

艾滋病疫情估计和预测方法研究进展

彭志行 汪宁 王璐 喻荣彬 丁国伟

【关键词】 艾滋病；疫情估计；疫情预测

Development of methods for estimation and prediction on epidemic situation of HIV/AIDS PENG Zhi-xing^{*}, WANG Ning, WANG Lu, et al. 'Department of Epidemiology & Biostatistics, Nanjing Medical University, Nanjing 210029, China

Corresponding author: WANG Ning, Email: wangnbj@163.com

【Key words】 Acquired immunodeficiency syndrome; Estimation on epidemic situation; Prediction on epidemic situation

自 1981 年美国首次在同性恋人群中检测出第一例艾滋病患者以来,全世界每年因艾滋病死亡的人数逐年上升, HIV 感染者/AIDS 患者(HIV/AIDS)已成为一个严重的公共卫生问题^[1,2]。截至 2007 年底,中国估计现存 HIV/AIDS 约 70 万,全人群感染率为 0.05%(0.04%~0.07%)^[3]。为准确掌握艾滋病流行状况和趋势,从而为政策制定、项目设计、干预措施评价以及资源配置提供依据,国内外在艾滋病疫情估计和预测方面进行了大量的理论和实践研究,取得了明显进展^[4]。

但由于艾滋病高危人群的特殊性、感染人群的隐匿性以及遭受的社会歧视,导致目前 HIV/AIDS 病例报告数和实际感染数差距较大^[5,6]。而我国不同地区 HIV/AIDS 的感染途径、源头人群、传播方式、社会人口学和行为学特征等各不相同,流行特点较其他国家和地区复杂得多^[7],因而深入研究并建立适合我国艾滋病疫情特点的疫情评估和预测模型就显得尤为必要。

世界范围内曾较广泛使用的 HIV/AIDS 疫情估计和预测的方法主要有:疫情模型(epimodel)、组分法(component model)、工作簿方法(workbook)、反向计算法(back-calculation)、亚洲艾滋病流行模型(Asian epidemic model, AEM)、德尔菲法(Delphi method)、估计与预测软件包(estimation and projection package, EPP)和 Spectrum 模型等。现综合大量国内外文献,按模型功能主要侧重于疫情估计或者预测将这些方法分为两大类对其统计学原理、应用背景和优缺点等方面进行综述。

1. 侧重于艾滋病疫情估计的方法: 艾滋病疫情估计

DOI:10.3760/cma.j.issn.0254-6450.2009.03.024

作者单位: 210029 南京医科大学流行病与卫生统计学系(彭志行、喻荣彬); 中国疾病预防控制中心艾滋病性病预防控制中心(汪宁、王璐、丁国伟)

通信作者: 汪宁, Email: wangnbj@163.com

是综合利用病例报告、哨点监测、专项调查等资料来估计艾滋病流行的实际情况,从群体水平测算艾滋病感染的实际例数。侧重于艾滋病疫情估计的方法有:反向计算法、组分法、工作簿方法、疫情模型等。

(1)反向计算法:该方法利用每年报告的艾滋病病例数(通常是对报告不完整或迟报进行调整后的数据)和每年 HIV 感染者发展为艾滋病病例的比率,反向推算每年有多少 HIV 感染者成为当年估计的艾滋病病例,从而得到 HIV 感染者人数的一种短期预测方法^[8,9]。其基本思想是运用 t 时刻的累积病例数 $A(t)$ 、 s 时刻的 HIV 发病率 $g(s)$ 和潜伏期分布函数 $F(t)$ 构成方程:

$$A(t) = \int_0^t g(s)F(t-s)ds$$

若病例数 $A(t)$ 已知(可以从病例报告中得到),潜伏期分布 $F(t)$ 可以经过流行病学调查估计得到,那么通过上方方程反积分可以估计 HIV 发病率 $g(s)$; 如果已知 HIV 发病率 $g(s)$ 和潜伏期分布 $F(t)$, 那么艾滋病病例数 $A(t)$ 可以用积分方程进行估计和预测。从 HIV 感染到艾滋病发病的潜伏期分布可以用参数法(一般认为艾滋病的潜伏期呈 Weibull 分布)或者非参数法。

反向计算法有两种,称为“旧法”和“新法”。两种方法的计算过程一致,区别在于输入数据和潜伏期定义的不同。“旧法”的潜伏期定义为“从感染 HIV 到发病的时间”,要求输入的数据为历年报告的艾滋病报告数。“新法”的潜伏期定义为“从感染 HIV 到被首次检测出 HIV 阳性的时间”,输入的数据为历年的 HIV 感染例数。“新法”中定义的潜伏期短,大大减少了抗病毒治疗的干扰作用,因而,能相对更加准确地估计 HIV 感染者数。

反向计算法仅利用 HIV 疫情报告数据即可进行 HIV/AIDS 疫情估计,为缺乏监测资料的国家(和地区)提供一个可供参考的疫情估计方法。疫情报告数据的准确性和代表性决定此方法的成败,因而适用于艾滋病疫情报告系统完善的国家,但在一些发展中国家,由于缺乏艾滋病死亡诊断设施、卫生信息系统相对落后以及政治、文化的限制等多种因素,使得艾滋病报告病例数远低于实际例数,从而造成艾滋病疫情估计过低,因此不适合这种方法^[10]。随着能够延缓或者阻止 HIV 感染者发展为艾滋病的有效药物治疗的出现,多数发达国家认为利用艾滋病报告病例数来估计 HIV 感染者反向计算法较为可信。

(2)组分法:是用来估算艾滋病流行现状的方法,其主要思想是:先将国家(或地区)可能的高危人群分组,某类高危人群的 HIV 感染者数=该类高危人群的估计基数×该人群

的HIV感染率;然后再将各类高危人群的估计感染例数进行累加,得到国家(或地区)的HIV感染者估计总数^[11]。该方法能比较直观地估计艾滋病的流行现状,操作简单明确,但在实际应用中,各类高危人群HIV感染率的代表性可能存在一定问题。在20世纪80年代中期该方法曾广泛应用于美国的HIV/AIDS疫情估计,我国也曾在1993—2001年主要采用该方法估计全国HIV/AIDS累计病例数^[12]。

(3)工作簿方法:是UNAIDS专家基于组分法原理开发的适用于聚集性流行或低流行地区的HIV/AIDS疫情估计方法。该方法是在Excel的基础上设计的,由疫情估计工作簿和流行曲线工作表两部分组成^[13]。应用该方法首先需根据当地艾滋病的流行情况,确定艾滋病流行集中的高危人群的种类,如注射吸毒者(IDU)、女性性工作者(FSW)、男男性行为者(MSM)和嫖客等,然后确定各类高危人群基数和HIV感染率的高值和低值,对当地存活的HIV感染例数进行估计。

在使用工作簿方法时,将每个人群基数的高值和低值与感染率的高值和低值分别相乘,求出人群基数低值×感染率低值(A值),人群基数低值×感染率高值(B值),人群基数高值×感染率低值(C值),人群基数高值×感染率高值(D值);将A值和B值的平均值、C值和D值的平均值分别作为最终估计感染例数的低值和高值。使用该模型需收集的数据包括:当地当年15~49岁成年人口数、性别构成、城镇人口比例;各类高危人群基数和感染率的高值和低值;各类低危人群(主要是高危人群的固定性伴)基数和感染率的高值和低值等^[14]。如果该地区没有可参考的感染率数据,则可借鉴文化、民族、经济、人口等方面相似地区的数据。

工作簿模型既能对艾滋病进行点估计从而了解当前的HIV/AIDS流行趋势,而且可以形成一条流行曲线来预测出短期(5年)内HIV/AIDS的发展趋势;另外该模型还具有自动检查系统,有助于消除错误使预测的结果更可靠。但是该方法对所利用数据的质量要求较高,不能很好地处理各类人群之间的相互作用和交叉感染。UNAIDS从2005年至今一直用该方法对亚洲、南美洲、东欧等地多个国家和地区的艾滋病疫情进行估计^[15,16]。我国2003年开始使用该方法在全国层面上进行疫情估计,2004年艾滋病监测系统得到完善并达到世界较高水平;使可利用的数据增加,开始用工作簿方法对全国的艾滋病疫情进行分省估计,2005年及2007年我国的艾滋病疫情估计和预测都运用该方法^[3];使疫情估计工作达到世界先进水平,并引导WHO/UNAIDS重新认识全球的艾滋病疫情。

(4)疫情模型:是由WHO在20世纪80年代末开发的一个简单的计算机程序,也是一个简单的艾滋病动力学模型,它基于单一传染源在人际之间的传播服从伽马分布的原理,假设任何人群中的艾滋病累计感染情况也服从伽马分布,并随着时间的推移其分布曲线有一条很长的右尾^[17]。但在实际应用中,曲线峰值以后的下降趋势不能很好地代表真实的HIV流行情况^[18]。

疫情模型主要用于资料不完整地区的过去和现时的艾滋病患病率的估计以及艾滋病病例和死亡的短期预测。1997年UNAIDS将疫情模型用于区域性模型对全球各国的艾滋病疫情进行了估计,并利用估计所得数据及艾滋病开始流行的时间生成一条能较好地描述各国艾滋病传播的流行曲线。这次估计未能很好地刻画各国之间及各地区之间艾滋病流行的多样性,使估计结果过高,但是却使我们认识到艾滋病在世界不同地区的流行有着较大差异,世界不同地区艾滋病监测系统的差距及不足^[19]。

2. 侧重于艾滋病疫情预测的方法:艾滋病疫情预测是利用预测模型等方法预测未来艾滋病流行的总体趋势、不同人群的流行趋势等。在国内外主要用于艾滋病疫情估计的方法有:AEM、德尔菲法、EPP和Spectrum模型等。

(1)德尔菲法:德尔菲专家咨询法是基于专家的知识、经验和分析判断,在历史和现实有关资料综合分析基础上,对未来事物的变动趋势作出一种定量和定性相结合的预测、评价方法。一般采取匿名的方式广泛征求专家的意见,经过反复多次的信息交流和反馈修正,使专家的意见逐步趋向一致,最后根据专家的综合意见,对所研究问题的现状进行判断,对未来发展中的各种可能出现的前景作出预测。

运用德尔菲法主要包括三个阶段:准备阶段、轮番征询阶段和结果处理阶段。准备阶段主要完成四个方面的工作,即明确预测主题和目的、选择专家、准备背景材料、设计调查咨询表。轮番征询专家意见一般分为四轮进行,各轮的反馈不是简单的重复,每一轮专家都可能得到新的信息,通过不断在前一轮基础上的交流,专家的意见一般可达到基本一致的程度。咨询结果统计处理和每一轮的专家意见处理常用几种统计方法: $M \pm$ 四分位数法、算术平均统计处理法、主观概率法和非量化评价结果的统计处理法;各项指标应结合专家权威程度、熟悉程度和积极系数进行权重处理^[19]。

总之,选择专家是应用德尔菲法的关键,如果专家选择良好,则预测结果的权威性高。在2008年应用德尔菲法进行疫情预测的实践中我们发现,专家咨询表的设计和结果处理也是至关重要的,表格的设计必须做到言简意赅、表达准确、解释清楚;而且该方法也可应用于各种亚人群规模和危险因素估计,得出对疫情预测非常有价值的建议。但该方法是一种主观预测和估计方法,结果的准确性较难测量。专家的选择不是一件容易的事情,专家意见未达成一致时,结果的准确性和可靠性不够好;另外,如果严格按照德尔菲法的原理进行估计,需要通过几轮的调查,费时费力^[20]。

(2)EPP模型:EPP是由UNAIDS专家组在2001年开发的一个常用的预测模型,代替了先前所用的疫情模型,该模型分不同人群单独绘制流行曲线,通过不同年份的HIV感染率数据,进行流行病学模拟,得到一条反映HIV感染率的时间序列的最佳拟合曲线;然后再与不同人群的流行曲线相结合,生成一条反映该地区HIV感染率水平的流行曲线^[21]。

EPP模型利用要满足监测资料的4个参数:① t_0 :HIV/AIDS开始流行的年份,起始年份越早则拟合的曲线起点越

早;② r : HIV 感染传播的速率, r 值越大说明感染率增加越迅速;③ f_0 : 流行的初期成年人中高危人群所占的比例, 该参数决定流行曲线的峰值;④ ϕ : 高危因素调整参数, 该参数决定成年人进入高危人群的比例随时间变化的情况, 如果该值大于零, 说明人群感染的危险性随着时间在增加, 感染率下降较慢或维持在一个较高水平上; 如果该值等于零, 说明危险人群比例是一个常数, 感染率曲线在峰值后随患者死亡而下降; 如果该值小于零, 说明人群在流行开始后其危险性逐渐降低。

模型为这 4 个参数提供了标准值, 但可以根据国家的具体特征进行校正, 直到拟合出反映成年人 HIV 感染率随时间演化发展的最佳曲线^[22]。2005 年经过改进的 EPP 最新版本适用于广泛性流行及聚集性流行地区。处于广泛流行的地区, 其估计以产前门诊的监测数据为基础; 处于聚集性流行地区, 其估计以特定人群的检测数据为基础, 如 MSM、IDU 等。2005 年 UNAIDS 采用 EPP 模型对很多国家的艾滋病疫情进行了预测分析。

但是该方法的缺点是如果所利用的数据的质量及代表性差, 将在很大程度上影响疫情的估计及预测^[21]。在实际应用中我们发现, 安全套的推广使用、治疗方法的改进、人群规模由于疫苗的使用而改变等, 这些不确定性都能影响估计的结果, 使操作者下结论时需要更加谨慎。

(3) Spectrum 模型: 这是一个策略模型, 核心是人口统计学预测 (DemProj), 同时包括其他模块的预测和分析: 艾滋病影响模型 (AIM)、计划生育 (FamPlan)、效益-成本 (Benefit-cost)、高生育率和人口增长的社会经济影响 (RAPID)、预防 HIV 母婴传播 (PMTCT) 等。其中以 DemProj 为基础, DemProj 是一个用于国家或地区人口预测的计算机程序, 与其他的 Spectrum 模块结合使用可以预测艾滋病的人口统计学影响; AIM 是预测艾滋病流行影响的计算机程序, 模型可以读取由 EPP 产生的感染率预测值, 并结合 HIV 感染者发病死亡规律资料及人口学估计数据, 根据假定的成年人 HIV 流行率预测 HIV 感染者未来的数量、艾滋病病例数和死亡数, 主要应用于许多非洲、拉丁美洲和亚洲国家^[24]。

Spectrum 的预测过程是首先利用 DemProj 进行人口统计学预测, 收集数据, 建立假设, 然后录入数据, 开始 HIV/AIDS 预测; 所有 AIM 的计算预测过程都依赖于对未来 HIV 流行率的估计, 在 EPP/Workbook 的基础上直接计算和预测所需结果, 为政策的制定提供有用的信息^[18]。

在疫情预测的具体工作中我们发现, Spectrum 使用的模型参数和指标基本都是在一段时间内相对稳定的指标, 这对预测工作是比较有利的。而且, 随着越来越多的 HIV 感染者接受 ART 治疗, 预测中要对国家 (或地区) 未来的 ART 治疗信息做进一步的研究和假设。

(4) AEM 模型: 目前在泰国、印度、马来西亚等大多数亚洲国家中, IDU、FSW、MSM 是导致该地区 HIV 传播和流行的主要因素。而随着流行的不断发展, HIV 由 FSW 传染给嫖客以及 IDU 传染给异性性伴, 逐渐成为该地区经异性性途径

传播的重要组成部分。AEM 就是基于对亚洲国家的这些流行特点理解的基础上开发的^[25]。与其他模型相比, AEM 是一种半经验化的过程模型。该模型通过内部的数学拟合过程, 动态模拟和重现驱动亚洲 HIV 的流行情况, 从而预测 HIV 的各种流行结果。

AEM 主要由两部分组成, 一个是需要进行安装的可执行文件 (EXE 文件); 另一个是用来进行基本参数和数据输入及预测结果输出的 Excel 文件。EXE 文件是 AEM 的核心, 主要功能包括总体比较不同方面的数据, 调整各种传播参数和影响因素, 拟合出有用的曲线和数据结果。同时, AEM 也需要更广泛的 HIV 流行病学和行为学指标数据输入, 构成三个密切联系的输入板块: 一类是当地的 HIV 流行病学相关指标; 另一类是影响当地 HIV 传播的行为学指标; 第三类是相对变化不大的生物学参数。AEM 主要考虑 15~49 岁人群中的 HIV 传播和感染, 儿童中 HIV 感染例数可以通过女性人群的感染率和生育率得到。

AEM 软件的产出有直接的数值输出和图形输出。数值输出: 包括不同年份成年人及儿童 HIV 感染者的估计数 (新发感染例数、累计感染例数、当前存活的感染例数); 不同年份成年人及儿童的艾滋病估计病例数、死亡估计数; 不同年份关于 HIV、艾滋病和死亡估计数的年龄分布; 不同年份不同传播途径的新发感染估计数; 以及根据上述输出数值计算出来的不同人群中的 HIV 患病率和发病率估计值。图形输出: 包括不同人群中的发病率图、HIV 感染例数图以及传播途径的绝对数构成和相对比例的构成图等^[26]。

相对于其他 HIV/AIDS 预测模型, AEM 的优势在于: AEM 能为应用地区提供较多的疫情数据, 为当地的艾滋病防治机构提供较有针对性的策略建议; 模型输入数据是以年为单位序列数据, 能够及时地把握 HIV 流行的动态, 增加预测结果的真实性, 参数和数据均是根据当地实际情况进行输入和调整, 拟合的流行曲线更能反映当地的流行趋势; 模型可以预测出不同行为水平下的流行情况, 从而可以比较不同干预措施的效果, 为决策者制定更有效的干预策略提供依据。

但是, AEM 最大的局限性是对数据的需求较大。因而数据的收集、评估和筛选过程非常重要, 在决定使用该模型之前应该对当地有关数据的可获得性和代表性进行评估。只有具备了充足的数据来源, 才能够确定真正适合的参数, 否则, 预测结果会极大地偏离当地的实际情况^[27]。而且, 使用该模型进行预测时, 当地的主要传播途径都应包含在模型之中, 否则可能会导致错误的预测结果。2005 年 UNAIDS 首次用 AEM 对泰国和柬埔寨的艾滋病疫情进行了估计和预测, 我国云南和广西等地区也在 2004 年开始使用 AEM 对艾滋病疫情进行预测, 并尝试用模型评价各种干预措施的效果, 2008 年我国首次在国家层面上运用 AEM 进行艾滋病疫情的预测。

3. 小结: HIV/AIDS 疫情估计和预测方法的选择、设计和应用与该地区监测资料的完整程度密切相关。在疫情估

计和预测方法选择的时候,首先需要对该地区 HIV/AIDS 流行情况及其各类影响因素进行分析,探索其流行规律,并对现有的监测数据、专题调查、国内外合作项目或者其他途径获得的各类人群数据进行评估。只有在此基础上选择适宜的方法,才能得到科学准确的疫情估计和预测结果。在艾滋病的预防与控制的工作中,这些方法发挥了重要作用。但由于资料来源的日益多样性,以及 HIV/AIDS 的流行形势的变化,监测系统的日益完善,疫情估计和预测的各种方法与模型都需要不断的改进、完善,才能为艾滋病防治工作的顺利进行提供可靠的信息,不断适应艾滋病防治工作的需要。

参 考 文 献

- [1] Hayes R, Weiss H. Epidemiology. Understanding HIV epidemic trends in Africa. *Science*, 2006, 311(5761): 620-621.
- [2] 汪宁. 我国艾滋病预防控制的形势与面临的挑战. *中华预防医学杂志*, 2004, 38(5): 291-293.
- [3] 国务院防治艾滋病工作委员会办公室, 联合国艾滋病中国专题组. 中国艾滋病防治联合评估报告(2007). 2007.
- [4] UNAIDS Reference Group on Estimates, Modelling and Projections. Improved methods and assumptions for estimation of the HIV/AIDS epidemic and its impact: Recommendations of the UNAIDS Reference Group on Estimates, Modelling and Projections. *AIDS*, 2002, 16(9): W1-14.
- [5] Rangsin R, Chiu J, Khamboonruang C, et al. The natural history of HIV-1 infection in young thai men after seroconversion. *J Acquir Immune Defic Syndr*, 2004, 36(1): 622-629.
- [6] 曹晓彬, 庞琳, 吴尊友, 等. AIDS 相关歧视产生的原因、表现形式及消除策略. *中国艾滋病性病*, 2005, 11(3): 235-240.
- [7] Cock KM, Weiss HA. The global epidemiology of HIV/AIDS. *Trop Med Int Health*, 2000, 5(7): A3-9.
- [8] Joshua AS, Christopher JLM. Modeling HIV/AIDS epidemics in Sub Saharan Africa using seroprevalence data from antenatal clinics. *Bull WHO*, 2001, 79(7): 596-607.
- [9] Downs AM, Heisterkamp SH, Brunet JB, et al. Reconstruction and prediction of the HIV/AIDS epidemic among adults in the European Union and in the low prevalence countries of central and Eastern Europe. *AIDS*, 1997, 11(5): 649-662.
- [10] 中国性病艾滋病防治协会. 中国非政府组织预防控制艾滋病工作通讯. 2007: 16.
- [11] 郑锡文. 流行病学进展(第八卷). 北京: 中国科学技术出版社, 1995: 32-35.
- [12] National Research Council. AIDS, sexual behavior and intravenous drug use. Washington, DC: National Academy Press, 1989.
- [13] Lu F, Wang N, Wu Z, et al. Estimating the number of people at risk for and living with HIV in China in 2005: methods and results. *Sex Transm Infect*, 2006, 82 (Suppl 3): iii87-91.
- [14] 中国疾病预防控制中心性病艾滋病预防控制中心. HIV/AIDS 疫情估计与预测方法. 2004.
- [15] UNAIDS, WHO. The workbook method manual [DB/OL]. <http://data.unaids.org/Topics/Epidemiology/Manuals/rkBookManual-05-en.pdf>, 26/4/2005.
- [16] Walker N, Stover J, Stanecki K, et al. The workbook approach to making estimates and projecting future scenarios of HIV/AIDS in countries with low level and concentrated epidemics. *Sex Transm Infect*, 2004, 80 (Suppl 1): i 10-13.
- [17] Chin J, Lwanga SK. Estimation and projection of adult AIDS cases: a simple epidemiological model. *Bull WHO*, 1991, 69(4): 399-406.
- [18] Schwartländer B, Stanecki KA, Brown T, et al. Country-specific estimates and models of HIV and AIDS: methods and limitations. *AIDS*, 1999, 13(17): 2445-2458.
- [19] 蔡辉, 张颖, 倪宗赞. Delphi 法中评价专家的筛选. *中国卫生事业管理*, 1995, 1: 49-50.
- [20] Harold A, Linstone, Murray Turoff. The Delphi method: Techniques and Applications. 2002.
- [21] Brown T, Grassly NC, Garnett G, et al. Improving projections at the country level: the UNAIDS estimation and projection package 2005. *Sex Transm Infect*, 2006, 82 (Suppl 3): iii34-40.
- [22] Rehle TM, Shisana O. Epidemiological and demographic HIV/AIDS projections: South Africa. *African J AIDS Res*, 2003, 2(1): 1-8.
- [23] Somi GR, Matee MI, Swai RO, et al. Estimating and projecting HIV prevalence and AIDS deaths in Tanzania using antenatal surveillance data. *BMC Public Health*, 2006, 6: 120.
- [24] Stover J. Projecting the demographic consequences of adult HIV prevalence trends: the Spectrum projection package. *Sex Transm Infect*, 2004, 80 (Suppl 1): i 14-18.
- [25] Chin J, Bennett A, Mills S. Primary determinants of HIV prevalence in Asian-Pacific countries. *AIDS*, 1998, 12 Suppl B: S87-91.
- [26] Brown T, Peerapatanapokin W. The Asian epidemic model: a process model for exploring HIV policy and programme alternatives in Asia. *Sex Transm Infect*, 2004, 80 (Suppl 1): i 19-24.
- [27] 喻达, 刘民, 明中强. 一种探索 HIV 预防控制政策和项目有效性的过程模型. *中国艾滋病性病*, 2008, 14(1): S96-97.

(收稿日期: 2008-08-28)

(本文编辑: 尹廉)