

# 基于 SEIAQR 动力学模型预测 湖北省武汉市新型冠状病毒肺炎疫情趋势\*

丁中兴<sup>1</sup> 宋文煜<sup>1</sup> 方欣玉<sup>1</sup> 王 凯<sup>2</sup> 鲍倡俊<sup>3</sup> 陈 峰<sup>1</sup> 沈洪兵<sup>1</sup> 武 鸣<sup>3△</sup> 彭志行<sup>1△</sup>

**【提 要】** **目的** 基于新型冠状病毒传播特点,使用实际数据拟合并预测武汉市疫情发展趋势。**方法** 考虑新型冠状病毒肺炎传播规律及采取的隔离措施,构建 SEIAQR 动力学模型;研究使用 2020 年 1 月 10 日至 1 月 31 日武汉市累积发病人数及死亡人数数据对模型进行拟合,并使用 2 月 1 日至 2 月 10 日的实时数据对模型结果进行验证;最后,研究通过模型拟合结果评价现有抗疫防控措施的效果。**结果** 武汉市发病人数在 2 月 4 日左右达到峰值,随后持续下降,2 月 14 日至 2 月 16 日出现一个短暂的反弹期,直至 5 月疫情将基本结束流行。防控措施不变的情况下,疫情结束时预计 50780 人感染,2449 人死亡,粗死亡率约为 4.82%;及时有效的防控措施降低了 75.08% 的感染率和 72.40% 的因病死亡率;模型拟合偏差约为 4.08%。**结论** SEIAQR 模型拟合效果良好,可以较为准确地预测疾病流行趋势。

**【关键词】** 新型冠状病毒肺炎; 传染病动力学模型; 疫情预测

## Using SEIAQR Dynamic Model to Predict the Epidemic Trend of Novel Coronavirus Pneumonia in Wuhan, Hubei Province

Ding Zhongxing, Song Wenyu, Fang Xinyu, et al( *Epidemiology and Health Statistics, School of Public Health, Nanjing Medical University*( 211166) , Nanjing)

**【Abstract】 Objective** Based on the characteristics of the new coronavirus, the actual data were used to predict the epidemic trend in Wuhan. **Methods** The SEIAQR dynamic model was constructed by considering the transmission rule and isolation measures of the novel coronavirus pneumonia. The cumulative number of cases and deaths in Wuhan from January 10th to January 31st, 2020 were used to fit the model and the results were verified with the real-time data from February 1st to February 10th. Finally, the results of model fitting were used to evaluate the effectiveness of prevention and control measures. **Results** The number of cases in Wuhan reached the peak around February 4th, then continued to decline, and there was a significant rebound from February 14th to February 16th, until the outbreak would be basically controlled in May. With the prevention and control measures unchanged, it is estimated that 50780 people were infected and 2449 died finally, with an crude mortality rate of about 4.82%. Timely and effective prevention and control measures reduced the infection rate by 75.08% and the mortality rate by 72.40%. The model fitting deviation was about 4.08%. **Conclusion** The SEIAQR model has a good fitting performance and can predict the epidemic trend accurately.

**【Key words】** Novel coronavirus; Dynamic model; Outbreak forecast

21 世纪以来全世界共发生三起冠状病毒大流行, 分别是 2003 年严重急性呼吸道综合征( SARS)、2014 年中东呼吸综合征( MERS) 和 2020 年新型冠状病毒感染肺炎( COVID-19), COVID-19 较前两者具有更高的传染性及其较低的致死率<sup>[1-2]</sup>。截至 2020 年 5 月 1 日, 中国已报告 82875 例确诊, 77685 例治愈和 4633 例死亡, 其中约 87.70% ( 68128 例) 确诊和 97.39% ( 4512 例) 死亡来自湖北省<sup>[3-4]</sup>。1 月 23 日, 中国政府启动突发公共卫生事件一级响应, 上午 10 点起武汉的城市公交、地铁、机场和火车站等交通方式暂停运营, 1 月 24 日, 湖北省启动重大突发公共卫生事件一级响应, 陆续封锁其余城市的公共交通。基于传

染病动力学模型, 本研究拟合并预测武汉市新型冠状病毒肺炎的疫情发展趋势, 为科学防控提供参考。

### 资料与方法

1. 数据来源: 数据来自于国家卫健委、武汉市卫健委的发布信息, 整理得到 2020 年 1 月 10 日至 2 月 10 日武汉市新型冠状病毒感染肺炎累积报告病例数、累积治愈病例数及累积死亡病例数, 见表 1。

2. 模型构建: 已有学者使用 SEIR 模型较好地拟合 COVID-19 流行趋势<sup>[5-6]</sup>, 在此基础上, 我们构建 SEIAQR 模型, 总人群  $N$  被分为 6 个仓室: 易感者(  $S$  ), 潜伏者(  $E$  ), 感染者(  $I$  ), 无症状感染者(  $A$  ), 发现并被隔离者(  $Q$  ) 以及康复者(  $R$  ); 即  $N = S + E + I + A + Q + R$ 。新冠肺炎的传播流程图如图 1 所示, 因潜伏者和染病者均具有感染性, 易感者接触潜伏者或感染者可成为新冠肺炎的病毒携带者, 考虑疾病的发生率为标准发生率, 易感者与潜伏者或感染者的接触率分别是  $\beta(t)$ ,  $\beta_E(t)$ , 其中  $\beta(t) = 1 - (1 - \beta)^{C \times (1 - f(t))} \times$

\* 基金项目: 国家自然科学基金( 81673275, 82041024); 国家科技重大专项( 2018ZX10715002-004-002, 2018ZX10713001-001)

1. 南京医科大学公共卫生学院流行病与卫生统计学系( 211166)

2. 新疆医科大学医学工程技术学院数学教研室

3. 江苏省疾病预防控制中心

△通信作者: 武鸣, E-mail: jswuming@vip.sina.com; 彭志行, E-mail: zhihangpeng@njmu.edu.cn

表1 2020年1月10日至1月31日武汉市新型冠状病毒感染肺炎病例

时间	1/10	1/11	1/12	1/13	1/14	1/15	1/16	1/17	1/18	1/19	1/20	1/21	1/22	1/23	1/24	1/25
累积确诊人数	41	41	41	41	41	41	45	62	121	198	258	363	425	495	572	618
累积治愈人数	2	6	7	7	7	12	15	19	24	25	0	28	0	31	32	40
累积死亡人数	1	1	1	1	1	2	2	2	3	3	5	8	16	23	38	45
时间	1/26	1/27	1/28	1/29	1/30	1/31	2/1	2/2	2/3	2/4	2/5	2/6	2/7	2/8	2/9	2/10
累积确诊人数	698	1590	1905	2261	2639	3215	4109	5142	6384	8351	10117	11618	13603	14982	16902	18454
累积治愈人数	42	47	80	87	108	144	176	229	308	373	436	539	703	882	1049	1211
累积死亡人数	63	85	104	129	159	192	224	265	313	362	414	478	545	608	681	748

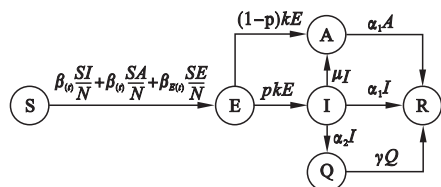


图1 COVID-19 传播动力学 SEIAQR 模型

$((1-\delta) \times \beta)^{c \times f(t)}$ ,  $\beta$  表示易感者每次与感染者接触后被感染的平均概率,  $c$  表示人群日接触次数,  $\delta$  表示佩戴各类口罩预防病毒感染的平均有效率, 函数  $f(t)$  在  $t$  时刻时口罩的使用率;  $\beta_E(t) = \varepsilon \beta(t)$ , 这里  $\varepsilon (0 < \varepsilon < 1)$  是一个衰减因子, 表示潜伏者相对于感染者传染力度的差异; 潜伏者以速率  $k$  发展成为感染者 (其中  $p$  部分为有症状感染者,  $1-p$  为无症状感染者); 感染者和无症状感染者均以  $\alpha_1$  的速率恢复, 其中感染者因为出现临床症状以速率  $\alpha_2$  被发现和隔离, 因新冠肺炎的病死率为  $\mu$ 。在严格的隔离防控措施下, 不考虑隔离者与易感者发生接触, 隔离感染者以速率  $\gamma$  恢复。进而建立如下新冠肺炎的动力学模型:

$$\begin{cases} \frac{dS}{dt} = -\beta_{(t)} \frac{SI}{N} - \beta_{(t)} \frac{SA}{N} - \beta_{E(t)} \frac{SE}{N} \\ \frac{dE}{dt} = \beta_{(t)} \frac{SI}{N} + \beta_{E(t)} \frac{SE}{N} - pkE - (1-p)kE \\ \frac{dI}{dt} = pkE - \alpha_1 I - \alpha_2 I - \mu I \\ \frac{dA}{dt} = (1-p)kE - \alpha_1 A \\ \frac{dQ}{dt} = \alpha_2 I - \gamma Q \\ \frac{dR}{dt} = \gamma Q + \alpha_1 I + \alpha_2 I \end{cases}$$

记累积感染人数为  $C$ , 则可得到  $\frac{dC}{dt} = kE$ 。

3. 参数估计: 新型冠状病毒的潜伏期平均为 5.2 天<sup>[5]</sup>, 发病到隔离、治疗的时间平均为 7 天<sup>[3]</sup>, 患者住院的中位数为 10 天<sup>[7-9]</sup>; 人际日接触系数  $c$  初始设定为 10 (1~50)<sup>[10]</sup>; 以既往文献研究为参考, 口罩的有效保护率  $\delta$  设定为 0.85 (0.8~0.99)<sup>[11]</sup>。见表 2, 表 3。

表2 SEIAQR 模型参数含义及取值

参数	描述	取值范围或 95% 置信区间	取值来源
$1/k$	平均潜伏时间	5.2 (4.1~7.0)	文献[5]
$1/\alpha_1$	患者自然康复平均时间	7 (4~8)	文献[3]
$1/\alpha_2$	发病至隔离的平均时间		MCMC
$1/\gamma$	隔离至康复的平均时间	10 (7~14)	文献[7,8,9]
$\beta$	易感者与有症状感染者每次接触后感染的概率		MCMC
$\varepsilon$	潜伏者相对于感染者传染性的差异	0.5 (0.10~0.90)	专家意见*
$c$	人群日接触数	10 (1~50)	文献[10]
$\mu$	因病死亡率	0.043	文献[9]
$\delta$	口罩的有效保护率	0.85 (0.8~0.99)	文献[11]
$p$	有症状感染者占比	0.7	专家意见*

\*: 专家意见为咨询中华预防医学会新型冠状病毒肺炎防控专家组后得到。

表3 SEIAQR 模型仓室含义及初值

仓室	含义	起始数值	依据
$S(0)$	易感人群的初始值	9083500	武汉人口数, 全人群普遍易感
$E(0)$	潜伏期人群的初始值	估计值	马尔科夫链蒙特卡洛算法拟合
$I(0)$	感染人群的初始值	38	截至 1 月 10 日共计 38 人正处于感染阶段
$A(0)$	无症状感染者人群的初始值	0	未发现存在无症状感染者
$Q(0)$	隔离人群的初始值	0	未采取有效隔离措施
$R(0)$	康复人群的初始值	2	截至 1 月 10 日共计 2 人康复
$C(0)$	累积发病人群的初始值	41	截至 1 月 10 日累积发病 41 人次

本研究从已有数据中截取武汉市 1 月 10 日至 1 月 31 日的数据进行参数拟合, 采用非线性最小二乘的

方法拟合潜伏者仓室初始值  $E(0)$ , 患者自然康复速率  $\alpha$  以及传染率系数  $\beta$ 。将初步估计出的参数值用作

MCMC 方法<sup>[12]</sup>中 M-H 算法的先验信息,使用 Matlab 软件进行 10000 次的迭代,退火后选取后面 2000 次迭代结果计算各参数平均数及相应 95% 可信区间。模型预测 2 月 1 日后肺炎疫情的流行趋势,通过比较预测结果与 2 月 1 日至 2 月 10 日实际数据的差异验证参数准确性和合理性。

结 果

1. 口罩使用率变化趋势: 疾病流行初期,由于卫生部门重视程度不足,人群防控意识普遍淡薄,口罩使用率保持在较低水平且增长速率缓慢;随着疫情加剧蔓延且各方媒体的大力宣传教育,口罩使用率迎来高速增长阶段,并在短期内达到较为饱和的状态。为了尽可能表现这种非线性变化,我们构造一个 logistic 增长型函数  $f(t)$  进行拟合,设  $f(t) = \frac{1}{1 + e^{-k_0(t-t_1)}}$ ,  $k_0 = 0.4(0.3 \sim 0.5)$ <sup>[11]</sup>,  $t_1 = 13$ (1 月 10 日到 1 月 23 日实施封城的时间间隔)。结果显示,函数拟合效果满足口罩佩戴率增速先慢后快,最终接近饱和的实际情况,见图 2。

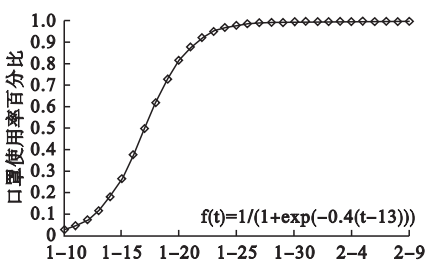


图 2 口罩使用率变化趋势图

2. 新型冠状病毒肺炎疫情趋势模拟: 我们在图 3 中使用湖北省武汉市 1 月 10 日至 1 月 31 日累积发病数据及死亡数据对模型参数进行拟合。左上图中 data1 为实际累积发病者, data2 为实际累积死亡者; 蓝色实线表示使用模型拟合的累积发病预测值, 红色实线表示累积死亡预测值。为了让结果体现更清晰,我们对 y 轴刻度值取对数后得右上图。结果显示,累积发病数和累积死亡数将在 3 月 1 日(第 50 天)趋于稳定。随着疫情逐渐被控制、消灭,疫情流行结束时预计武汉市会有 50780 人感染,累积死亡 2449 人,因病死亡率约为 4.82%。使用均方误差(MSE)函数评估模型拟

表 4 SEIAQR 动力学模型拟合效果评价

时间	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10
评价指标 MSE(%)	17.04	11.69	10.32	7.33	1.62	1.03	0.57	-2.66	-3.74	-2.34

合优度,经计算 2 月 1 日至 2 月 10 日累积确诊病例数的平均偏差值为 4.08%,差异处于可接受范围内,随着日期延长,差异逐渐变小且稳定,故认为本模型中所估计的各参数值较为合理。模型拟合效果见表 4。

图 3 下行两图显示,模型预测武汉市发病人数在 2 月 4 日(第 24 天)左右达到峰值,预计当天确诊 1512 人,随后出现一个短暂下降趋势,这个过程持续 6~7 天。2 月 14 日-2 月 16 日(第 34 天~37 天)疫情可能出现一个反弹期,2 月 16 日(第 37 天)后持续下降,直至今年 5 月(第 120 天)将基本得到控制。随着时

间推移,人群防控意识日益淡薄,疫情持续减弱过程中可能出现小幅度回升,但衰减趋势保持不变。实际数据显示,2 月 4 日附近出现峰值,当天共有 1967 人次确诊,随后出现快速下降,与模型预测结果相吻合。2 月 12 日,由于诊断方式的改变,临床诊断的患者数量增加到新增病例的数量中,日确诊人数激增 1 万多人,图 3 中出现一个极端的离群值。2 月 14 日至 2 月 16 日,疫情出现小幅度反弹(新增病例分别为 1001 例,1405 例,1548 例),与模型预测结果相一致(预测新增病例分别为 1334 例,1368 例,1391 例)。

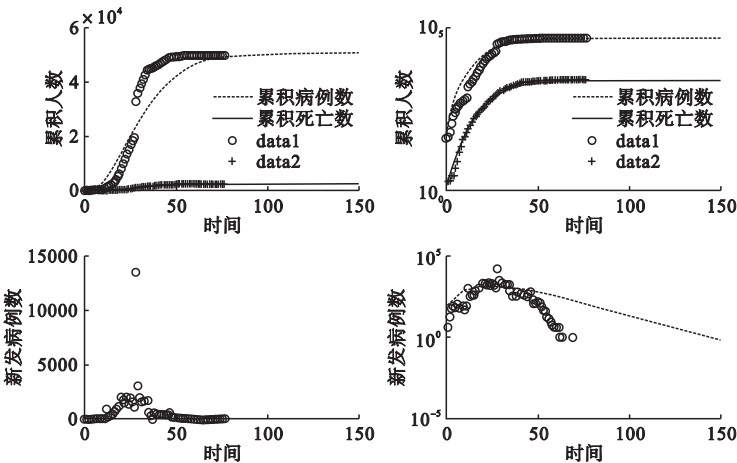


图 3 武汉市 COVID-19 实际日发病人数与模型拟合日发病人数

疫情发生后,武汉市 1 月 23 日暂时关闭全市公交、地铁、机场和火车站等交通方式,进入全面封禁状

态。本研究比较防控与未防控不同条件下疫情的蔓延趋势,结果见图 4。截至疫情流行结束,不采取或未及

时采取预防手段时,武汉市将有 20.38 万人感染,8874 人死亡,及时采取防控至少避免 15.30 万人感染和

6425 人死亡,有效降低 75.08% 感染率和 72.40% 因病死亡率。

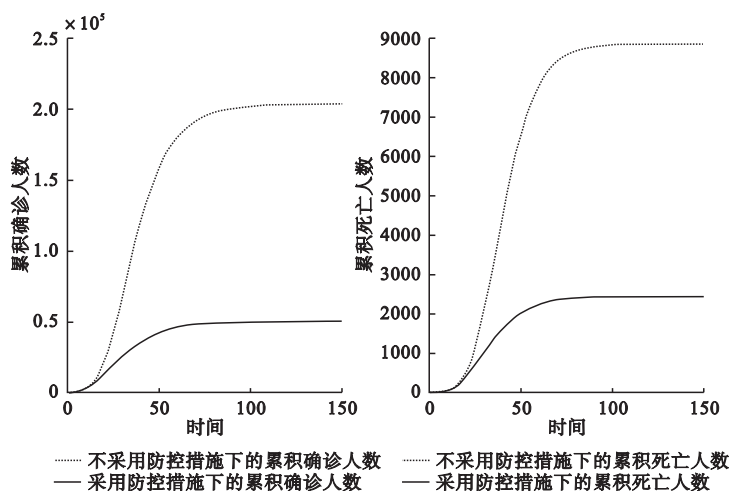


图4 武汉市两种措施下 COVID-19 疫情流行趋势

## 讨 论

已有研究发现<sup>[5]</sup>, COVID-19 潜伏期较长且具有传染性,新型肺炎病毒感染者分为有症状和无症状两种类型,无症状感染者在实际工作中很难被发现,这些不利因素导致传染病监测数据复杂多变,给模型评估预测带来难度。面对诸多挑战,本研究基于 COVID-19 疾病流行特征,集中分析预测武汉市的肺炎流行趋势。相比于既往文献常用的 SEIR 模型,本模型新增无症状感染者、隔离者、死亡者等因素,在模型结构上取得完善与创新,更符合实际应用需求。

通过拟合累积发病人数及累积死亡人数,我们得到较好的拟合结果,累积确诊病例预测值与实际值的平均偏差约为 4.08%。模型对累积死亡人数的预测十分准确,但累积发病人数的早期预测存在较小差异。考虑到实际工作过程中,部分 COVID-19 感染者不能及时被发现、上报,甚至存在漏报<sup>[13]</sup>,实时确诊报告人数存在滞后性,所以累积发病人数的预测值表现为开始略高于实际值,随后趋于一致。从日确诊发病数趋势来看,预测值也呈现“先高后低”现象,亦证实这种可能。

由模型拟合结果可知,截至疫情结束武汉市预计 50780 人感染,2449 人死亡,因病死亡率约为 4.82%,及时防控降低了 75.08% 感染率和 72.40% 因病死亡率。从分析结果来看,武汉的肺炎疫情已经过平台期,正处于逐渐衰退阶段。COVID-19 已经成为 21 世纪以来我国危害性最大,社会影响性最强的重大突发性传染病之一。从防控手段来看,早发现、早隔离、早治疗依然是阻断传染病蔓延的关键思路,对传染病的早期重视和防控能够有效阻止“小病”发展为“大疫”,避免后续产生更多伤亡和损失。同时,我们需要正视预

防工作的价值和意义。公共卫生体系的任务使命在于迅速准确地捕捉传染病蔓延的“蛛丝马迹”,及时提供反馈并依据这些信息部署正确合理的防控策略,这一点在未来的体系建设中应当加强。此外,模型结果显示,即使当前疫情拐点已过仍不能放松警惕,将来可能出现连续的小幅度反弹,保持现有防控力度和重视程度是合理且必要的。

本研究依然存在一些局限性。采取防控措施的同时,各方媒体大力宣传和报道,人群对疫情的预防措施愈发重视,导致我们可能忽略个人行为对疫情流行产生的干预<sup>[14]</sup>,过度估计防控手段对疫情控制的正面效果。另一方面,本研究仅考虑武汉市的疫情变化,没有关注到封城等措施对人口流动的积极影响,忽视了人口流动对于减缓疫情向省外城市蔓延的重要意义<sup>[15]</sup>。此外,模型中大多数参数在疫情流行期间呈现动态变化,本研究尚未考虑参数的随机化过程,可能会影响模型精确性。后续研究可以从卫生经济学的角度分析,探究采取有效防控措施所避免的经济损失。

## 参 考 文 献

- [1] Huang CL, Wang YM, Li XW, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020; DOI: [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [2] Wang C, Horby PW, GF Gao, et al. A novel coronavirus outbreak of global health concern. Lancet. 2020; DOI: 10.1016/S0140-6736(20)30185-9.
- [3] World Health Organization. Novel Coronavirus—China, Disease outbreak news: Update. [Accessed 1 May 2020] <https://www.who.int/csr>.
- [4] National Health Commission of the People's Republic of China. Daily report about 2019-nCoV. <http://www.nhc.gov.cn/>.
- [5] Li Q, Guan XH, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med. 2020. DOI: 10.1056/NEJMoA2001316.

(下转第 334 页)

归模型对数据的拟合效果更好的原因可能在于: 数据中零观测值相对较少, 为 15.5%, 这对零观测值只有一个来源并与非零计数截然分开的 hurdle 回归模型更合适。

本文只讨论了零膨胀和 Hurdle 模型在老年人医疗保健次数影响因素研究中的应用并进行比较, 实际的调查研究中还存在许多零膨胀计数资料。在应用回归模型进行数据拟合时, 不仅要考虑数据的性质和分布, 还要综合考虑实际情况和专业性, 从而选择最优模型。

### 参 考 文 献

- [1] 赵丽华, 刘桂芬, 田娇妮. 重复测量计数资料的随机效应 ZINB 模型. 中国卫生统计, 2011, 28(6): 665-667.
- [2] Lambert D. Zero-inflated Poisson Regression, With an Application to Defects in Manufacturing. Technometrics, 1992, 34(1): 1-14.
- [6] Wu J, Leung K, Leung G. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet, 2020. DOI: 10.1016/S0140-6736(20)30260-9. [published Online First: 2020/01/31].
- [7] Jiangxi News. The first confirmed case has recovered in Xingcan city, Jiangxi province. [Accessed 31 Jan 2020] [http://jx.cnr.cn/2011jxfw/xwj/20200127/t20200127\\_524950854.shtml](http://jx.cnr.cn/2011jxfw/xwj/20200127/t20200127_524950854.shtml).
- [8] Jiangxi News. The confirmed case has recovered in Pingxiang city, Jiangxi province. [Accessed 31 Jan 2020] <http://jx.people.com.cn/n2/2020/0127/c190260-33746426.html>.
- [9] Wang DW, Hu B, Hu C et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA. 2020. DOI: 10.1001/jama.2020.1585.
- [10] Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. P Roy Soc B - Biol Sci. 2014; 281: 20140268.
- [11] Offeddu V, Yung CF, Fong Low MS, et al. Effectiveness of masks

- [3] Xie H, Tao J, McHugo GJ, et al. Comparing statistical methods for analyzing skewed longitudinal count data with many zeros: An example of smoking cessation. Journal of Substance Abuse Treatment, 2013, 45(1): 99-108.
- [4] 许正. 零膨胀数据的几种模型及应用. 扬州大学, 2014.
- [5] Neelon B, Ghosh P, Loebs PF. A Spatial Poisson Hurdle Model for Exploring Geographic Variation in Emergency Department Visits. J R Stat Soc Ser A Stat Soc, 2013, 176(2): 389-413.
- [6] Adnan R, Saffari SE, Greene W. Hurdle Negative Binomial Regression Model with Right Censored Count Data. 2012.
- [7] 李蕊. Hurdle 回归和零膨胀回归在交通事故人员伤亡中的应用研究. 太原理工大学, 2015.
- [8] 朱高培, 朱乐乐, 孟马承, 等. 零膨胀负二项回归模型在共存疾病影响因素研究中的应用. 中华疾病控制杂志, 2018, 22(10): 1063-1066.
- [9] 李丹丹, 韩娜, 崔壮, 等. 潜在类别分析在糖尿病脆弱性人群分类中的应用. 中国卫生统计, 2018, 35(1): 11-13.

(责任编辑: 邓 妍)

(上接第 330 页)

- and respirators against respiratory infections in healthcare workers: A systematic review and meta-analysis. Clin Infect Dis, 2017, 65(11): 1934-1942.
- [12] Haario H, Laine M, Mira A, et al. Dram: Efficient adaptive MCMC. Stat Comput, 2006, 16: 339(54).
- [13] Zhao S, Musa SS, Lin QY 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. DOI: 10.3390/jcm9020388.
- [14] 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. DOI: 10.3390/jcm9020462.
- [15] Zhao S, Zhuang ZA, Ran JJ, et al. The association between domestic train transportation and novel coronavirus(2019-nCoV) outbreak in China from 2019 to 2020: A data-driven correlational report. Travel Med Infect Dis, 2020. DOI: 10.1016/j.tmaid.2020.101568.

(责任编辑: 郭海强)