

# 2011—2013年全国传染病报告信息管理系统中乙型肝炎病例重复报告现状分析

郑徽 王富珍 张国民 缪宁 孙校金 崔富强

100050 北京, 中国疾病预防控制中心免疫规划中心 流行病学二室(郑徽、王富珍、张国民、缪宁、孙校金), 免疫规划中心(崔富强)

通信作者: 崔富强, Email: cui fuq@126.com

DOI: 10.3760/cma.j.issn.0254-6450.2016.09.013

**【摘要】** **目的** 分析全国传染病报告信息管理系统(National Notifiable Disease Report System, NNDRS)中2011—2013年乙型肝炎(乙肝)报告病例中的重复报告现状,了解重复报告对乙肝报告发病的影响。**方法** 在NNDRS中,按照“报告地区浏览”下载“发病日期”在2011年1月1日至2013年12月31日期间所有报告病种为乙肝的个案,剔除“删除个案”后,根据识别个案的5项主要指标(身份证号码、姓名、性别、出生日期、现住址编码)的不同组合,形成6种查重方案,通过SAS 9.1.3软件编程实现重复病例查找。**结果** 2011—2013年NNDRS报告的3 389 374例乙肝个案中,重复病例249 047例,重报率为7.35%,其中6种查重方案可发现重报率范围为0.88%~6.38%。以2011年个案库为基准,3年重报率分别为2.91%、8.11%和11.72%。重报个案中,年度内重复报告占33.04%(71 709例),跨年重复报告占66.96%(145 329例)。去除重报个案后,2013年新疆维吾尔自治区和青海省乙肝报告发病率可分别下降30.80/10万和26.26/10万。**结论** 2011—2013年NNDRS报告乙肝病例中约有7.36%的重报,重复报告以跨年重报为主。乙肝重复报告对新疆、青海等省份的发病率统计影响较大。

**【关键词】** 乙型肝炎; 传染病报告

**基金项目:** 国家科技重大专项(2008ZX10002001; 2012ZX10002001)

**Analysis on duplicate reporting of hepatitis B in national notifiable communicable disease report system in China, 2011–2013** Zheng Hui, Wang Fuzhen, Zhang Guomin, Miao Ning, Sun Xiaojin, Cui Fuqiang

2<sup>nd</sup> Epidemiology Division (Zheng H, Wang FZ, Zhang GM, Miao N, Sun XJ), Department of National Immunization Program (Cui FQ), Chinese Center for Disease Control and Prevention, Beijing 100050, China  
Corresponding author: Cui Fuqiang, Email: cui fuq@126.com

**【Abstract】** **Objective** To analyze the duplicated reporting of hepatitis B in the national notifiable communicable disease report system (NNDRS) in China during 2011–2013. **Methods** The data of hepatitis B reported from January 1, 2011 to December 31, 2013 were collected from NNDRS, the duplicated reporting were searched with software SAS 9.1.3 according to the cases' ID number, name, gender, birth data and current address code. **Results** During 2011–2013, a total of 3 389 374 hepatitis B cases were reported through NNDRS, but the duplicated reporting rate was 7.35% (249 047). According to 6 protocols, the duplicated reporting rate ranged from 0.88% to 6.38%. If using the data in 2011 as the baseline, the duplicated reporting rate in the three years was 2.91%, 8.11% and 11.72%, respectively. In the duplicated repeating, 33.04% (71 709) occurred in the same year, and 66.96% (145 329) occurred in other year. Without the duplicated repeating, the hepatitis B incidence in Xinjiang and Qinghai provinces in 2013 could decreased by 30.80/100 000 and 26.26/100 000, respectively. **Conclusions** The analysis indicated that the duplicated repeating rate of hepatitis B was 7.36% in NNDRS during 2011–2013, and the duplicated reporting mainly occurred in other years. The duplicated reporting of hepatitis B in NNDRS had greatly affected the accuracy of surveillance data in some provinces, such as Xinjiang and Qinghai.

**【Key words】** Hepatitis B; Infectious disease report

**Fund programs:** National Science and Technology Major Project of China (2008ZX10002001, 2012ZX10002001)

乙型肝炎(乙肝)病毒在我国的感染率较高,是我国重点防控的四个重大传染病之一<sup>[1]</sup>。由于慢性乙肝患者病情迁延,反复就诊、异地就诊等现象普遍,因而在全国传染病报告信息管理系统(NNDRS)中易造成重复报告(重报)。目前临床医生对急、慢性乙肝病例存在混乱报告现象,系统中报告的急性乙肝病例也可能存在重报。我国2004年建立NNDRS,对促进传染病监测的完整性和及时性起到了至关重要的作用,但由于该系统的查重功能仅限在年度内病例中查找,因此对于类似乙肝等存在跨年度反复报告的慢性传染病的监测数据中重报问题一直未能解决。为了解NNDRS中全国乙肝病例重报情况,初步了解重报对NNDRS报告乙肝发病率的影响,本研究对2011—2013年NNDRS中上报的乙肝病例个案进行查重分析,从而为评估我国乙肝报告现状提供参考。

### 资料与方法

1. 数据来源:在NNDRS中,按照“报告地区浏览”下载2011年1月1日至2013年12月31日(发病日期统计)报告病种为乙肝的全部个案信息,剔除“已删除个案”。

2. 个案定义:①原始个案:从NNDRS中按照要求下载的全部乙肝个案(含实验室诊断、临床诊断、病原携带、阳性检测和疑似病例上报的所有急性、慢性和未分类乙肝病例)。②有效分析个案:在原始个案中,NNDRS的自动统计功能仅分析“实验室诊断”和“临床诊断”乙肝病例,因此本文将原始个案中的“实验室诊断”和“临床诊断”乙肝病例个案定义为有效分析个案。

3. 数据处理:乙肝疫情报告卡记录的基本特征信息包括身份证号码、姓名、出生日期、性别、现住址编码等。对于个案信息中的姓名汉字,利用数据库函数,将汉字转换为16进制计算机编码,作为姓名的替代特征信息。将3年个案数据库整理为相同的数据结构后合并数据,应用建立的不同查重标准进行重复数据查询。

4. 查重标准:根据查找的精准程度不同,将个案特征信息变量进行组合,设定6种查重标准进行数据库中的个案比对:(A)身份证号码信息相同(不考虑其他信息因素),即视为重复报告病例;(B)出生日期相同+性别相同+姓名16进制编码完全相同+现住址前6位相同的个案;(C)出生日期相同+性别相同+姓名16进制编码完全相同;(D)出

生日期相同+现住址前6位相同+姓名16进制编码2/3以上相同;(E)姓名16进制编码相同+性别相同+现住址前6位一致+出生日期相差在12个月内;(F)姓名16进制编码2/3以上相同+性别相同+现住址前6位一致+出生日期相差在24个月内。

5. 查重方法:从合并后的总数据库中任取一条个案记录R1,按照6种查重规则与总库中剩余个案逐一对比,将所有方法查出的重复个案取出另建新数据集。将重复个案的新数据集中任一条记录,按照6种查重规则再与总库中剩余个案逐一对比,将比对后所有查出的重复个案取出再另建新数据集,将此过程重复循环3次,标记所有与R1重复的个案。每组重复个案中以报告日期最早的个案为首次报告,其余为重复报告。

6. 统计学分析:使用SAS 9.1.3软件实现数据库查重过程。重报1次计1例,重报2次计2例,其余类推。使用多种查重方案查找出的同一重复个案计1例。重报率(%)=重报病例数/报告病例数×100。矫正发病率(/10万)=(报告病例数-重报病例数)/人口数×100 000。

### 结 果

1. 基本情况:2011—2013年NNDRS中共上报3 389 374例原始乙肝个案(包含病原携带者123 046例、临床诊断病例352 355例、实验室诊断病例2 895 056例、阳性检测172例和疑似病例18 745例),其中有效分析个案3 247 411例。原始个案库的5个主要查重指标中,26.54%(899 666例)个案有身份证号码信息,其中2011年244 382例(7.21%),2012年309 170例(9.12%),2013年346 114例(10.21%)。其余4项指标填写率均为100%。

2. 病例重复报告情况:原始个案库中,应用6种查重方案能够发现29 777~216 388例重复报告个案,6种方法查找的3年重报率为0.88%~6.38%。将查找结果合并,共有249 047例重报病例,3年重报率为7.35%;重报病例分别为34 882例、95 636例和118 529例,重报率分别为2.91%、8.11%和11.72%(表1)。

按照本研究查重方法设置,2011—2013年有效分析乙肝个案中,急性乙肝病例216 499例,其中5.57%(12 065例)为重报病例;慢性乙肝病例2 054 753例,其中7.91%(162 567例)为重报病例;未分类乙肝病例976 159例,其中6.58%(64 275例)为重报病例。以2011年数据库为基准纵向比较急

表 1 2011—2013 年 NNDRS 原始乙肝病例个案中重报情况

查重方案	2011 年(原始个案 1 199 504 例)		2012 年(原始个案 1 178 843 例)		2013 年(原始个案 1 011 027 例)		合计(原始个案 3 389 374 例)	
	重报例数	重报率(%)	重报例数	重报率(%)	重报例数	重报率(%)	重报例数	重报率(%)
A	2 343	0.20	10 450	0.89	16 984	1.68	29 777	0.88
B	4 173	0.35	16 818	1.43	25 932	2.56	46 923	1.38
C	7 072	0.59	23 389	1.98	34 813	3.44	65 274	1.93
D	6 922	0.58	22 435	1.90	32 972	3.26	62 329	1.84
E	21 863	1.82	62 710	5.32	77 169	7.63	161 742	4.77
F	30 053	2.51	84 019	7.13	102 316	10.12	216 388	6.38
合计 <sup>a</sup>	34 882	2.91	95 636	8.11	118 529	11.72	249 047	7.35

注: <sup>a</sup>同一重报病例可用 A~F 六种方法中的 1 种或多种方法查出, 合计中将多种方案查找出的同一重复个案计为 1 例重报

性乙肝的重报情况, 2012 年急性乙肝病例中 6.08% 为既往报告病例, 2013 年急性乙肝病例中 8.58% 为既往报告病例(表 2)。

3. 重复报告病例的重报次数分析: 2011—2013 年 NNDRS 原始报告乙肝病例个案中, 有 217 038 例个案报告次数 ≥ 2 次, 其中 190 777 例(87.90%) 被报告 2 次; 21 949 例(10.11%) 被报告 3 次; 3 342 例(1.54%) 被报告 4 次; 970 例(0.45%) 被报告 ≥ 5 次(表 3)。被多次报告的个案中, 年度内重复报告个案有 71 709 例(33.04%), 其中 2011—2013 年度内重复报告病例分别为 26 146、23 892 和 21 671 例, 分别占当年报告病例总数的 2% 左右。剩余 145 329 例(66.96%) 个案均为跨年度多次报告(表 3)。

4. 重报对乙肝报告发病率的影响分析: 在有效

分析乙肝个案中, 删除重复报告个案, 2011—2013 年乙肝报告发病率分别从 79.45/10 万、80.68/10 万和 71.12/10 万下降到 77.13/10 万、74.12/10 万和 62.79/10 万, 降幅分别达 2.92%、8.13% 和 11.71%。其中, 3 年急性乙肝报告发病率分别从 5.60/10 万、5.66/10 万和 4.70/10 万下降到 5.46/10 万、5.32/10 万和 4.30/10 万, 降幅分别达 2.50%、6.01% 和 8.51%(表 4)。

删除 2013 年全国 31 个省(自治区、直辖市)报告乙肝病例中的重报病例, 对各省(市、自治区)乙肝报告发病率进行矫正。新疆、青海、山西、广东和福建乙肝报告发病率经矫正后分别从 190.5/10 万、231.2/10 万、137.4/10 万、131.9/10 万和 126.1/10 万, 下降到 159.7/10 万、204.9/10 万、115.3/10 万、115.7/10 万和 112.5/10 万, 分别下降了 30.80/10 万、26.26/10 万、

表 2 2011—2013 年 NNDRS 中有效分析乙肝病例的重报情况

年份	有效分析个案数				有效个案中的重报病例							
	急性	慢性	未分类	合计	急性		慢性		未分类		合计	
					重报例数	重报率(%)	重报例数	重报率(%)	重报例数	重报率(%)	重报例数	重报率(%)
2011	75 861	695 790	375 896	1 147 547	1 926	2.54	21 127	3.04	10 399	2.77	33 452	2.92
2012	76 855	727 326	327 123	1 131 304	4 669	6.08	62 660	8.62	24 670	7.54	91 999	8.13
2013	63 783	631 637	273 140	968 560	5 470	8.58	78 780	12.47	29 206	10.69	113 456	11.71
合计	216 499	2 054 753	976 159	3 247 411	12 065	5.57	162 567	7.91	64 275	6.58	238 907	7.36

表 3 2011—2013 年 NNDRS 中报告乙肝病例重报次数情况

报告次数	重复报告例数	年度内被重报病例(例)				年度间重报例数				
		2011 年	2012 年	2013 年	合计	2012 年与 2011 年	2013 年与 2012 年	2013 年与 2011 年	2011—2013 年	合计
2	190 777	24 750	22 736	20 566	68 052	48 997	42 512	31 216	-	122 725
3	21 949	1 251	1 061	1 013	3 325	5 189	4 277	3 006	6 152	18 624
4	3 342	120	88	82	290	741	526	372	1 413	3 052
5	697	22	7	8	37	119	100	55	386	660
6	182	2	0	2	4	19	19	19	121	178
7	53	1	0	0	1	7	8	3	34	52
8	15	0	0	0	0	2	2	0	11	15
9	9	0	0	0	0	0	0	0	9	9
10	6	0	0	0	0	2	0	1	3	6
11	2	0	0	0	0	0	0	1	1	2
12	3	0	0	0	0	0	0	0	3	3
13	2	0	0	0	0	0	0	0	2	2
23	1	0	0	0	0	0	0	0	1	1
合计	217 038	26 146	23 892	21 671	71 709	55 076	47 444	34 673	8 136	145 329

表4 2011—2013年删除重报病例后  
乙肝报告发病率变化(/10万)

年份	原始报告发病率				矫正发病率			
	急性	慢性	未分类	合计	急性	慢性	未分类	合计
2011	5.60	48.38	27.55	79.45	5.46	46.91	26.79	77.13
2012	5.66	51.09	23.93	80.68	5.32	46.69	22.13	74.12
2013	4.70	46.28	20.15	71.12	4.30	40.51	18.00	62.79

22.08/10万、16.25/10万和13.61/10万。北京、天津和江苏矫正后乙肝发病率下降小于1/10万(图1)。

### 讨 论

我国于2004年1月1日起正式启动NNDRS,法定传染病直接通过网络报告,实现了疫情监测资料从县级到国家级通过计算机网络的实时报告与管理<sup>[2]</sup>。自网络直报以来,我国乙肝报告发病一直处于较高水平。尽管全国乙肝血清流行病学调查结果显示,人群HBsAg流行率已显著下降<sup>[1,3]</sup>,但NNDRS中乙肝年报告发病人数和发病率仍未出现明显下降,乙肝报告病例历年均高达100万左右,年发病率平均80/10万~100/10万,在全国法定报告的传染病中历年排前3位。NNDRS中的乙肝病例报告准确性不高,重报、误报等诸多问题严重影响监测数据质

量已成为共识。

本文对2011—2013年所有NNDRS报告的乙肝个案进行重报分析,3年系统中存在7.35%的重报病例,以2011年乙肝个案数据库为基准,重报病例比例逐年上升,从2011年的2.92%上升到2013年的11.71%。重复报告病例以跨年度重复报告为主,占66.96%;年度内重复报告占33.04%。福建、河南和内蒙古等地区均对乙肝病例的重复报告做过研究,由于查重标准设置不同,重报率有所差异,范围浮动在2.00%~18.75%<sup>[4-8]</sup>。郑家耿等<sup>[8]</sup>对三亚市医疗机构乙肝病例报告质量核实发现,12.7%的复诊乙肝病例也作为初诊病例进行网络直报,提示乙肝病例的重复报告主要是因为病例反复就诊、异地就诊等原因造成。此外,本文分析发现,除慢性乙肝病例存在反复报告外,急性乙肝病例中也存在5.57%的重复报告,且有跨年重报现象,提示目前乙肝监测系统中存在着误报问题,推测部分临床医生将慢性乙肝的急性发病期患者作为急性乙肝病例进行网络直报,但还有待于进一步现场证实。

本次分析发现,目前NNDRS系统中各省报告乙肝发病率的准确性差异较大,去除重报病例后,新

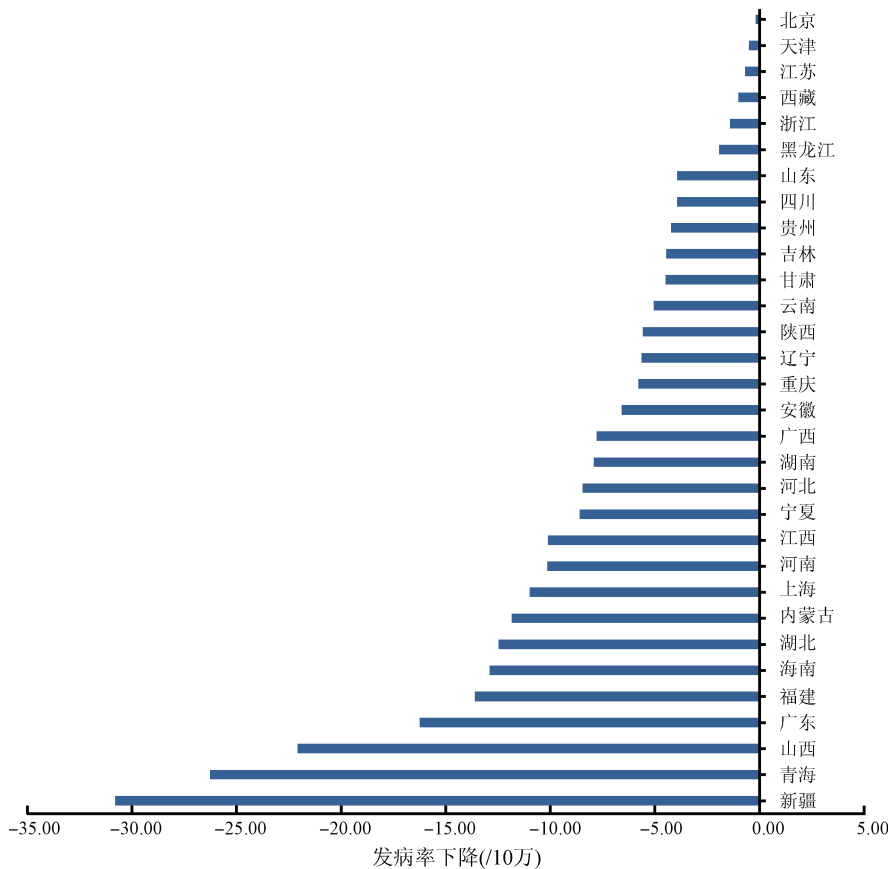


图1 2013年我国乙肝报告发病率与调整后的发病率变化情况



疆、青海、山西、广东等地的2013年乙肝发病率显著下降,而北京、天津、江苏等地的乙肝发病率较准确,重复报告比例较低,表明在不同省份医疗机构执行乙肝病例报告标准的尺度不一致,因病例重报而造成部分省份乙肝高发,提示在开展基于NNDRS的乙肝监测及数据分析工作中,应重点关注新疆、青海等地区。NNDRS数据库中的乙肝病例重报比例随着年份的跨度增加而增高,因而推测如增加数据库查重的年份,将进一步降低全国以及高发省份的乙肝报告发病率。

本文首次尝试利用全国乙肝病例个案的大数据进行重报分析,并设置了多重查找标准,最大限度利用计算机发现可能存在的重复病例,从方法学上具有一定的借鉴意义。但本文也存在一定的局限性:首先,“姓名”汉字的书写对数据库重复病例查找影响较大。本研究将数据库中姓名汉字转化为计算机语言进行比较分析,但由于汉字中有同音不同字或字形相同读音不同的现象,尤其是少数民族姓名是否含有“·”或空格等符号,均对转换后的姓名变量产生影响,在一定程度上可能造成低估查重结果,为了尽量减少影响,参照董蒲梅等<sup>[5]</sup>的方法将姓名变量2/3以上相同作为查重变量之一。第二,“出生日期”是查重的重要指标之一,本文考虑到中国人对年龄有虚岁的表述以及阴历等对“出生日期”准确性的影响,将查重标准放宽到“出生日期相差24个月内”,但与卢爱桃等<sup>[7]</sup>使用“年龄浮动在3年以内”结果相比较,本文的查重标准更严格,因而对全国乙肝重报估计还存在低估的可能性。第三,本文结论中对预先设置的6种查重方案结果进行合并统计,在一定程度上会扩大了查重过程中的错误判断,从而高估了重报比例。因此在查重标准设置时尽量严格设置,从而降低累积的错判风险。第四,本文未能通过开展相应现场调查对数据库查重结果进行外部核实,因而从方法学上还有待于进一步探讨。

综上所述,我国2011—2013年NNDRS报告乙肝病例中存在7.35%的重复病例,且3年重报率分别为2.91%、8.11%和11.72%。乙肝重报病例以跨年重复报告为主。乙肝重复报告对新疆、青海等省份的发病率统计影响较大。

利益冲突 无

#### 参 考 文 献

[1] 卫生部疾病预防控制局,中国疾病预防控制中心. 全国人群乙

型肝炎血清流行病学调查报告[M]. 北京:人民卫生出版社,2011:52-56.

Bureau of Disease Prevention and Control Department, Ministry of Health of the People's Republic of China, Chinese Center for Disease Control and Prevention. The national report of hepatitis B sero-survey in China [M]. Beijing: People's Medical Publishing House, 2011: 52-56.

[2] 马家奇,王丽萍,戚晓鹏,等. 2004年法定传染病报告信息质量分析[J]. 疾病监测, 2005, 20(5): 264-266. DOI: 10.3784/j.issn.1003-9961.2005.05.017.

Ma JQ, Wang LP, Qi XP, et al. Quality analysis on the reports of notifiable diseases in 2004 [J]. Dis Surveill, 2005, 20(5): 264-266. DOI: 10.3784/j.issn.1003-9961.2005.05.017.

[3] 戴志澄,祁国明. 中国病毒性肝炎——血清流行病学调查(上卷):1992—1995[M]. 北京:科学技术文献出版社,1997:39-59.

Dai ZC, Qi GM. Viral hepatitis in China sero-serum epidemiology survey (part A): 1992-1995 [M]. Beijing: Scientific and Technological Literature Publishing House, 1997: 39-59.

[4] 张廷良. 镇平县2009年以来乙肝病例重复报告情况调查[J]. 中国现代药物应用, 2014, 8(10): 243-244.

Zhang TL. The hepatitis B duplicate report in Zhenping county from 2009 [J]. Chin J Mod Drug Appl, 2014, 8(10): 243-244.

[5] 董蒲梅,张肖肖,范军星,等. 河南省六个县(区)2010年报告乙型肝炎病毒性肝炎病例分析[J]. 中国疫苗和免疫, 2013, 19(5): 444-447, 472.

Dong PM, Zhang XX, Fan JX, et al. Analysis on hepatitis B in six counties/districts in Henan Province in 2010 [J]. Chin J Vac Immun, 2013, 19(5): 444-447, 472.

[6] 黄文龙,陈长华,洪荣涛. 福建省传染病网络直报重卡及解决方法探讨[J]. 实用预防医学, 2010, 17(2): 272-274. DOI: 10.3969/j.issn.1006-3110.2010.02.023.

Huang WL, Chen CH, Hong RT. Discussion on case repeated report in notifiable disease report system in Fujian province [J]. Pract Prev Med, 2010, 17(2): 272-274. DOI: 10.3969/j.issn.1006-3110.2010.02.023.

[7] 卢爱桃,郭卫东,任利民,等. 内蒙古地区乙型肝炎报告质量调查及分析[J]. 中国实用医药, 2010, 5(29): 239-240. DOI: 10.3969/j.issn.1673-7555.2010.29.212.

Lu AT, Guo WD, Ren LM, et al. Survey and analysis of hepatitis B cases in Inner Mongolia [J]. China Pract Med, 2010, 5(29): 239-240. DOI: 10.3969/j.issn.1673-7555.2010.29.212.

[8] 郑家耿,李大章,杨海文,等. 三亚市乙型肝炎报告质量调查[J]. 中国热带医学, 2008, 8(4): 613-615. DOI: 10.3969/j.issn.1009-9727.2008.04.044.

Zheng JG, Li DZ, Yang HW, et al. Results of report quality of hepatitis B infections in Sanya city [J]. China Trop Med, 2008, 8(4): 613-615. DOI: 10.3969/j.issn.1009-9727.2008.04.044.

(收稿日期:2016-03-08)

(本文编辑:斗智)