入隐形传播者的 SEIR 模型的 COVID-19 疫情分析和预测[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020083.
- [17] 傅家旗, 刘敏, 邓春燕, 等. 复杂人流网络下的 COVID-19 传播模型[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020109.
- [18] 汪剑眉, 李钢. 新冠肺炎非均匀感染力传播模型与干预分析[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020158.
- [19] LITVINOVA M, LIU Q H, KULIKOV E, et al. Reactive school closure weakens the network of social interactions and reduces the spread of influenza[J]. *PNAS*, 2019, 116: 13174-13181.
- [20] 孙皓宸, 徐铭达, 许小可. 基于真实人际接触数据的新冠肺炎校园传播与防控[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020172.
- [21] CHEN Duan-bing, ZHOU Tao. Control Efficacy on COVID-19[EB/OL]. [2020-5-23]. <https://arxiv.org/abs/2003.00305>.
- [22] 郑文, 赵偲, 李泽堃, 等. 基于 Web 数据挖掘的 COVID-19 流行病学特征分析[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020079.
- [23] 张欣, 郁佳亮, VODENSKA I. 新冠肺炎疫情影响下区域产业网络风险传导效应研究[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020092.
- [24] 廖敬仪, 杨子曦, 周涛. 新冠肺炎家庭聚集性病例分析揭示女性的护理角色和高易感性[J]. *电子科技大学学报*, 2020, DOI: 10.12178/1001-0548.2020201.
- [25] HELLEWELL J, ABBOT S T, GIMMA A, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts[J]. *Lancet Global Health*, 2020, 8: e488-e496.
- [26] FERRETTI L, WYMANT C, KENDALL M, et al. Quantifying SARS-COV-2 transmission suggests epidemic control with digital contact tracing[J]. *Science*, 2020, DOI: 10.1126/science.abb6936.
- [27] BALCAN D, GONÇALVES B, HU H, et al. Modeling the spatial spread of infectious diseases: The global epidemic and mobility computational model[J]. *Journal of Computational Science*, 2010, 1: 132-145.
- [28] LI R, PEI S, CHEN B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus(SARS-CoV2)[J]. *Science*, 2020, DOI: 10.1038/s41586-020-2349-y.
- [29] PINTO D, PARK Y J, BELTRAMELLO M, et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody[J]. *Nature*, 2020, DOI: 10.1038/s41586-020-2349-y.
- [30] WANG W, LIU Q H., LIANG J, et al. Coevolution spreading in complex networks[J]. *Physics Reports*, 2019, 820: 1-51.
- [31] GAO J, ZHANG Y C, ZHOU T. Computational socioeconomics[J]. *Physics Reports*, 2019, 817: 1-104.
- [32] KISSLER S M, TEDIJANTO C, GOLDSTEIN E, et al. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period[J]. *Science*, 2020, DOI: 10.1126/science.abb5793(2020).
- [33] ZHANG H, YANG Z, WU Z, et al. Braess's paradox in epidemic game: Better condition results in less payoff[J]. *Scientific Reports*, 2013, 3: 3292.