

# 公共卫生危机信息媒体监测平台的建立与运用

袁俊, 王鸣, 郝爱华, 刘于飞, 毛新武, 甘建哲, 秦鹏哲, 蔡文锋, 李海麟, 景钦隆

**摘要:** **目的** 为了进一步丰富公共卫生监测信息的来源,提高疾病预防控制系统对突发公共卫生事件的早期预警能力,充分利用媒体资源信息,开展早期干预,及时应对。**方法** 利用网络平台技术及 RSS 新闻聚合器,对目前互联网 1000 多个中文新闻媒体网站进行监测,搜索大多数公共卫生危机信息进行计算机智能分类并及时发布于网站。**结果** 该系统建立能达到预期目标,可加载大部分网站的公共卫生相关新闻,并及时在公共卫生风险沟通平台(www.riskcommunication.org.cn)分类发布公共卫生危机信息。系统工作人员接受性较好。不同事件新闻可实现按地区、事件类别、时间、点击率进行分类排列,并提供 RSS 信息订阅。**结论** 该系统能根据公众和专业技术人员所处地区、关心的事件种类、关心热点等提供量身订做的公共卫生危机信息,作为公共卫生信息聚合和媒体监测的工具,为预防疾病、早期发现事件、及时采取应急响应等均可以起到一定作用。

**关键词:** 监测; 公共卫生; 互联网; 危机管理

中图分类号:R174<sup>+1</sup>

文献标识码:A

文章编号:1003-9961(2009)08-0559-04

**Establishment and application of media surveillance system for public health risk information** YUAN Jun, WANG Ming, HAO Ai-hua, LIU Yu-fei, MAO Xin-wu, GAN Jian-zhe, QIN Peng-zhe, CAI Wen-feng, LI Hai-lin, JING Qin-long. Guangzhou Municipal Center for Disease Control and Prevention of Guangdong Province, Guangzhou 510080, China

**Corresponding author:** WANG Ming, Email: wangming@gzcdc.org.cn

**Abstract:** **Objective** To increase the source of public health surveillance information, improve the early warning ability for public health emergency in CDC system, make full use of media information and conduct early intervention and timely response. **Methods** By using net-web platform technique and RSS (Really Simple Syndication), the surveillance on more than 1000 Chinese news media websites in Internet was conducted to search public health risk information, and the information was classified by computer and timely released on website. **Results** The expectancy to establish this surveillance system has been realized, public health risk information can be obtained from the most of media websites and released timely on the public health risk communication website (www.riskcommunication.org.cn) after classification. The system has good acceptability. The information can be classified by region, type, time and visiting rate, and the RSS information service can be provided. **Conclusions** The media surveillance system can provide the needed public health information to the public or professional personnel by area, interest and hot point. As a tool of public health information collection and media surveillance, the system will play a role in disease prevention, early detection and timely response of public health emergency.

**Key words:** surveillance; public health; internet; risk management

开展媒体监测即开展基于新闻报道的公共卫生相关信息监测对于专业人员和公众及时了解公共卫生领域最新进展可起到一定的作用。国内外不少专家和学者都进行过公共卫生媒体监测这方面的尝试和研究<sup>[1-4]</sup>。公共卫生媒体监测信息来源于新闻报道,明显不同于疾病常规监测来源于医疗和疾病预防控制机构的特点,为疾病常规监测提供

多方位线索,因此在获取线索、开展预警和及时了解群众对于事件的各种反应方面,媒体监测包含着更加丰富的内容。

为了进一步丰富公共卫生监测信息的来源,提高疾病预防控制系统对突发公共卫生事件的早期预警,充分利用媒体资源,早期干预,及时开展应对,笔者结合目前计算机网络信息技术和公共卫生工作进展,建立了基于网络的公共卫生危机信息媒体监测平台。

## 1 方法

**1.1 解决技术方案** 本系统是基于目前较为流行的 RSS 新闻聚合服务、强大的新闻搜索引擎以及应用

基金项目:中国/WHO 合作项目(No. WP/2008/CHN/02;2702)

作者单位:广东省广州市疾病预防控制中心,广东 广州 510080

作者简介:袁俊,男,湖南省人,主要从事现场流行病学和突发公共卫生事件应急处置工作

通信作者:王鸣,Tel:020-83872497,Email: wangming@gzcdc.org.cn

收稿日期:2008-11-24

计算机智能识别分类功能开发的媒体信息收集分类发布系统。RSS 是在线共享内容的一种简易方式(也叫聚合内容, Really Simple Syndication), 通常在时效性比较强的内容上使用 RSS 订阅能更快速获取信息, 目前大部分主流网站均提供 RSS 输出, 有利于用户获取网站内容的最新更新。广泛使用的免费新闻搜索引擎, 如百度新闻、Google 新闻都提供了关键词检索新闻的功能, 这两类新闻搜索引擎新闻来源均超过 1000 多家中文新闻媒体网站。被检索到的新闻类别通过计算机智能识别系统, 预先将新闻条目按照设定的关键词进行包括地区、事件类别等的分类, 待管理员审核后发布于网络平台, 供公众和专业人员浏览。正式发布于平台的新闻条也同样提供 RSS 订阅, 方便公众第一时间得到感兴趣的信息。

1.2 数据来源 基于大部分主流网站均设有 RSS 输出, 数据主要来源分为两类。第一类: 按照本系统选择的关键词表由免费的百度新闻搜索引擎搜索到的新闻条目。这部分新闻条目可以直接进入到系统进行分类。第二类: 部分未被百度新闻引擎所收录网站的新闻条目, 由系统直接接入, 经过系统过滤, 选择包含关键词的新闻条目进入到系统进行分类。

1.3 数据库 本系统共有 5 个数据库, 分别为 RSS 种子数据库、地名关键词数据库、事件类别关键词数据库、健康相关新闻类别关键词数据库、新闻数据库。RSS 种子数据库由百度新闻搜索引擎针对公共卫生事件紧密相关的 200 多个关键词制作生成。地名关键词数据库由国内市以上地名和国际上较大国家和城市名称构成。事件类别关键词数据库, 由事件类别和相对应的关键词组成。健康相关新闻类别关键词数据库由健康相关新闻类别和相应的关键词组成。新闻数据库由 RSS 种子收集的

所有新闻条组成。RSS 种子数据库用于收集新闻。地名关键词数据库、事件类别关键词数据库、健康相关新闻类别关键词数据库用于新闻条的匹配分类。新闻数据库用于查重和保存所有新闻条目相关信息, 包括标题、来源、时间、内容摘要、原文网址、发布类别、浏览量等。

1.4 公共卫生危机信息搜索发布流程 公共卫生危机信息搜索发布流程包括 3 个步骤。第一步, 筛选有关键词的新闻条目, 如已收录在百度新闻搜索引擎内的网站新闻通过百度引擎进行筛选, 而未收录到百度新闻搜索引擎的网站新闻则进入后由系统过滤筛选出有关键词的新闻。第二步, 智能分类, 在新闻标题、摘要、来源中利用地名关键词数据库、事件类别关键词数据库、健康相关新闻类别关键词数据库进行关键词匹配, 如匹配到关键词, 智能分配本新闻条目到关键词对应的地区和事件类别或健康相关新闻类别中。第三步, 管理员审核发布。自动分类的新闻, 由具有专业知识的管理员对新闻进行审核, 内容和分类均合适的按自动分类类别发布, 不合适的更换类别发布或放弃不发布。对于未匹配到关键词的, 待管理员审核时认为可发布后对本条新闻进行人工分类。见图 1。

## 2 结果

2.1 媒体监测结果发布形式 媒体监测数据借助于“公共卫生风险沟通平台”(网址: www.riskcommunication.org.cn) 进行公共卫生危机信息的分类发布。该平台运用 ASP.NET 程序语言设计, 由广州市疾病预防控制中心在 WHO 的资助下建立, 于 2008 年 10 月 7 日正式开通。虽然开通时间不长, 但发布效果基本达到要求。

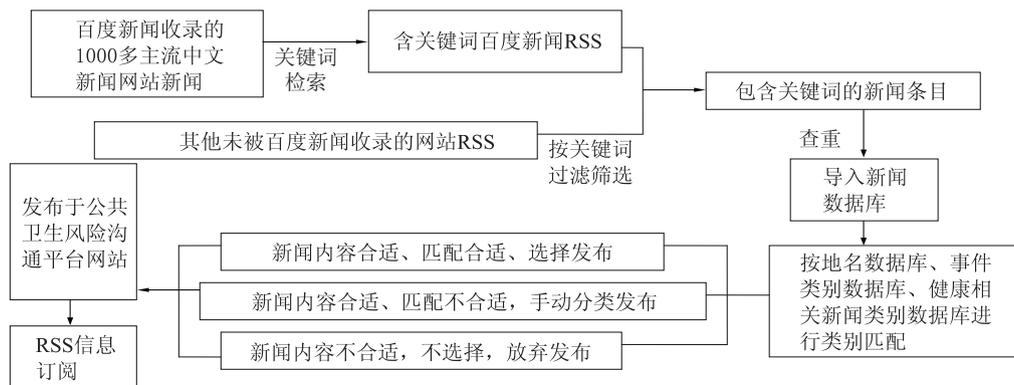


图 1 媒体监测系统公共卫生危机信息流程  
Figure 1 Flow-sheet of media surveillance for public health information

2.2 媒体监测系统可接受性 目前从互联网每日搜集包含关键词的新闻条目约 2000 条左右,经审核确认发布 100 条左右。其中大部分工作由计算机自动化信息处理,提高了获取信息的速度和质量,工作人员仅需点击审核发布或仅少量修改类别,可在短时间内完成针对全球 1000 多个主流中文新闻媒体的公共卫生危机信息监测,工作人员对系统接受性较好。

2.3 实现的信息分类及功能 为适合专业人员和公众等不同人群的信息需求,平台发布的信息共分为两大类:第一部分为公共卫生事件新闻,包括未达到突发公共卫生事件级别的公共卫生相关事件和危机信息都收入到这个类别,便于各地公众和专业人员早期预警采取行动,也可作为各地疾病预防控制部门、卫生行政部门、政府应急部门长期开展媒体监测预警的工具,发现苗头及时调查核实或处理,掌控公共卫生的媒体事件动向。这部分危机信息新闻按不同的地区(主要为各省份)和不同的事件类别进行分类,事件类别分类参照突发公共卫生事件信息报告系统分类并进行适当调整,见表 1。地区按区域共分为 8 个区域,无法区分区域的为其他,见表 2。第二部分为预警及健康相关新闻,主要是新闻媒体针对近期公共卫生情况发出的预警讯息及公共卫生相关领域研究进展和最新防治知识等信息,共分为七类公共卫生相关领域。

表 1 公共卫生危机信息媒体监测结果事件类别  
Table 1 Type of event by media surveillance for public health information

事件种类	事件类别
呼吸道传染病	传染性非典型肺炎、麻疹、肺结核、流行性脑脊髓膜炎、百日咳、白喉、猩红热、流行性感、风疹、流行性腮腺炎、水痘、其他呼吸道传染病
肠道传染病	霍乱、甲型肝炎、戊型肝炎、脊髓灰质炎、痢疾(细菌性和阿米巴性)、伤寒/副伤寒、手足口病、急性出血性结膜炎、其他感染性腹泻病、其他肠道传染病
自然疫源及虫媒传染病	流行性乙型脑炎、登革热、疟疾、黑热病、禽流感、鼠疫、出血热、狂犬病、炭疽、布鲁氏菌病、钩端螺旋体病、流行性和地方性斑疹伤寒、血吸虫病、丝虫病、包虫病、人感染猪链球菌、其他自然疫源及虫媒传染病
血源及性传播传染病	艾滋病、乙型肝炎、丙型肝炎、淋病、梅毒、其他血源及性传播传染病
其他传染病	麻风病、新生儿破伤风、其他传染病
中毒类	食物中毒与相关、职业中毒、其他中毒
其他	环境因素、不明原因或群体性不明原因、预防接种或服药事件、医源性感染和其他医院相关事件、意外辐射照射及核相关事件、其他公共卫生事件及疫情信息

表 2 公共卫生危机信息媒体监测结果地区分类  
Table 2 Region of event by media surveillance for public health information

区域	地 区
华北	北京市、天津市、河北省、山西省、内蒙古自治区
东北	辽宁省、吉林省、黑龙江省
华东	上海市、江苏省、浙江省、安徽省、福建省、江西省、山东省
中南	河南省、湖北省、湖南省、广东省、广西壮族自治区、海南省
西南	重庆市、四川省、贵州省、云南省、西藏自治区
西北	陕西省、甘肃省、青海省、宁夏回族自治区、新疆维吾尔自治区
港澳台	香港特别行政区、澳门特别行政区、台湾省
国外	
其他	注:某些讯息无法确定区域可列入该类

公众和专业技术人员可以通过本网站平台获取感兴趣的分类讯息,订阅方式灵活,可订阅不同地区或不同病种甚至于不同地区的不同病种分类的信息。如发生于某省的事件如广东的事件,关于某类别的事件如麻疹的事件,发生于某省的某类事件如广东的麻疹事件等,且均可提供免费 RSS 信息订阅推送服务。专业技术人员还可根据目前系统收集的媒体报道的疫情情况,提出当地相应的预警信息和干预措施,如发现近期出现流感的报道较多,提出流感等呼吸道传染病的预警,或者针对近期媒体报道煤气中毒较多,采取一系列干预措施等。

### 3 讨论

基于网络的媒体信息越来越丰富,其中包含着大量疾病监测和早期预警潜在而有用的信息,世界卫生组织(WHO)参与调查的几乎所有暴发信息都是首先从这些信息中获得的<sup>[4,5]</sup>。媒体信息对于突发公共卫生事件的调查处理来说是个双刃剑,报道失实或导向错误可以造成恐慌,增加疾病预防控制部门工作难度,及时正确引导媒体导向有助于妥善处理和平息事件。及时跟踪媒体动向对疾病预防控制部门、卫生行政部门以及政府来说都是相当重要的,以便于相关部门及时了解动态,加强应对。

目前各行业开展媒体监测的方式有很多种,从最原始的剪报方式,直到目前大量运用网络蜘蛛技术或网络机器人技术的搜索引擎,以及专业开发的各类情报竞争系统,均可达到行业内的媒体监测的功能,广泛用于针对商业竞争对手市场动向监测、品牌的危机事件报道和危机预警监测、自身产品市场反应和信誉监测以及政府部门开展的舆情监测等。本系统利用的是免费的搜索引擎,即网络蜘蛛技术开展相关关键词的新闻搜索,利用 RSS 网络技术将新闻条引入到系统中,并自动分类发布。计算机技术可使工作人员在

短时间从大量媒体中抓取到公共卫生危机信息。利用本系统,受众不需逐个浏览媒体或报纸就可了解媒体报道的大多数公共卫生疫情新闻。该系统能根据公众和专业技术人员的所处地区、关心热点,提供量身订做的信息,作为公共卫生信息聚合和媒体监测的工具,为预防疾病、早期发现事件、及时采取应急反应等均可以起到一定作用。

Freifeld 等<sup>[6]</sup>开发的 HealthMap 系统对全球的媒体报导的疾病暴发进行监测,运用的也是 RSS 新闻聚合器和 Google 新闻搜索引擎技术,并且 HealthMap 系统还对 ProMED Mail 进行收录监测。该系统采用的技术与本系统有着相同的基础,但又各有不同。该系统的来源为 1000 多个 Google 新闻 RSS 种子,主要收集非中文新闻。本系统采用的是知名的中文搜索引擎百度新闻的 RSS 种子,来源于 1000 多个中文新闻媒体网站,而且按照不同省份进行事件或信息的分类,搜索的国内事件或预警信息范围更全,地区定位更具体。

本系统作为中国公共卫生媒体监测的尝试还需要进一步改进,本系统仅使用 RSS 信息订阅作为接口,目前仅能接收大多数主流平面媒体的新闻信息,暂未包括电视、BBS 论坛、博客等公共卫生危机信息的来源。对于无 RSS 订阅功能网站的新闻也不能收录,尽管这部分的大型网站目前较少。另外,本系统不能完全实行自动化处理,审核发布仍需专业管理员进行人工操作,因此,公共卫生危机信

息收集发布时间与新闻网站发布时间存在一定的时间差。

## 参考文献

- [1] Jiang T, Li L, Jin BB. Discussion on surveillance system of internet media information in disease control and prevention system[J]. *Zhejiang Journal of Preventive Medicine*, 2008, 20(3):77-78. (in Chinese)  
江涛,李莉,金彬彬. 建立疾病预防控制机构网络媒体监测信息系统探讨[J]. *浙江预防医学*, 2008, 20(3):77-78.
- [2] Li GJ, Luo HM, Huang GH, et al. Analysis of media surveillance in areas of disease control[J]. *Chinese Journal of Disease Control & Prevention*, 2006, 10(1):54-57. (in Chinese)  
李桂娇,罗会明,黄国华,等. 疾病控制领域网络媒体监测及其实施结果分析[J]. *疾病控制杂志*, 2006, 10(1):54-57.
- [3] Feng ZC, Wang SJ, Peng XD, et al. Discussion on the function of news media in development of disease control field in Qingdao[J]. *Chinese Journal of Public Health Management*, 2002, 18(6):482-485. (in Chinese)  
逢增昌,汪韶洁,彭向东,等. 新闻媒体在青岛市疾病控制事业发展中的作用探讨[J]. *中国公共卫生管理*, 2002, 18(6):482-485.
- [4] Grein TW, Kamara KB, Rodier G, et al. Rumors of disease in the global village: outbreak verification[J]. *Emerg Infect Dis*, 2000, 6:97-102.
- [5] Heymann DL, Rodier GR. Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases[J]. *Lancet Infect Dis*, 2001, 1:345-353.
- [6] Freifeld CC, Mandl KD, Reis BY, et al. HealthMap: global infectious disease monitoring through automated classification and visualization of internet media reports[J]. *J Am Med Inform Assoc*, 2008, 15:150-157.

## ◇ 科普小知识 ◇

### 鼠 疫

鼠疫(plague)是由啮齿动物和蚤把鼠疫菌传给人和动物的一种特殊的人兽共患病,其开始的症状与体征是无特征性的发热、畏寒、不适、头及四肢疼痛、恶心、喉痛,常见的是最早被蚤叮咬的地方其相关的淋巴结出现淋巴腺炎,这就是腺鼠疫,其90%发生在腹股沟淋巴结,少数发生在腋下或颈部,受害的淋巴结肿胀、红肿、变软、化脓,发热是最常见的。

鼠疫的发病有明显的季节性,南方的鼠疫(主要是腺鼠疫)多发生在春夏季节,青藏高原等地区的鼠疫(主要是肺鼠疫)多发生于夏秋季节。

鼠疫的传播途径主要有三种:1. 经鼠蚤传播。鼠蚤叮咬是主要的传播途径,啮齿动物→蚤→人的传播是腺鼠疫的主要传播方式。主要的媒介是印鼠客蚤(*Xenopsylla*

*cheopis*)等10余种蚤类。2. 经皮肤传播。剥食患病啮齿动物的皮、肉或直接接触患者的脓血或痰,经皮肤伤口而感染。3. 经呼吸道飞沫传播。肺鼠疫患者是通过呼吸、谈话、咳嗽等,借飞沫形成“人-人”的方式传播,并可造成人间鼠疫的大流行。一般情况下,腺鼠疫并不造成对周围的威胁。

日常的预防措施主要是减少人被感染的蚤叮咬、直接接触被感染的组织或者暴露于肺鼠疫患者的可能性。通过使用杀虫剂和驱避剂避免蚤的叮咬。在农村避免处理老鼠。实行“三报三不”制度,鼠疫“三报”是指:报告病死鼠、报告疑似鼠疫患者、报告不明原因的高热患者和急死患者。鼠疫“三不”是指:不私自捕猎疫源动物、不剥食疫源动物、不私自携带疫源动物及其产品出疫区。