新型冠状病毒基本再生数研究进展

万时雨, 刘珏*, 刘民*

北京大学公共卫生学院流行病与卫生统计学系, 北京 100191

* 联系人, E-mail: jueliu@bjmu.edu.cn; liumin@bjmu.edu.cn

2020-04-14 收稿, 2020-06-03 修回, 2020-06-03 接受, 2020-06-04 网络版发表

国家重点研发计划(2020YFC0846300)和国家自然科学基金(71934002)资助

摘要 新型冠状病毒肺炎是由一种新型冠状病毒(SARS-CoV-2)引起的新发传染病. 在传染病流行病学中, 基本再生数是可用于衡量传染病传播能力的一个定量参数, 若 $R_0 > 1$, 提示病例数会呈指数增长造成疾病暴发或流行; 若 $R_0 = 1$, 提示该种疾病将在人群中持续存在并保持稳定; 若 $R_0 < 1$, 提示传染病将会逐渐消失. R_0 可通过传播动力学模型构建并求解微分方程, 或采用最大似然法估计、随机模拟等方法估计. 随着全球新型冠状病毒肺炎疫情快速蔓延, 有关SARS-CoV-2基本再生数的研究证据日益增多. 经过同行评议的国际期刊发表文献所估计的中位 R_0 约为3.15 (95%CI: 2.26~6.20), 而预印本文献估计 R_0 为3.01 (95%CI: 1.99~5.44), 中文期刊估计为2.55 (95%CI: 1.61~3.55), 英文期刊估计为3.10 (95%CI: 2.09~6.05). 随着干预措施的实施, 再生数呈下降趋势. 本研究综述了SARS-CoV-2的基本再生数 R_0 相关研究, 以期了解病毒相关传播动力学特征提供科学依据.

关键词 SARS-CoV-2, COVID-19, 基本再生数, 新发传染病

新型冠状病毒肺炎是由一种新型冠状病毒(SARS-CoV-2)引起的新发传染病. 截至2020年5月27日, 全球累计报告确诊病例5491678例, 累计死亡349190例^[1]. 在传染病流行病学中, 基本再生数是可用于衡量传染病传播潜力的一个定量参数^[2]. 随着全球新型冠状病毒肺炎疫情快速蔓延, 有关基本再生数(basic reproduction number, R_0)的研究证据日益增多. 本文综述了SARS-CoV-2的基本再生数 R_0 相关研究, 以期了解病毒相关传播动力学特征提供科学依据.

1 基本再生数的概念

基本再生数是指一个感染者在具有传染性的时期内平均传染的人数, 主要用于衡量某种病原体的传播能力^[3]. 若 $R_0 > 1$, 提示病例数会呈指数增长, 疾病将暴发或流行; 若 $R_0 = 1$, 提示该种疾病将在人群中持续存在并保持稳定; 若 $R_0 < 1$, 提示传染病将会逐渐消失. 通常

随着防控措施的实施, R_0 也会发生变化^[4].

2 基本再生数的估计方法及常用模型

基本再生数可通过模型进行估计, 常用的模型包括传染病传播动力学模型(SIR(susceptible infected recovered)、SEIR(susceptible exposed infectious recovered)及其衍生模型等)、指数增长模型等. 通过在传播动力学模型中构建并求解微分方程^[5-7], 或采用最大似然法估计^[8]、随机模拟^[9]等方法估计基本再生数.

2.1 SIR模型

SIR模型的基本思想是将总人口分为3个仓室: S 代表易感者, 即未感染且不具有免疫力的人; I 代表感染者, 这部分人具有一定的传染性; R 代表移出者, 即痊愈的人, 这部分人具有免疫力. 在人口总数不变, 所有个体易感性、感染能力、痊愈能力相同, 均匀接触的假

引用格式: 万时雨, 刘珏, 刘民. 新型冠状病毒基本再生数研究进展. 科学通报, 2020, 65: 2334-2341

Wan S Y, Liu J, Liu M. Progress on the basic reproduction number of SARS-CoV-2 (in Chinese). Chin Sci Bull, 2020, 65: 2334-2341, doi: [10.1360/TB-2020-0413](https://doi.org/10.1360/TB-2020-0413)

设下, 3个仓室间人数可以通过微分方程表示。通过统计软件求解微分方程, 并通过预测病例数与实际病例数的差异来优化模型, 即可计算出基本再生数^[5]。

2.2 SEIR模型及其衍生模型

SEIR模型是在SIR模型的基础上, 考虑潜伏期患者。易感人群在感染后先进入潜伏期 E 仓室, 此期无传染性, 之后有一定概率发病进入 I 仓室。同样地, SEIR 4个仓室之间的关系可由一系列微分方程表示, 通过最小二乘法拟合累计病例数来估计模型中参数, 即可计算出基本再生数^[8]。

由于SARS-CoV-2潜伏期具有传染性, 存在地区间及人员流动等一系列不满足经典SEIR假设的现实条件, 因此有研究者在经典SEIR模型基础上改良, 加入其他仓室后拟合模型, 并通过马尔可夫链蒙特卡洛(Markov Chain Monte Carlo, MCMC)等方法进行随机采样以获得参数估计。例如, Wu等人^[10]在SEIR模型的微分方程中加入了国内与国际间陆地及航空旅行数据, 并考虑初始的动物源性传播; Shen等人^[11]建立了考虑有症状感染者(I , infectious individuals with symptoms)与接受治疗隔离者(J , isolated individuals with treatment)的SEIJR模型; Liu等人^[12]建立了考虑城市间人员流动与核酸检测假阴性患者(H , hidden)的Flow-SEHIR模型; Li等人^[13]建立了加入旅行数据且含有分层人群(记录的感染者与未被记录的感染者)的SEI模型; Choi等人^[14]建立了考虑住院者(H , hospitalized)的SEIHR模型; 白尧等人^[15]构建了考虑无症状感染者(A , asymptomatic but infectious)的SEIAR模型。不同的研究者结合SARS-CoV-2的特点, 通过对经典SEIR模型改良, 构建了不同的SEIR衍生模型, 使建立的模型更符合现实情况。

2.3 指数增长模型

在疫情初始阶段, 累计病例数往往呈指数增长^[8], 因此可以对累积病例数拟合指数增长模型以估计指数增长率, 并根据一定公式使用指数增长率计算基本再生数^[16]。考虑到防控措施的影响, Zhao等人^[17]基于报告率调整累计病例数, 并分别拟合指数增长模型, 以估计不同干预手段下基本再生数的差异。

2.4 其他模型

也有学者构建其他模型进行 R_0 的估计。Majumder等人^[18]通过构建病例延迟与指数调整(incidence decay and

exponential adjustment, IDEA)模型估计新型冠状病毒再生数。Chen等人^[19]通过建立蝙蝠-宿主-病毒储蓄地-人(bats-hosts-reservoir-people, BHRP)模型的简化RP模型估计新型冠状病毒基本再生数。Li等人^[20]使用与华南海鲜市场无关的病例建立考虑动物源性传播的更新方程传播模型。此外, 还可采用最大似然法以及随机模型等方法进行估计 R_0 。例如, 假设每日新增病例服从泊松分布, 则基本再生数可以由最大化对数似然估计函数估计得到^[8,21]。也有学者基于既往研究或流行病学调查数据获取代际间隔等关键参数的分布, 通过随机模拟的方法模拟传染病流行轨迹即累计预测病例数, 通过计算预测病例数落入一定区间的比例获取基本再生数的估计^[9,22]。

3 不同研究中对新型冠状病毒基本再生数的估计

世界卫生组织在疫情早期估计我国新型冠状病毒的 R_0 约为1.4~2.5^[4]。Li等人^[20]利用截至1月22日425例确诊患者的流行病学调查资料, 基于与华南海鲜市场无关的病例, 使用更新方程构建包含动物源性感染的传播模型, 估计我国SARS-CoV-2的 R_0 为2.2(95%CI: 1.4~3.9)。他们使用了人畜共患传染病传播模型, 较为真实地模拟了暴发初期疾病的传播状态; 但由于该估计仅基于10个确诊病例的信息, 且构建模型时使用了SARS的代际间隔, 因此结果可能不够精确。Tang等人^[21]利用1月10~22日国家卫生与健康委员会官方网站公布的报告病例数, 基于传染病动力学-仓室模型估计我国SARS-CoV-2的 R_0 为6.47(95%CI: 5.71~7.23)。该研究在SEIR模型基础上增加了新的仓室, 将易感者 S 、暴露者 E 与感染者 I 分层为被隔离的易感者 S_q 、被隔离的暴露者 E_q 与被隔离的感染者 I_q , 并加入了住院者 H 这一仓室, 以对疾病的流行状况进行更佳的模拟, 更全面地揭示了疾病早期流行特征。但由于疾病早期可用信息较少, 且对 R_0 的估计较为依赖所选择的数据集, 也有可能对 R_0 估计产生一定影响。Wu等人^[10]利用2019年12月31日~2020年1月28日报告病例数估计我国SARS-CoV-2的基本再生数约为2.68(95%CI: 2.47~2.86)。该研究在SEIR模型的基础上加入陆地旅行与航空数据, 模拟了人员流动对疾病暴发的影响, 但由于该模型有较强的理论假设, 如旅行行为不受疾病状态的影响、所有的感染者都有症状等, 这些假设可能在模拟疾病传播时引入误差。Kucharski等人^[22]估计, 截至2020年1月中旬, 我国 R_0 的中位数在1.6~2.6之间波动。该研究在SEIR仓

室模型基础上,考虑部分武汉暴露者通过国际旅行后在其他国家确诊的情况,结合武汉确诊数据与产生于武汉的国际病例,更好地模拟了武汉疾病暴发早期的流行情况。但由于该模型使用了既往其他研究中对潜伏期等关键流行病学参数的估计,随着对病毒认知的深入,有必要更新完善 R_0 估计结果。Chen等人^[19]基于潜在中间宿主-人传播网络模型估计我国 R_0 约为3.58。该研究中所用模型完整地模拟了病毒从蝙蝠-中间宿主-病毒仓库-人的传播链,但由于模型涉及参数较多,部分参数只能基于既往文献,也有可能造成结果的有偏估计。Cereda等人^[23]估计意大利伦巴第大区的 R_0 为3.1(95% CI: 2.9~3.2)。该研究使用疾病暴发早期的每日新增数据拟合泊松分布,较好模拟了暴发早期的流行特征。不同研究结果的差异性可能与所采用的模型方法、使用参数、拟合分布、使用的数据资料时间范围不同等因素有关。Liu等人^[24]分析了截至2月7日前发表的有关新型冠状病毒 R_0 的早期研究文献,发现采用模型方法估算 R_0 范围为2.2~2.68(平均2.44),采用数理统计

估计 R_0 的研究得出的范围为1.5~6.49(平均4.2),采用指数增长等统计学方法进行的 R_0 估计范围为2.2~3.58(平均2.67)。国内外新型冠状病毒 R_0 的估计结果见表1。

国内各省市的 R_0 估计也不尽相同。Liu等人^[33]利用1月23日前广东省所有确诊病例的数据,采用指数增长估计广东省新型冠状病毒 R_0 为2.90(95% CI: 2.32~3.63),同时采用最大似然法估计 R_0 为2.92(95% CI: 2.28~3.67)。李承倬等人^[5]利用浙江省卫生与健康委员会官网获取的2020年1月21日~2月16日数据构建SIR模型,估计2020年1月23日~2月13日期间浙江省新型冠状病毒 R_0 在1.6~2.6之间波动,且呈下降趋势。但由于SIR模型假设人口总数不变,所有个体易感性、感染能力、痊愈能力相同,均匀接触,该假设并不完全符合SARS-CoV-2的特性,因此容易造成有偏估计。各省新型冠状病毒 R_0 的估计结果见表2。

4 新型冠状病毒 R_0 动态变化估计

通常随着防控措施的实施, R_0 也会发生动态变

表1 国内外新型冠状病毒 R_0 的估计

Table 1 Estimation of R_0 in China and other countries

文献	地区	研究日期	方法	参数估计	R_0 (95%CI)
Chen等人 ^[19]	中国	2020.12.7~2020.1.1	建立蝙蝠-宿主-病毒储蓄地-人的简化RP模型	根据已有研究和报告确定一部分参数的取值,使用Berkeley Madonna方法拟合曲线,依据公式计算再生数	2.3~3.58
Liu等人 ^{[12]a)}	中国	~2020.1.20	建立Flow-SEHIR模型,Flow为城市间人员流动,H为核酸检测假阴性的患者	使用2020年1月20日之前的确诊及人员流动数据模拟疾病的自然传播	3.56(2.31~4.81)
Li等人 ^[20]	中国	~2020.1.22	基于考虑动物源性传播的更新方程对与华南海鲜市场无关的病例建立模型	根据暴露史和出现症状日期拟合对数正态分布估计潜伏期、根据出现症状的日期、第一次去医院的时间、医院确诊时间拟合Weibull分布来估计相应时间间隔的分布,使用更新方程构建模型并估计再生数	2.2(1.4~3.9)
Shen等人 ^{[11]a)}	中国	2019.12.12~2020.1.22	建立SEIJR模型,I为有症状感染者,J为接受治疗的隔离者	使用非线性最小二乘获得参数的点估计,使用马尔可夫链蒙特卡洛方法与Metropolis-Hastings算法获取参数的分布	4.71(4.50~4.92)
Tang等人 ^[21]	中国	2020.1.10~22	建立改良SEIR模型,包括住院者H,无症状感染者A,被隔离的易感者Sq,被隔离的暴露者Eq与被隔离的感染者Iq	使用马尔可夫链蒙特卡洛方法与Metropolis-Hasting采样拟合仓室模型并拟合模型关键参数,基于估计参数使用公式计算基本再生数	6.47 (5.71~7.23)
WHO ^{[4]a)}	中国	~2020.1.23	/	/	1.4~2.5
Li等人 ^[13]	中国	2020.1.10~23	拟合加入旅行数据的,含有分层人群(记录的和未记录的感染者)的SEI模型	通过随机采样进行模型试验和参数估计,并根据延迟报告调整预测数与报告数之间的关系,使用IF-EAFK方法评估模型有效性	2.38(2.04~2.77)
Zhao等人 ^[17]	中国	~2020.1.24	基于报告率调整病例总数,并拟合指数增长模型	使用均值为7.5,标准差为3.4的Gamma分布估计序列间隔,并根据非线性最小二乘拟合指数模型	2.13~3.33
Zhao等人 ^[25]	中国	~2020.1.24	基于指数泊松过程拟合流行曲线	使用SARS与MERS的平均代际间隔估计SARS-CoV2的代际间隔,并使用泊松先验的对数似然法估计指数增长率,并代入公式计算再生数	2.56(2.49~2.63)

(续表1)

文献	地区	研究日期	方法	参数估计	R_0 (95%CI)
Jung等人 ^[26]	中国	2019.12.8与2020.1.24	拟合指数增长模型	通过马尔可夫链蒙特卡洛方法(MCMC)拟合指数模型, 根据拟合参数计算再生数	2.10~3.19
Wu等人 ^[10]	中国	2019.12.31~2020.1.25	构建加入陆地及航空旅行数据与动物源性传播的SEIR模型	使用马尔可夫链蒙特卡洛方法(MCMC)与吉布斯采样估计基本再生数	2.68(2.47~2.86)
Majumder等人 ^{[18]a)}	中国	2019.12.8, 1.26	病例延迟与指数调整模型	使用SARS与MERS的中间代际间隔6~10 d建立IDEA模型	2.2~3.1
周涛等人 ^[27]	中国	~2020.1.26	构建SEIR模型	结合已有研究、报告数据, 带入公式计算	2.8~3.9
Riou等人 ^{[9]a)}	中国	2020.1.28	对早期人传人轨迹进行随机模拟	基于负二项分布生成二代病例, 基于Gamma分布生成代际间隔, 考虑指示病例数和动物传人的时间, 2019年11月27日对病例周围进行均匀采样, 将模拟结果相加, 与先前预测的感染人数进行比较	2.2(1.4~3.8)
Sanche等人 ^[28]	中国	2020.1.15~30	假定暴露/感染的武汉个体数符合指数增长, 建立“第一次到达”数学模型	假定暴露/感染的武汉个体数符合指数增长, 建立第一个病例到达各省的时间的概率分布函数, 通过使用国内旅行数据计算观察到的到达时间来估计指数增长率, 再使用已有研究的代际间隔与潜伏期, 通过公式计算基本再生数	5.7 (3.8~8.9)
宋倩倩等人 ^[8]	中国	2020.1.10~31	拟合指数增长模型、使用最大似然法、构建SEIR模型	使用Weibull、Gamma、对数正态模型拟合从报告中获取的潜伏期和代际间隔的概率分布并拟合模型, 根据模型参数计算 R_0	3.74(3.63~3.87, 指数) 3.16(2.90~3.43, 似然) 3.91(3.47~4.11, SEIR)
Tang等人 ^[29]	中国	~2020.2.3	构建基于代际间隔与每日出现症状数的更新方程	通过确诊者的流行病学数据拟合症状出现至确诊的时间分布, 使用武汉1月22日之前确诊病例数与该分布估计武汉在1月4日前出现症状的病例数; 使用流行病学数据拟合代际间隔的Gamma分布并代入更新方程估计 R_0	3.27(2.98, 3.58)
耿辉等人 ^[6]	中国	2020.1.21~2.8	构建SEIR模型	根据报告数据及模型假设确定参数, 代入公式计算	2.38~2.72
Lai等人 ^[30]	中国	~2020.2.13	拟合指数增长模型	根据指数增长率计算	2.6(2.1~5.1)
Rocklov等人 ^[31]	钻石公主号	2020.1.21~2.20	构建SEIR模型	根据预测的累计病例与观测到的累计病例进行矫正	14.8
Fang等人 ^[32]	中国	~2020.2.29	建立SEIR模型	根据确诊病例数拟合三次函数, 根据拟合的三次函数的预测病例数来模拟SEIR模型并获得参数估计, 根据公式计算再生数	2.35~3.21
Choi等人 ^[14]	韩国	~2020.3.5	建立SEIHR模型, H 指住院者	根据预测病例数与实际病例数调整模型参数	4.27(湖北) 3.47~3.54(韩国)
Tang等人 ^{[21] a)}	意大利伦巴第大区	2020.1.14~3.8	假设每日新增病例符合泊松分布, 根据似然函数估计再生数	利用流行病学数据估计代际间隔等分布, 假设每日新增病例符合泊松分布, 使用马尔可夫链蒙特卡洛方法Metropolis-Hastings采样获取再生数的后验分布	3.1(2.9~3.2)

a) 预印本. 非标注的文献均已在线发表; /表示未提及

化^[4]. 既往有研究者分析了不同时期新型冠状病毒的 R_0 动态变化情况. 黄丽红等人^[39]根据COVID-19疫情网络公开数据, 利用序贯Bayes方法估计全国、各省市以及武汉市不同时间 R_0 动态变化情况, 结果发现, 疫情初期全国整体 R_0 并不稳定, 数值较大, 误差也大. 随着防控措施的加强和疫情的发展, R_0 普遍在1月下旬开始呈现下降趋势, 2月份之后下降趋势稳定, 并已经在2月8日小于1, 预计在3月11日后接近0. 研究者分析还发现, 截至2020年2月10日, 纳入分析的24个省份中已有18个省

份 R_0 降至1以下. Tian等人^[40]估计在1月23日启动一级响应之前, 新型冠状病毒的 R_0 为3.15; 1月23日开始, 随着防控措施实施范围扩大, 高、中、低防控力度水平的省份 R_0 分别下降至0.97、2.01、3.05. 当干预措施在所有省份实施完成95%时, 平均 R_0 下降至0.04.

5 小结与展望

综上, 经过同行评议的国际期刊发表的SARS-CoV-2的 R_0 中位数估计为3.15(95% CI: 2.26~6.20), 而

表2 各省新型冠状病毒 R_0 的估计结果比较Table 2 Comparison of the estimation of R_0 in different provinces

文献	研究地区	研究日期	方法	参数估计	R_0 (95%CI)
Liu等人 ^{[33] a)}	广东省	~2020.1.23	拟合指数分布、使用最大似然法	使用SARS的生成时间通过泊松回归拟合指数增长模型, 使用指数增长率计算再生数、假设二代病例数服从泊松分布, 基于最大似然计算再生数	2.90(2.32~3.63, 指数) 2.92(2.28~3.67, 似然)
Zhao等人 ^[7]	湖北省	2020.1.10~24	基于报告率调整病例总数, 并拟合指数模型	分别根据MERS的均值为7.7、标准差为3.4的Gamma分布, SARS的均值为8.4、标准差为3.8的Gamma分布和平均之后的均值为8.0、标准差为3.6的Gamma分布来估计代际间隔, 并根据非线性最小二乘拟合指数模型	2.24~5.71
Sanche等人 ^[34]	湖北省	~2020.1.30	指数增长模型(模型一); 基于人群分层的SEIR模型(模型二)	模型一假设武汉本地的病例指数增长, 使用旅行数据和感染者在各省市首次出现的时间拟合指数增长模型; 模型二假设武汉病例指数增长, 使用随机的基于主体的模型描述感染者的轨迹, 拟合SEIR模型. 参数假设一为传染期等于潜伏期; 参数假设二为潜伏期比传染期短2 d	6.2(3.3~11.3, 指数1) 4.7(2.8~7.6, 指数2) 6.6(4~10.5, SEIR1) 4.9(3.3~7.2, SEIR2)
武文韬等人 ^[35]	广东省	2020.1.4~2.8	构建SIR模型	最小化 t 时刻真实值和预测值之间的方差	1.65~2.48
白尧等人 ^[15]	山西省	2019.12.31~2020.2.13	构建SEIAR模型, 其中A指无症状感染者	通过文献、实际疫情确定参数范围, 使用Berkeley Madonna软件构建模型	2.95
李承倬等人 ^[5]	浙江省	2020.1.21~2.16	构建SIR模型	R软件构建和求解微分方程, 使用optim函数优化	1.65~2.6
唐三一等人 ^[36]	陕西省	~2020.2.17	构建基于代际间隔与每日新增输入病例、新增本地感染者的更新方程	使用确诊病例的流行病学数据估计输入病例进入陕西省到感染本地易感者的时长与代际间隔的参数分布, 并根据该分布与每日新增确诊数估计每日新增本地感染者数, 代入更新方程计算控制再生数	1.48~1.69
林挺葵等人 ^[37]	广东省西部及其地级市	2020.1.21~2.18	构建SIR模型	R软件deSolve包构建模型	1.54~2.36
李承倬等人 ^[38]	河南省	2020.1.21~2.19	构建SIR模型	R软件构建和求解微分方程, 使用optim函数优化	1.25~1.7

a) 预印本. 非标注的文献均已在线发表.

未经同行评议的预印本文献对 R_0 的估计为3.01(95% CI: 1.99~5.44), 中文期刊中对 R_0 的估计为2.55(95% CI: 1.61~3.55), 英文期刊中对 R_0 的估计为3.10(95% CI: 2.09~6.05). 且随着干预措施的实施, R_0 呈下降趋势. 值得注意的是, 在估计SARS-CoV-2基本再生数时, 模型假设、仓室设置、采用参数、使用数据的不同均会影响 R_0 的估计结果. 在传播动力学模型中, 原始SIR模型由于其基本假设的限制, 无法考虑人口变动、个体差异及病毒潜伏期具有传染性等因素, 未经模型改良而直接估计 R_0 时会造成有偏估计. 有部分研究对SEIR进行了改进, 引入了新的仓室使之更符合

SARS-CoV-2特征, 所拟合的 R_0 更佳. 不过, 也应注意避免过多引入仓室而导致模型过度拟合等问题. 在采用指数增长模型、随机模拟等其他方法进行 R_0 估计时, 由于不受传播动力学模型的各种假设条件限制, 估计更稳定. 但由于未考虑传染病传播的内在特性, 因此更容易受到数据本身的影响(如数据分布、采用参数等), 因此采用适宜的数据分布和参数尤为关键. 无论采用何种方法, 笔者建议在建模估计时应尽可能充分考虑病毒特征, 使模型的假设与现实情况更贴近, 并选择尽可能详实的数据集, 以期实现对 R_0 的准确估计.

参考文献

- 1 Tian J, Yuan X, Xiao J, et al. Clinical characteristics and risk factors associated with COVID-19 disease severity in patients with cancer in Wuhan, China: A multicentre, retrospective, cohort study. *Lancet Oncol*, 2020, doi: 10.1016/S1470-2045(20)30309-0
- 2 Chowe G, Hyman J M, Bettencourt L M A, et al. *Mathematical and Statistical Estimation Approaches in Epidemiology*. London: Springer,

- Dordrecht, 2009
- 3 Liu M, Liu X, Liang W N. Handbook of Basic Epidemiological Knowledge on Coronavirus Disease 2019 (COVID-19)-110 Questions & Answers. Beijing: Peking University Medical Press, 2020 [刘民, 刘霞, 梁万年. 新型冠状病毒肺炎流行病学知识110问. 北京: 北京大学医学出版社, 2020]
 - 4 Special Expert Group for Control of the Epidemic of Novel Coronavirus Pneumonia of the Chinese Preventive Medicine. An update on the epidemiological characteristics of novel coronavirus pneumonia (COVID-19) (in Chinese). *Chin J Epidemiol*, 2020, 41: 139–144 [中华预防医学会新型冠状病毒肺炎防控专家组. 新型冠状病毒肺炎流行病学特征的最新认识. *中华流行病学杂志*, 2020, 41: 139–144]
 - 5 Li C Z, Wu W T, Pan Z Y, et al. Prevention and control of coronavirus disease 2019 in Zhejiang Province based on SIR model and R_0 (in Chinese). *Zhejiang Med J*, 2020, 42: 311–314 [李承倬, 武文韬, 潘振宇, 等. 基于SIR模型和基本再生数的浙江省新型冠状病毒肺炎防控效果分析. *浙江医学*, 2020, 42: 311–314]
 - 6 Geng H, Xu A D, Wang X Y, et al. Analysis of the role of current prevention and control measures in the epidemic of corona virus disease 2019 based on SEIR model (in Chinese). *J Jinan Univ (Nat Sci Med Ed)*, 2020, 41: 175–180 [耿辉, 徐安定, 王晓艳, 等. 基于SEIR模型分析相关干预措施在新型冠状病毒肺炎疫情中的作用. *暨南大学学报(自然科学与医学版)*, 2020, 41: 175–180]
 - 7 Zhao S, Lin Q, Ran J, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int J Infect Dis*, 2020, 92: 214–217
 - 8 Song Q Q, Zhao H, Fang L Q, et al. Study on assessing early epidemiological parameters of COVID-19 epidemic in China (in Chinese). *Chin J Epidemiol*, 2020, 41: 461–465 [宋倩倩, 赵涵, 方立群, 等. 新型冠状病毒肺炎的早期传染病流行病学参数估计研究. *中华流行病学杂志*, 2020, 41: 461–465]
 - 9 Riou J, Althaus C L. Pattern of early human-to-human transmission of Wuhan 2019-nCoV. *BioRxiv*, 2020, doi: 10.1101/2020.01.23.917351
 - 10 Wu J T, Leung K, Leung G M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *Lancet*, 2020, 395: 689–697
 - 11 Shen M W, Peng Z H, Xiao Y N, et al. Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. *BioRxiv*, 2020, doi: 10.1101/2020.01.23.916726
 - 12 Liu Q H, Liu Z C, Li D Q, et al. Assessing the tendency of 2019-nCoV (COVID-19) outbreak in China. *MedRxiv*, 2020, doi: 10.1101/2020.02.09.20021444
 - 13 Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 2020, 368: 489–493
 - 14 Choi S, Ki M. Estimating the reproductive number and the outbreak size of novel coronavirus disease (COVID-19) using mathematical model in Republic of Korea. *Epidemiol Health*, 2020: e2020011
 - 15 Bai Y, Liu K, Chen Z J, et al. Early transmission dynamics of COVID-19 in Shaanxi Province (in Chinese). *Chin J Nosocomiol*, 2020, 30: 834–838 [白尧, 刘昆, 陈志军, 等. 陕西省新型冠状病毒肺炎疫情的早期传播动力学研究. *中华医院感染学杂志*, 2020, 30: 834–838]
 - 16 Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc B*, 2007, 274: 599–604
 - 17 Zhao S, Lin Q, Ran J, et al. The basic reproduction number of novel coronavirus (2019-nCoV) estimation based on exponential growth in the early outbreak in China from 2019 to 2020: A reply to Dhungana. *Int J Infect Dis*, 2020, 94: 148–150
 - 18 Majumder M, Mandl K D. Early transmissibility assessment of a novel coronavirus in Wuhan, China. *SSRN J*, 2020, doi: 10.2139/ssrn.3524675
 - 19 Chen T M, Rui J, Wang Q P, et al. A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infect Dis Poverty*, 2020, 9: 24
 - 20 Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N Engl J Med*, 2020, 382: 1199–1207
 - 21 Tang B, Wang X, Li Q, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *J Clin Med*, 2020, 9: 462
 - 22 Kucharski A J, Russell T W, Diamond C, et al. Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *Lancet Infect Dis*, 2020, 20: 553–558
 - 23 Cereda D, Tirani M, Rovida F, et al. The early phase of the COVID-19 outbreak in Lombardy, Italy. *ArXiv*, 2020, doi: arXiv.2003.09320.
 - 24 Liu Y, Gayle A A, Wilder-Smith A, et al. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med*, 2020, 27: taaa021
 - 25 Zhao S, Musa S S, Lin Q, et al. Estimating the unreported number of novel coronavirus (2019-nCoV) cases in china in the first half of January 2020: A data-driven modelling analysis of the early outbreak. *J Clin Med*, 2020, 9: 388
 - 26 Jung S, Akhmetzhanov A R, Hayashi K, et al. Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: Inference using exported cases. *J Clin Med*, 2020, 9: 523

- 27 Zhou T, Liu Q H, Yang Z M, et al. Preliminary prediction of the basic reproduction number of the novel coronavirus 2019-nCoV (in Chinese). *Chin J Evid Based Med*, 2020, 20: 359–364 [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. *中国循证医学杂志*, 2020, 20: 359–364]
- 28 Sanche S, Lin Y T, Xu C, et al. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis*, 2020, , doi: 10.3201/eid2607.200282
- 29 Tang B, Xia F, Tang S, et al. The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China. *Int J Infect Dis*, 2020, 95: 288–293
- 30 Lai A, Bergna A, Acciarri C, et al. Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2. *J Med Virol*, 2020, 92: 675–679
- 31 Rocklöv J, Sjödin H, Wilder-Smith A. COVID-19 outbreak on the diamond princess cruise ship: Estimating the epidemic potential and effectiveness of public health countermeasures. *J Travel Med*, 2020, 27: taaa030
- 32 Fang Y, Nie Y, Penny M. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *J Med Virol*, 2020, 92: 645–659
- 33 Liu T, Hu J X, Kang M, et al. Transmission dynamics of 2019 novel coronavirus (2019-nCoV). *BioRxiv*, 2020, doi: 10.1101/2020.01.25.919787
- 34 Sanche S, Lin T Y, Xu C G, et al. The novel coronavirus, 2019-nCoV, is highly contagious and more infectious than initially estimated. *BioRxiv*, 2020, doi: 10.1101/2020.02.07.20021154
- 35 Wu W T, Bai R H, Li D N, et al. Preliminary prediction of the epidemic trend of corona virus disease 2019 in Guangdong Province (in Chinese). *J Jinan Univ (Nat Sci Med Ed)*, 2020, 41: 181–185 [武文韬, 柏如海, 李达宁, 等. 广东省新型冠状病毒肺炎疫情流行趋势的初步预测. *暨南大学学报(自然科学与医学版)*, 2020, 41: 181–185]
- 36 Tang S Y, Tang B, Bragazzi N L, et al. Analysis of COVID-19 epidemic traced data and stochastic discrete transmission dynamic model (in Chinese). *Sci Sin Math*, 2020, 50: 1–16 [唐三一, 唐彪, Bragazzi N L, 等. 新型冠状病毒肺炎疫情数据挖掘与离散随机传播动力学模型分析. *中国科学: 数学*, 2020, 50: 1–16]
- 37 Lin T K, Wu J Y, Liu H F, et al. Prediction of epidemic trend of COVID-19 in western Guangdong and other prefecture-level cities: A study based on holt two-parameter exponential smoothing model (in Chinese). *Prac J Card Cereb Pneum Vasc Dis*, 2020, 28: 13–17 [林挺葵, 吴家园, 刘华锋, 等. 粤西地区及各地级市新型冠状病毒肺炎疫情发展趋势预测分析——基于Holt双参数指数平滑模型的研究. *实用心脑血管病杂志*, 2020, 28: 13–17]
- 38 Li C Z, Feng A Z, Wu W T, et al. Analysis epidemic prevention and control effect of COVID-19 in Henan (in Chinese). *China J Tradit Med*, 2020, 35: 497–500 [李承倬, 冯敖梓, 武文韬, 等. 河南省新型冠状病毒肺炎疫情防控效果分析. *中医学报*, 2020, 35: 497–500]
- 39 Huang L H, Shen S P, Yu P, et al. Dynamic basic reproduction number based evaluation for current prevention and control of COVID-19 outbreak in China (in Chinese). *Chin J Epidemiol*, 2020, 41: 466–469 [黄丽红, 沈思鹏, 余平, 等. 基于动态基本再生数的新型冠状病毒肺炎疫情防控现状评估. *中华流行病学杂志*, 2020, 41: 466–469]
- 40 Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 2020, 368: 638–642

Summary for “新型冠状病毒基本再生数研究进展”

Progress on the basic reproduction number of SARS-CoV-2

Shiyu Wan, Jue Liu^{*} & Min Liu^{*}

Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing 100191, China

^{*} Corresponding authors, E-mail: jueliu@bjmu.edu.cn; liumin@bjmu.edu.cn

Coronavirus disease 2019 (COVID-19) is an emerging infectious diseases caused by a new coronavirus named SARS-CoV-2. In the epidemiology of infectious diseases, the basic reproduction number (R_0) is a central quantitative parameter that can be used to measure the transmission potential of infectious diseases. R_0 represents the average number of new infections generated by an infectious person in a totally naïve population. If $R_0 > 1$, the number infected is likely to increase. If $R_0 < 1$, transmission is likely to die out. In this study, we present a review on the estimation of basic reproduction number of SARS-CoV-2, in order to provide a scientific basis for better understanding the dynamic characteristics of SARS-CoV-2 transmission. R_0 is usually estimated with various types of complex mathematical models. It can be calculated and estimated by differential equations in the propagation dynamics model (such as SIR model and SEIR model), or estimated using maximum likelihood method, stochastic model, etc. Modeled R_0 values are dependent on model structures and assumptions. With the rapid spread of the global COVID-19 epidemic, there is an increasing evidence on R_0 of SARS-CoV-2. However, the results of R_0 estimation varied a lot in different studies, due to different model assumptions, parameters settings, and data used in the models. The estimated median R_0 is about 3.15 (95%CI: 2.26–6.20) in studies published by peer-reviewed journals, 3.01 (95%CI: 1.99–5.44) in preprinted platform without peer-reviewed and online reports, and 2.55 (95%CI: 1.61–3.55) in studies published in Chinese language, and 3.10 (95%CI: 2.09–6.05) in studies published in English language. With the implementation of comprehensive intervention measures, R_0 showed a downward trend. In the basic propagation dynamics model, R_0 might be estimated with biases, due to the limitations of its basic assumptions (such as fixed population). Regardless of the methods used, characteristics of the virus should be considered fully during estimation. To better estimate R_0 , some researchers improved the basic propagation dynamics model and considered the assumption of unfixed population, population mobility, infectivity of the virus during incubation period, and quarantine measures, to make it more in line with the characteristics of SARS-CoV-2. When adopting other methods such as exponential growth model and stochastic model to estimate R_0 , the estimated results is relatively stable because that it is not restricted by assumptions of the propagation dynamics model. However, the estimation are more susceptible to the data used in the model (such as distribution of the data) because that the inherent characteristics of the spread of infectious diseases are not considered. Thus, it is particularly critical to use appropriate data distribution and parameters during estimation.

SARS-CoV-2, COVID-19, basic reproduction number, emerging infectious diseases

doi: [10.1360/TB-2020-0413](https://doi.org/10.1360/TB-2020-0413)