

中国2013—2014年艾滋病高危人群干预工作网络直报数据质量分析

李力 曹晓斌 罗巍 王常合 吴尊友 柔克明

102206 北京, 中国疾病预防控制中心性病艾滋病预防控制中心预防干预室

通信作者: 柔克明, Email: kemingrou@chinaaids.cn

DOI: 10.3760/cma.j.issn.0254-6450.2017.09.011

【摘要】 **目的** 分析中国2013—2014年艾滋病高危人群干预工作网络直报中出现的错误数据, 产生错误的原因, 探讨提高干预数据质量的办法。**方法** 整理2013—2014年全国艾滋病综合防治数据信息系统中高危人群干预信息报表的错误数据记录, 包括县(区)自查和国家审查错误, 采用描述性统计方法分析其分布、类型和产生原因。**结果** 2013—2014年共修正干预数据错误1 066个, 年均登记错误率 $<0.1\%$, 每年上半年出现的错误占全部登记错误的71.3%(760/1 066), 东中西部地区分别占全部登记错误的14.9%(159/1 066)、22.6%(241/1 066)和62.5%(666/1 066), MSM和吸毒人群中的干预数据错误多于暗娼人群。两年全部干预数据错误中, 县(区)自查发现的占86.4%(921/1 066), 国家审核发现的占13.6%(145/1 066), 出错最多的数据项是HIV本年首次检测人数和干预覆盖人数。常见的错误原因有漏报与延误、工作人员失误、计算错误和理解错误等。**结论** 总体上干预数据的登记错误率较低, 可通过完善干预活动记录, 加强统计和直报环节的数据质量控制, 定期培训信息岗位工作人员等措施, 来提高干预数据质量。

【关键词】 艾滋病; 高危人群; 网络直报; 数据质量

Quality of network direct reporting of information about intervention service in population at high risk for HIV infection in China, 2013–2014 Li Li, Cao Xiaobin, Luo Wei, Wang Changhe, Wu Zunyou, Rou Keming

Division of Prevention and Intervention, Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China

Corresponding author: Rou Keming, Email: kemingrou@chinaaids.cn

【Abstract】 **Objective** To analyze the problems in the network direct reporting of information about intervention service conducted in population at high risk for HIV infection in China during 2013–2014, and provide evidence for the improvement of the network direct reporting of the intervention information. **Methods** The wrong records of the intervention service in population at high-risk were collected from national AIDS prevention and treatment information system. The wrong records, including those found at county (district) level and those found at state level, were analyzed with descriptive statistical method. **Results** A total of 1 066 wrong records were found during 2013–2014, and average annual wrong record rate was $<0.1\%$. Up to 71.3% (760/1 066) of wrong records occurred in the first half year. The wrong records in eastern, central and western areas accounted for 14.9% (159/1 066), 22.6% (241/1 066) and 62.5% (666/1 066) of the total respectively. More wrong records were found in the intervention information for men who have sex with men and injecting drug users than in those for female sex workers. Among the total wrong records, 86.4% (921/1 066) were found at county level and 13.6% (145/1 066) were found at state level. The wrong records were mainly “annual number of persons receiving the first HIV test” and “annual number of persons covered by intervention”. Common causes of wrong records were underreporting and delay, staff fault, miscalculation and misunderstanding. **Conclusion** In general, the wrong record rate in intervention information for population at high risk for HIV infection reported directly through network was low in China. It is necessary to strengthen the recording of the intervention service, the quality control of statistics and network direct reporting and staff training to improve the quality of reported data of intervention service.

【Key words】 AIDS; High risk population; Network direct reporting; Data quality

我国自2007年艾滋病综合防治数据信息管理系统建立以来, 以县(区)为单位, 每月网络直报的艾

滋病高危人群干预数据(干预数据)一直是监测干预工作数量与质量, 评估预防干预效果, 发现艾滋病防

治难点的重要依据之一^[1]。经过多年实践调整,高危人群干预工作网络报表的可采集性较好,依靠直报系统中用户权限、报告时限及逻辑校验等约束条件,数据的及时性和完整性有一定保障。

然而,不是所有的网络直报干预数据质量问题都能通过预置条件加以避免。一方面,常规防治工作干预的艾滋病高危人群,主要包括暗娼人群、MSM和吸毒人群等,一般是不易接触到的敏感人群,有很高的隐私保密需求,这就决定了干预数据以收集人群信息为主,干预活动一旦结束,想要追溯和还原当时收集的信息非常困难,给发现数据问题和督导数据质量带来了挑战。另外,正确理解报表中数据项的含义,及时准确地记录、收集、汇总和直报干预数据,涉及人员培训、分工合作、干预实施、数据管理等多个环节,在任一环节上干预数据的准确性都会面临考验。

为数不多的研究关注了中国艾滋病疫情报告^[2]、实验室检测^[3]、随访管理^[4]、抗病毒治疗^[5]等数据的质量,但未见干预信息的质量报告。本研究主要从干预数据的准确性入手,分析中国暗娼、MSM和吸毒人群等高危人群干预工作网络直报过程中出现的数据错误,产生错误的原因,探讨提高干预数据质量的办法。

资料与方法

1. 资料来源:①2013—2014年艾滋病综合防治数据信息系统的高危人群干预信息报表中,主要经县(区)自查发现的错误数据记录(自查错误)。自查错误是数据上报后发现的错误,一般是县(区)CDC在回顾、检查、统计及使用本县(区)已上报数据时发现了错误,由县(区)CDC填写错误数据修改单,注明错误原因,提交省级CDC和中国CDC审核,审核通过的登记并修正,形成了自查错误数据记录。②2013—2014年艾滋病综合防治数据信息系统的高危人群干预信息报表中,经审核发现的错误数据记录(审核错误)。每月1—11日国家审核全国当月上报数据,发现存疑数据,主要是逻辑错误和异常值等,立即反馈各省,各省联系县(区)核对,确认有误的登记并修正,形成国家审核数据错误记录。

2. 研究方法:对记录资料进行整理,提取错误数据发生的时间(年份、月份),空间(东、中和西部地区)和人群(暗娼、MSM和吸毒人群)信息。其中,东部地区指北京、天津、上海、江苏、浙江、福建、山东、广东、辽宁9个省份,中部地区指河北、山西、海南、

吉林、黑龙江、安徽、江西、河南、湖北、湖南10个省份,西部地区指内蒙古、广西、重庆、四川、贵州、云南、西藏、陕西、甘肃、青海、宁夏、新疆12个省份。保留错误数据的来源类型(自查、审核),错误的的数据项(干预覆盖人数、HIV阳性检测人数等),梳理每个数据错误的主要成因(延迟统计和报告、工作人员失误、计算错误等)。

3. 统计学分析:利用Excel 2010和SAS 9.0软件对资料做描述性统计分析。描述错误数据的时间、空间及人群分布。采用最大值、构成比和直观线图展示2013—2014年错误数据的时间分布情况;采用排序和构成比描述错误数据的空间和人群分布。按照错误发现途径,分类描述自查和审核错误的分布情况,包括最常出错的数据项。阐述数据错误的主要原因,并分人群展示其频数和构成比。

结果

1. 高危人群干预工作信息报表概况:2013年全国有2 808个县(区)上报了高危人群干预工作信息报表,登记并修正的数据错误为588个;2014年全国有2 810个县(区)上报了高危人群干预工作信息,登记并修正的数据错误为478个,两年共修正干预数据错误1 066个。高危行为干预工作信息报表中,县(区)每月最多上报高危人群干预工作数据26个(暗娼人群9个、MSM人群9个和吸毒人群8个),全年最多上报工作数据312个,全国年均上报干预数据约87万个,年均登记错误率<0.1%。

2. 数据错误分布:2013—2014年全国高危人群干预网络直报数据登记错误最多的月份是5月,共152个,占两年登记错误的14.3%(152/1 066),最少的是10月,共16个,占1.5%(16/1 066);前两季度发生的错误占登记错误的71.3%(760/1 066)。见图1。

空间上,西部地区高危人群干预网络直报登记

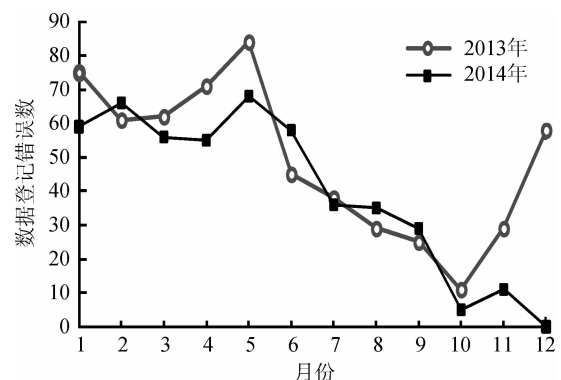


图1 2013—2014年全国干预数据登记错误的月份分布

数据错误最多,2013和2014年分别登记297个和369个,共666个,占两年登记错误的62.5%(666/1 066)。其次是中部地区,2013和2014年分别登记数据错误166个和75个,共241个,占两年登记错误的22.6%(241/1 066)。登记数据错误最少的是东部地区,2013和2014年分别登记数据错误125个和34个,共159个,占两年登记错误的14.9%(159/1 066)。

人群上,2013年登记错误数据最多的是MSM人群,占当年登记错误的41.5%(244/588),最少的是暗娼人群,占当年登记错误的27.9%(164/588)。2014年全国登记错误数据最多的是吸毒人群(包括参加针具交换的吸毒者),占当年登记错误的49.1%(235/478),最少的是暗娼人群,占当年登记错误的18.8%(90/478)。

3. 数据错误类型:从错误发现途径上,高危人群干预网络直报数据的错误分为两类:自查错误和审核错误,自查发现的数据错误为主。2013—2014年登记的自查错误分别有485个和436个,共921个,占两年登记错误的86.4%(921/1 066),其中自查发现错误最多的是MSM人群,占全部自查错误的40.1%(369/921)。2013—2014年登记的审核错误分别有103个和42个,占两年登记错误的13.6%(145/1 066),其中审核发现错误最多的是吸毒人群,占51.0%(74/145)。见表1。

2013年自查错误中出错最多的数据项是HIV本年首次检测人数(104个),2014年自查错误中出错最多的数据项是干预覆盖人数(82个)。2013年和2014年审查错误中,出错最多的数据项都是HIV检测阳性人数,分别为62个和40个。

4. 错误原因:干预数据出错的主要原因,一是延迟统计和报告,即部分数据未在规定的报表截止时间内完成统计和上报,占全部错误原因的34.9%(321/921),例如有的是因为“漏了2个社区的干预数据”、“统计漏了同伴教育员干预人数”或者“依托MSM感染者开展(干预),未及时上报所有数据”。二是工作人员失误,占全部错误原因的21.8%(201/921),包括键盘输入错误、填错行、填错列、填错月和

漏填了等,例如“录数据时将8错看成3”、“工作人员录入错误,听电话(时,将)‘没有一个’听成了‘有一个’”或者“录入时漏填一个零”。三是计算错误,占21.2%(195/921),包括重复、误算等,例如“把(其他)高危人群中的检测数重复报到针具交换(吸毒者)报表中”、“将2例VCT(自愿咨询检测)数填报成暗娼数”或者“QQ群里有部分人不是MSM,也有部分人有两个号”。此外,对数据项不理解、概念不清楚也常导致错误,尤其是在工作人员发生变动或报表调整的时候,例如“(工作人员)对报表(定义)理解不够(清楚)”或者“未理解到此新增项目(HIV本年首次检测人数)的意思”。见表2。

表2 2013—2014年全国干预数据错误的主要原因^a

错误原因	暗娼人群	MSM人群	吸毒人群	合计
延迟统计和报告	45(21.3)	133(36.0)	143(41.9)	321(34.9)
工作人员失误	52(24.7)	61(16.5)	88(25.8)	201(21.8)
计算错误	45(21.3)	103(27.9)	47(13.8)	195(21.2)
理解错误	27(12.8)	17(4.6)	12(3.5)	56(6.1)
缺失 ^b	13(6.2)	5(1.4)	8(2.3)	26(2.8)
其他	29(13.7)	50(13.6)	43(12.6)	122(13.2)
合计	211(100.0)	369(100.0)	341(100.0)	921(100.0)

注:^a数据来源于自查错误的原因登记,括号外数据为例数,括号内数据为构成比(%);^b未填写错误原因

讨 论

分析2013—2014年高危人群干预工作网络直报数据的错误发现,登记的总错误率较低,有进一步减少错误的空间。每年上半年发生并修正的干预数据错误多于下半年,这与每年三季度各省和国家为提高防治工作信息质量,开展艾滋病综合防治数据质量评估有关。一是上半年的数据正在评估范围内,受关注度高且使用较多,发现和修正错误的可能性大;二是经历评估后一些数据问题暴露出来,为下半年减少错误打下基础;三是评估能起到为工作人员答疑解惑的作用,对提升数据统计和报告质量有宜;然而,不排除评估结束后核查和修正数据的频率下降等其他可能性导致的下半年修正错误减少。中西部地区的干预数据错误多于东部地区,或与前者的艾滋病防治和信息工作基础较后者薄弱有关。重点减少MSM和吸毒人群的直报错误,及早发现并修正错误,将提高干预数据的质量。

干预数据错误主要来自3个方面:
①干预活动记录错误,包括记录不完整、不及时、不准确和不真实等。②数

表1 2013—2014年全国干预数据自查错误和审核错误的人群分布

人群	自查错误		合计	审核错误		合计
	2013年	2014年		2013年	2014年	
暗娼	127(26.2)	84(19.3)	211(22.9)	37(35.9)	6(14.3)	43(29.7)
MSM	225(46.4)	144(33.0)	369(40.1)	19(18.4)	9(21.4)	28(19.3)
吸毒	133(27.4)	208(47.7)	341(37.0)	47(45.6)	27(64.3)	74(51.0)
合计	485(100.0)	436(100.0)	921(100.0)	103(100.0)	42(100.0)	145(100.0)

注:括号外数据为例数,括号内数据为构成比(%)

据统计错误,包括重复、错算、漏算和误解数据项等。③网络直报错误,包括填报不及时,键盘输入错误,网络不畅、中断、响应缓慢等问题造成乱码或数据上传存储失败等。从这三个方面加强数据质量管理,尤其是减少统计报告的延误和工作人员的失误,有助于提高干预数据质量。此外,信息人员频繁更换且缺乏相应培训,极易损害数据质量,蔡睿等^[6]对预防艾滋病母婴传播信息质量控制体系的研究中有类似发现。加强对信息岗位工作人员的定期培训,使其对数据定义的理解准确一致,确保各级直报网络通畅及数据安全,跟上信息系统的设置、操作和管理等方面的更新步伐,提供必要的技术支持等,都有利于干预数据质量。

干预数据质量有赖于基层从事干预和信息工作的人员,在实地开展干预服务的同时,还要负责干预数据的收集、整理、汇总和上报,身兼数职在精力上有可能顾此失彼。一个办法是通过购买服务充分发挥社区基层卫生服务机构和社会组织的力量,细化分工,提高资源利用效率,使基层信息人员专注于数据管理和质量控制。不过,从社会购买服务也隐含了一些问题,如督导和评估等难度可能增加,因此,数据质量的重要性将越发突出。减少潜在影响干预数据质量的动机,科学完善绩效评价机制,是干预数据质量的长效保障。

利益冲突 无

参 考 文 献

[1] 董薇,周楚,葛琳,等. 2008—2014年中国预防艾滋病经性传播干预措施落实情况分析[J]. 中华流行病学杂志, 2015, 36(12): 1337-1340. DOI: 10.3760/cma.j.issn.0254-6450.2015.12.003.
Dong W, Zhou C, Ge L, et al. Implementation of intervention programs on AIDS-related sexual transmission in China[J]. Chin J Epidemiol, 2015, 36(12): 1337-1340. DOI: 10.3760/cma.j.issn.0254-6450.2015.12.003.

[2] 秦倩倩,王璐,丁正伟,等. 2010年中国艾滋病疫情网络直报质量分析及指标探讨[J]. 疾病监测, 2012, 27(12): 1007-1012. DOI: 10.3784/j.issn.1003-9961.2012.12.024.
Qin QQ, Wang L, Ding ZW, et al. Evaluation of AIDS network direct reporting quality and related indicators in China, 2010[J]. Dis Surveill, 2012, 27(12): 1007-1012. DOI: 10.3784/j.issn.1003-9961.2012.12.024.

[3] 郭志宏,丁晓贝,张佳峰,等. 2014年浙江省《艾滋病实验室信息管理系统》应用结果分析[J]. 中国艾滋病性病, 2015, 21(10): 846-848, 854. DOI: 10.13419/j.cnki.aids.2015.10.05.
Guo ZH, Ding XB, Zhang JF, et al. Analysis of application results of national HIV laboratory information management system in Zhejiang in 2014[J]. Chin J AIDS STD, 2015, 21(10): 846-848, 854. DOI: 10.13419/j.cnki.aids.2015.10.05.

[4] 韩晶,汤后林,许娟,等. 2009—2013年全国艾滋病防治随访管理数据质量评估结果分析[J]. 中华预防医学杂志, 2014, 48(12): 1104-1106. DOI: 10.3760/cma.j.issn.0253-9624.2014.12.017.
Han J, Tang HL, Xu J, et al. Evaluation of HIV/AIDS follow-up management data quality in China in 2009-2013[J]. Chin J Prev Med, 2014, 48(12): 1104-1106. DOI: 10.3760/cma.j.issn.0253-9624.2014.12.017.

[5] 赵德才,马焯,张福杰,等. 基于客户端/服务器结构的全国艾滋病抗病毒治疗信息系统的升级改造[J]. 中国艾滋病性病, 2013, 19(6): 445-447. DOI: 10.13419/j.cnki.aids.2013.06.020.
Zhao DC, Ma Y, Zhang FJ, et al. Upgrading of the national antiretroviral therapy information system based on client/server structure in China[J]. Chin J AIDS STD, 2013, 19(6): 445-447. DOI: 10.13419/j.cnki.aids.2013.06.020.

[6] 蔡睿,郭光萍,张燕,等. 预防艾滋病母婴传播信息质量控制体系对数据质量的影响分析[J]. 中国妇幼保健, 2009, 24(35): 4951-4953.
Cai R, Guo GP, Zhang Y, et al. Analysis on effects of quality control system of prevention of mother-to-child transmission of AIDS information on quality of data[J]. Matern Child Health Care China, 2009, 24(35): 4951-4953.

(收稿日期:2017-03-26)

(本文编辑:斗智)