

智能体模型在艾滋病流行病学中的应用*

吴声¹ 樊虹¹ 张彭燕¹ 徐艺耘¹ 施婷婷¹ 张欣¹ 张铁军^{1,2,Δ}

【摘要】目的 探讨智能体模型(agent-based model, ABM)在艾滋病流行病学中的应用价值。**方法** 介绍 ABM 的基本原理和模型特点,基于 NetLogo 构建以异性性行为为主体的人类免疫缺陷病毒(human immunodeficiency virus, HIV)传播模型,在不同 HIV 检测频率的条件下估算男男性行为(men who have sex with men, MSM)中的 HIV 流行参数。**结果** 模拟结果显示随着 HIV 检测频率的增加,年新增感染数和诊断比例明显改善,且当 HIV 检测频率为 1 次/年时,感染 HIV 并知晓病情的比例为 91%。**结论** 智能体模型适用于具有复杂传播网络的健康体系,可用于预测疾病传播趋势和比较不同政策组合的长期健康收益。

【关键词】 智能体模型 艾滋病 流行病学 政策评估

【中图分类号】 R511 **【文献标识码】** A **DOI** 10.11783/j.issn.1002-3674.2024.05.021

构建疾病传播的数学模型能够帮助人们更深入的洞察传染病的传播动力学以及不同干预策略的潜在效益^[1]。在过去的几十年间,随着国际间的联系日益紧密,传染病新发和流行的威胁也在增加。流行病学家致力于借助各种方法工具来预测疫情、制定相应防控对策,疾病传播模型的使用越来越受重视^[2]。传染性是一个包含大量异质性元素的复杂系统^[3],常用的建模方法包括系统动力学、网络分析和智能体建模。相比其他两种方法,智能体建模在构建实时动态系统、模拟多水平个体间的交互等方面有着独特的优势。随着计算机技术的发展,集成了易感-感染-恢复(susceptible-infected-recovered, SIR)模型和地理信息系统(geographic information system, GIS)等技术的智能体模型(agent-based model, ABM, 也称基于主体模型)在传染病领域取得了长足的发展^[4-5]。本文将从基本原理、建模特点等方面介绍 ABM,并结合艾滋病防治实例探讨其在动态建模分析中的应用。

基本原理

ABM 是一种通过微观层面上模拟智能体(agent)的行动,从而达到复现、表征以及深入理解人群现象的一种“自下而上”的模拟^[6-7]。agent 是一种智能实体(intelligent entities),可以代表个人或医院、学校等实体的集合^[8],智能体不受外部控制,基于交互规则自主采取行动。经过构建、校准和验证等过程后,agent 相互作用而产生传染事件,局部范围内的互作会产生人群水平的结果,如发病率、死亡率以及病毒传播途径等。

agent 之间的关系代表社会层面互动的过程,其行

为会相互干扰(interference),这导致了模拟结果是非独立性的。模拟人际互动从而阐明干扰对于系统的影响是 ABM 的预期目标和优势之一。ABM 可以实现 agent 之间关联的产生、解除或重建,这种具有社会互动意义的关联会产生系列动态网络结构。

模型设计和实现

根据 ABM 的定义,建模者需要明确研究目的,设定 agent 和相关的交互规则,同时在特定平台搭建和运行模拟,如图 1 所示。详细建模和分析步骤如下:

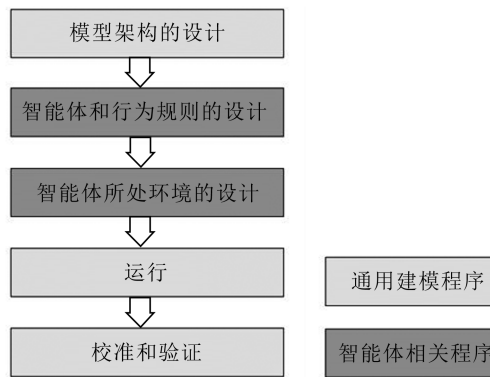


图 1 智能体模型设计和开发流程

1. 模型设计

ABM 是现实世界抽象的简化,遵循 Robert Axelrod 提出的 KISS(keep it simple, stupid)原则^[9]。建模者需要基于研究目的构建一个高度简化复杂系统的模型,在建模时应纳入相关传播机制,基于社会生态学理论的健康行为模型^[10]等概念框架以及因果分析图^[11]等有助于在建模时确定关键传播过程。模拟算法需遵循以下特征:①个体只能通过空间和时间上具有共同定位时相互作用和影响;②个体的健康状况变化可预先计算;③通常包含一个最短潜伏期 D_{min} 。

ABM 由一组代码在特定的计算平台中运行,编程语言包括 Python、Java 和 C++ 等,其运行速度较快,建

* 基金项目:国家科技基础资源调查专项(2019FY1011103)

1. 复旦大学公共卫生学院流行病学教研室,公共卫生安全教育部重点实验室(200032)

2. 上海市重大传染病和生物安全研究院

Δ通信作者:张铁军, E-mail: tjzhang@shmu.edu.cn

模灵活。或者基于已经开发的仿真平台,例如 NetLogo (<http://ccl.northwestern.edu/netlogo/>) 和 Repast (<https://repast.github.io/>) 等,建模者可基于各自的编程语言和模板功能进行相关研究,但会受限于运行速度和建模深度。

2. 定义 agent 和交互规则

属性是 agent 的特征变量,可随模拟时间改变(如疾病状态、年龄等),也可保持恒定(如性别)。agent 群体(N)中的个体 $i=1, \dots, N$ 被赋值 $m=1, \dots, M$ 一系列属性集合(如性别属性, $m=1$ 代表男性, $m=2$ 代表女性),在任一时间步长 $t(t=1, \dots, T)$, agent 集合可以使用 $N \times M$ 的矩阵 S_t 表示:

$$S_t = \begin{bmatrix} S_{t1,1} & S_{t1,2} & \dots & S_{t1,M} \\ S_{t2,1} & S_{t2,2} & \dots & S_{t2,M} \\ \vdots & \vdots & \ddots & \vdots \\ S_{tN,1} & S_{tN,2} & \dots & S_{tN,M} \end{bmatrix} \quad (1)$$

agent 交互行为通常被指定为一组 if-then 条件规则,个体 i 在任一时间步长 $t(t=1, \dots, T)$ 可与满足条件的 agent 集合($1, \dots, i-1, i+1, \dots, N$) 交互,从而以特定概率发生疾病传播事件。

3. 运行模拟

在运行时间步长 T 后,建模者可记录和分析属性变化结果,如疾病发病率、患病率和死亡率等。基于离散时间线的模型中个体的疾病暴露状态在持续改变,如单位时间步长后暴露状态下的 agent 数量为:

$$E(0, t+1) = \lambda \sum_{\tau=1}^{\tau_{max}} I(\tau, t) \quad (2)$$

其中 τ 代表处于感染期的时间长度。多次迭代模拟可以避免因 ABM 的随机性导致的单次输出异常值的产生。目前没有关于运行次数的明确统一指导,因而通常会提供变异性相关的数据。

4. 模型校准和验证

模型的校准和验证是一种通过调整可信度低的参数以保持模型输出与真实经验数据一致性的迭代的过程^[12]。校准后若模型输出和真实世界现象同等复杂,将会降低模型的可解释性,因此需要避免过度拟合(overfitting)。用于校准的参数通常是缺乏经验数据参考或真实性较低的参数^[13],用于校准的输出可以使用疾病相关变量,如发病率、死亡率等。

案例应用

本文以男男性行为(men who have sex with men, MSM)群体中人类免疫缺陷病毒(human immunodeficiency virus, HIV)的动态传播研究为例,介绍如何通过 NetLogo 6.2.0 软件构建 ABM。

1. 模型框架

拟探究 MSM 群体异性性行为介导的 HIV 感染动态变化趋势。HIV 的传播与高危性行为及性伙伴亲密关系的类型有关联^[14], agent 之间通过固定性关系和临时性关系构建连接,模型允许存在无交互及多个 agent 同时交互的情景。NetLogo 建模平台和内部代码见图 2。

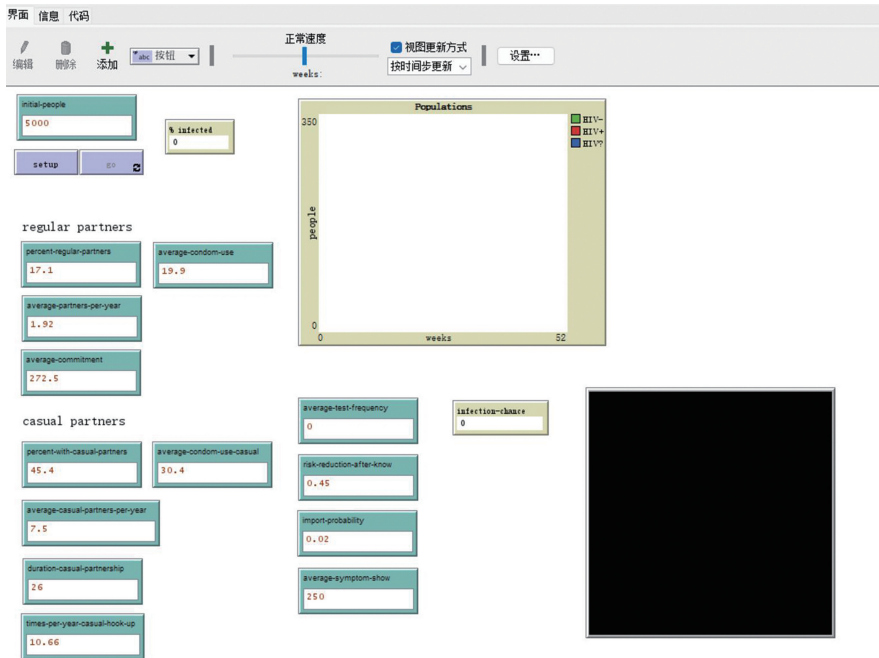


图 2 NetLogo 平台建模示意图

2. 属性及交互规则

单位时间步长被定义为一周。agent 被赋予 HIV 相关的可变疾病状态属性,包括未感染、感染但未被诊断、感染且被诊断。本例中任一时间步长下的 agent 数

量可使用以下差分方程表示:

$$S_{t+1} = S_t - \frac{\beta I_t S_t}{N} \quad (3)$$

$$E_{t+1} = E_t + \frac{\beta I_t S_t}{N} - \sigma E_t \quad (4)$$

$$I_{t+1} = I_t + \sigma E_t \quad (5)$$

其中 S_{t+1} 、 E_{t+1} 和 I_{t+1} 表示经过一个单位时间步长后

的易感者、感染但未被诊断(潜伏期)和感染且被诊断的 agent 数量; β 表示每次无保护性行为的感染概率; σ 表示从潜伏期到诊断的转移率, N 为研究人口总数,相关参数见表 1。

表 1 模型主要参数设置及其描述

模型参数	参数值	描述
Percent-regular-partners	61.9%	61.9%的 MSM 具有固定性伴侣 ^[15]
Average-commitment	272.5(周)	固定性关系的平均持续时间为 272.5 周 ^[16]
Average-condom-use	19.9% (14.1%~27.2%)	固定伴侣关系中的安全套使用概率为 19.9% ^[17]
Average-partners-per-year	1.92	在过去六个月中,MSM 的常规性伴侣平均数量为 0.96 ^[16] ,转换为每年平均 1.92 个常规性伴侣
Percent-with-casual-partners	45.4%	45.4%的 MSM 具有临时性伴侣 ^[15]
Average-casual-partners-per-year	7.5	过去六个月中,可以发生临时性关系的 MSM 平均拥有 3.75 个临时的性伴侣 ^[16] ,转换为平均每年 7.5 个临时性伴侣
Average-casual-condom-use	30.4% (20.5%~42.5%)	在临时性行为中安全套的使用概率为 30.4% ^[17]
Times-per-year-casual-hook-up	10.66	过去六个月中,与临时性伴侣发生性行为的次数平均值为 5.33 ^[16] ,转换为每年平均发生临时性行为 10.66 次
Risk-reduction-after-diagnose	0.45	调查表明约 45%的 MSM 在知晓感染 HIV 后会告知性伴侣 ^[18]
Percent-inject-drug	8%	约 8%的 MSM 在过去 12 个月内有过注射毒品的行为 ^[17]
Import-probability	0.02%	根据注射吸毒者每年感染的概率为 11.9% ^[18-19] ,估算得到每周因注射吸毒而感染的概率约为 0.02%
Infection-chance	-	每次无保护性行为的 HIV 感染概率,用于校准模型

3. 模型构建与校准

该模型在 NetLogo 模型库中的 HIV 经性传播 ABM^[20] 以及 Scott 等^[18] 的丙型肝炎病毒经性传播 ABM 模型的基础上修改完成。每次模拟 520 个单位时间步长(10 年),每组参数分别运行 100 次。使用 *nlrj* 包在 R 软件模拟分析,相应代码见附录。模型通过调整感染概率(infection-chance)使输出的 HIV 年发病率校准在 0.70/100~1.25/100 人年^[21]。

4. 运行结果

校准后的模型输出的 HIV 患病率和感染概率见图 3,黑色实线代表 HIV 患病率中位数,灰色条带表示相应的四分位数间距。HIV 患病率从初始的 1.770% 增至 10 年后的 11.965% (IQR: 10.398%~13.030%),年发病率约为 1.020%。校准后的感染概率呈右偏态分布,中位感染概率为 3.107% (IQR: 2.090%~4.595%),以此作为后续分析的先验分布。

不同情景下的模拟结果如图 4,在 10000 人的 MSM 模拟队列中,年新增感染人数随着 HIV 检测频率的增加呈下降趋势,表明主动进行 HIV 检测或定期大规模筛检可以显著降低 HIV 发病率。但是当检测频率为 2 次/2 年时,年新增感染人数的下降趋势减缓,在感染并知晓比例随时间的变化趋势中也可观察到相似结果,且当 HIV 检测频率为 2 次/2 年时,感染并知晓的人数约占总感染人数的 91%,达到世界卫生组织 2030 规划中的目标,这提示为控制 HIV 流行率,可在 MSM 人群中每年进行一次 HIV 检测。

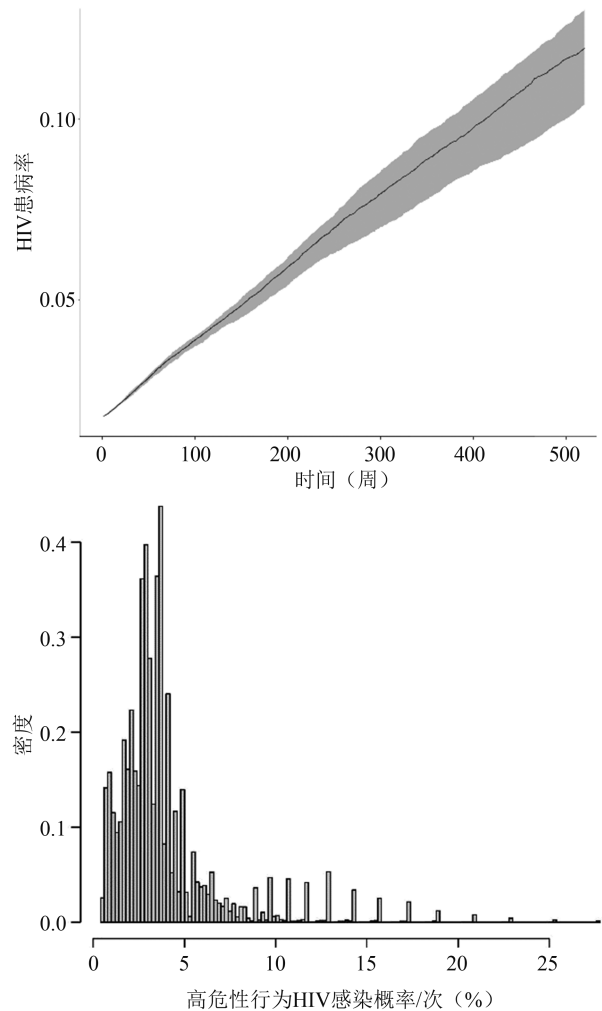


图 3 HIV 患病率趋势图和 HIV 感染率的密度直方图

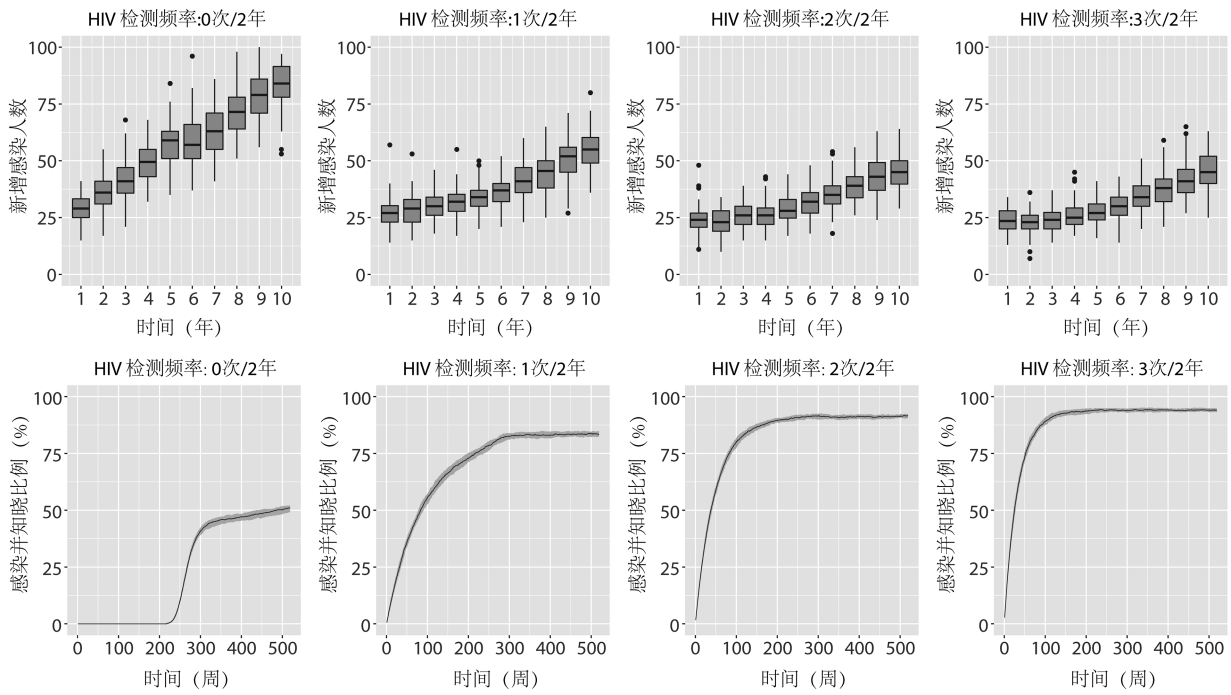


图4 MSM人群中新增感染人数和感染并知晓比例随时间变化图

优点与局限性

ABM的优点在于包含多样化属性的agent,可对影响健康结局的动态过程建模,从而基于微观层次的动力学产生宏观层次的复杂结果。ABM是基于离散事件的模拟,可帮助建模者理解复杂系统中的基本过程^[22]。ABM更聚焦于个体间的交互过程,而不仅仅是变量之间的数学关联。ABM可以修改微小的agent特征、行为、相互作用和疾病参数,实现真实世界中复合场景和政策的反事实(counterfactual)模拟,从而为政策制定者提供更全面的理论依据。

ABM的主要局限性在于建模时需要满足模型和参数的简化^[9],其可信度取决于模型是否充分地捕捉到现实运作的过程,且纳入研究的经验数据真实性和代表性无法保证。ABM在运行中会追踪和调度大量agent交互,也给ABM引入巨大的计算和参数负担。

讨论

本研究基于中国大陆MSM人群的性行为特征相关研究获取参数,依托NetLogo平台构建了时依性HIV病毒传播ABM,以尝试通过模型中的交互作用模拟HIV在MSM人群的传播过程。结果表明高危性行为的多重组合导致了MSM人群持续升高的HIV感染率,且每年一次HIV检测可以有效控制HIV的流行水平,模型有效捕捉性行为的异质性和性传播网络的高度动态特性。性传播疾病依赖感染者与易感者的接触而传播,ABM利用行为数据,生成个体间的性行为或

注射吸毒等接触行为,使HIV的模拟贴近实际场景,从而更深入的理解复杂传播系统,为HIV的防治策略提供理论基础。

ABM突出的核心特征包括涌现性(emergent)、随机性(stochasticity)和自主适应性(autonomy)等。涌现性指人群层面的健康结局(发病率、死亡率等)是agent交互所产生的不同的健康结局的集合,即将宏观现象看作各类异质性微观个体交互作用涌现的结果。随机性指以随机模拟为基础的ABM在初始条件下依据参数均值及其概率分布将不同属性数值随机赋予agent。自主适应性指暴露在疾病传播风险中的个体会以改变自身行为的方式极大地改变疾病传播的进程^[23],agent可根据当前环境和程序化的交互规则决定如何行动。

Grimm等^[24]在2010年提出了描述ABM的标准共识ODD(overview, design concepts, and details),以帮助研究者更好的理解和学习ABM。借助系统综述和meta分析等资源整合相关证据,合理利用人口普查数据和使用探索性模型检查结果的敏感性,可以避免模型可信度的限制。

综上,ABM通过模拟一段时间内异质性agent的交互,以疾病状态的转变反映群体健康水平的结局,通过对动态、非线性、异质性的时空过程的模拟和探索,促进对个体相互之间以及与环境交互方式的理解,适用于具有复杂传播网络的HIV等性传播疾病,可用于预测疾病传播趋势和比较不同政策组合的长期健康收益,因此对HIV/艾滋病防治具有重要意义。

参 考 文 献

- [1] Heesterbeek H, Anderson RM, Andreasen V, et al. Modeling infectious disease dynamics in the complex landscape of global health [J]. *Science*(New York, NY), 2015, 347(6227): aaa4339-aaa.
- [2] Chowell G, Sattenspiel L, Bansal S, et al. Mathematical models to characterize early epidemic growth: A review [J]. *Phys Life Rev*, 2016, 18: 66-97.
- [3] Mitchell M. Complexity: A Guided Tour [M]. Oxford University Press, 2009.
- [4] Gorman DM, Mezić J, Mezić I, et al. Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice [J]. *Am J Public Health*, 2006, 96(11): 2055-2060.
- [5] Epstein JM. Modelling to contain pandemics [J]. *Nature*, 2009, 460(7256): 687.
- [6] Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems [J]. *Proceedings of the National Academy of Sciences*, 2002, 99(suppl 3): 7280.
- [7] El-sayed AM, Scarborough P, Seemann L, et al. Social network analysis and agent-based modeling in social epidemiology [J]. *Epidemiol Perspect Innov*, 2012, 9(1): 1.
- [8] Tracy M, Cerdá M, Keyes KM. Agent-Based Modeling in Public Health: Current Applications and Future Directions [J]. *Annu Rev Public Health*, 2018, 39: 77-94.
- [9] Axelrod R. Advancing the Art of Simulation in the Social Sciences; proceedings of the Simulating Social Phenomena, Berlin, Heidelberg, F 1997//, 1997 [C]. Springer Berlin Heidelberg.
- [10] Singer M, Clair S. Syndemics and Public Health: Reconceptualizing Disease in Bio-Social Context [J]. *Medical Anthropology Quarterly*, 2003, 17(4): 423-441.
- [11] Greenland S, Pearl J, Robins JM. Causal Diagrams for Epidemiologic Research [J]. *Epidemiology*, 1999, 10(1).
- [12] Ngo TA, See L. Calibration and Validation of Agent-Based Models of Land Cover Change [M]//Heppenstall AJ, Crooks AT, See LM, et al. Agent-Based Models of Geographical Systems. Dordrecht; Springer Netherlands. 2012: 181-197.
- [13] Dowling R, Skabardonis A, Halkias J, et al. Guidelines for Calibration of Microsimulation Models: Framework and Applications [J]. *Transportation Research Record*, 2004, 1876(1): 1-9.
- [14] He J, Xu HF, Cheng WB, et al. Intimate relationship characteristics as determinants of HIV risk among men who have sex with regular male sex partners: a cross-sectional study in Guangzhou, China [J]. *BMC Infect Dis*, 2018, 18(1): 150.
- [15] 许娟. 四城市男男性行为人群性伴特征与艾滋病病毒/梅毒感染状况 [D]. 中国疾病预防控制中心, 2010.
- [16] Zhang L, Chow EPF, Wilson DP. Distributions and trends in sexual behaviors and HIV incidence among men who have sex with men in China [J]. *BMC public health*, 2012, 12: 546.
- [17] Chow EPF, Lau JTF, Zhuang X, et al. HIV Prevalence Trends, Risky Behaviours, and Governmental and Community Responses to the Epidemic among Men Who Have Sex with Men in China [J]. *BioMed Research International*, 2014, 2014: 607261.
- [18] Scott N, Stoové M, Wilson DP, et al. Eliminating hepatitis C virus as a public health threat among HIV-positive men who have sex with men: a multi-modelling approach to understand differences in sexual risk behaviour [J]. *J Int AIDS Soc*, 2018, 21(1).
- [19] Scott N, Doyle JS, Wilson DP, et al. Reaching hepatitis C virus elimination targets requires health system interventions to enhance the care cascade [J]. *Int J Drug Policy*, 2017, 47: 107-116.
- [20] Wilensky U. NetLogo HIV model [Z]. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. 1997. <http://ccl.northwestern.edu/netlogo/models/HIV>.
- [21] Zhang L, Chow EPF, Jing J, et al. HIV prevalence in China: integration of surveillance data and a systematic review [J]. *The Lancet Infectious Diseases*, 2013, 13(11): 955-963.
- [22] Axelrod R. The Complexity of Cooperation [M]. Princeton University Press, 1997.
- [23] Ceddia MG, Bardsley NO, Goodwin R, et al. A complex system perspective on the emergence and spread of infectious diseases: Integrating economic and ecological aspects [J]. *Ecological Economics*, 2013, 90: 124-131.
- [24] Grimm V, Berger U, Bastiansen F, et al. A standard protocol for describing individual-based and agent-based models [J]. *Ecol Model*, 2006, 198(1): 115-126.

附录:

```

Sys.setenv(JAVA_HOME="C:/Program Files/Java/jre1.8.0_231")
library(nlrx)
netlogopath <- file.path(D:/netlogo)
modelpath <- file.path("D:/ABM/HIV-1.nlogo")
outpath <- file.path("D:/ABM/output/test3")
nl <- nl(nlversion = "6.2.0",
        nlpath = netlogopath,
        modelpath = modelpath,
        jvmem = 1024)
nl@experiment <- experiment(expname = "test3-0.5",
                             outpath = outpath,
                             repetition = 1,
                             tickmetrics = 1,
                             idsetup = "setup",
                             idgo = "go",
                             runtime = 520,
                             evalticks = seq(1,520),
                             metrics=c("count turtles with [known?] / count turtles with [infected?]",
                                       "count turtles with [infected?] / count turtles",
                                       "count turtles with [infected?]"),
                             constants = list("initial-people" = 10000,
                                                "percent-regular-partners" = 17.1,
                                                "average-partners-per-year" = 1.92,
                                                "average-commitment" = 272.5,
                                                "average-condom-use" = 19.9,
                                                "percent-with-casual-partners" = 45.4,
                                                "average-casual-partners-per-year" = 7.5,
                                                "duration-casual-partnership" = 26,
                                                "times-per-year-casual-hook-up" = 10.66,
                                                "average-condom-use-casual" = 30.4,
                                                "average-test-frequency" = 0.5,
                                                "risk-reduction-after-know" = 0.45,
                                                "import-probability" = 0.02,
                                                "average-symptom-show" = 250,
                                                "average-infection-chance" = 3.1067))
nl@simdesign <- simdesign_simple(nl = nl, nseeds = 100)
eval_variables_constants(nl) #检查变量是否有效
print(nl)
results <- run_nl_all(nl)
setsim(nl, "simoutput") <- results
write_simoutput(nl)

```

(责任编辑:张悦)