

融合 12320 卫生热线数据构建突发传染病事件早期症状监测预警平台*

蔡文锋 袁 俊

(广州市疾病预防控制中心 广州 510440)

〔摘要〕 通过分析和挖掘广州市 12320 卫生热线电话数据信息,结合原有传染病疫情监测体系,建立基于热线数据的突发公共卫生事件监测预警系统,介绍预警方法和流程。该系统能利用 12320 咨询信息进行数据分析处理,可作为公共卫生信息聚合和舆情监测的工具。

〔关键词〕 12320 热线;症状监测;聚集性探测;传染病突发事件对比与验证

〔中图分类号〕 R-056 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2016.02.004

Constructing a Monitoring and Early Warning Platform for Early Symptoms of Emergent Infectious Diseases by Integrating Data from the Health Hotline 12320 CAI Wen-feng, YUAN Jun, Guangzhou Center for Disease Control and Prevention, Guangzhou 510440, China

〔Abstract〕 By analyzing and excavating data from the health hotline 12320 in Guangzhou, in combination with the original monitoring system of infectious diseases, the paper constructs a monitoring and early warning system of emergent public health affairs based on data from the hotline and introduces the method and process of early warning. This system can utilize consulting information from 12320 to analyze and process data. Also, it can be used as the tool for aggregating public health information and monitoring public opinions.

〔Keywords〕 12320 hotline; Symptom monitoring; Spatio-temporal aggregation for discovery; Infectious diseases outbreak comparison and verification

1 引言

12320 卫生热线是原卫生部于 2005 年 12 月正式启用的政府公益热线^[1],主要目的是及时有效地应对公共卫生突发事件和重大公共卫生问题,

向公众传播健康防病知识,解答公众的各种健康咨询,引导公众的健康行为,帮助公众建立科学的健康防病意识^[2]。截至 2015 年 9 月,全国已有 28 个省市开通了 12320,历经三鹿奶粉^[2]、北京奥运会^[3]、甲型 H1N1 流感^[4]等事件,以其信息的权威性和良好的沟通赢得了公众的信赖和认可,在及时掌握公众舆情、引导公众认知、疏导公众情绪、妥善化解危机和维护稳定等方面发挥着不可替代的作用^[4],成为一张“舆情直报网”和应对处理突发公共卫生事件最直接、最迅速的通道^[3]。广州于 2012 年 10 月底开通了 12320,累计

〔修回日期〕 2015-12-29

〔作者简介〕 蔡文锋,硕士,副主任医师,发表论文 5 篇;通讯作者:袁俊。

〔基金项目〕 广州市科技计划项目(项目编号:2014Y2-00063)。

受理各类咨询 6 万多件次。为了进一步丰富突发公共卫生监测信息的来源,提高 12320 对突发公共卫生事件监测的预警作用,充分挖掘咨询电话信息中所隐藏的突发公共卫生事件苗头和危机信息,及早开展干预及应对,笔者结合广州市目前公共卫生领域传染病监测与预警的工作进展,初步构建了融合 12320 卫生热线数据的突发传染病事件早期症状监测预警平台。

2 方法

2.1 数据来源与预处理

2.1.1 数据来源 广州市 12320 卫生热线接听的咨询电话,座席员在接听每一个咨询电话后均按照不同传染病或咨询症状等进行分类,同时还将咨询电话具体内容输入话务系统。

2.1.2 数据筛选 一种是经座席员分类的咨询电话,按咨询内容属性已归类为不同的传染病等相关的分类项目。另一种是涉及发热、咳嗽、腹泻、呕吐、皮疹、死亡等症状,则座席员将该条咨询电话按症状关键词分类录入系统,可用于开展按传染病症状分类的症状监测进行分析。

2.1.3 数据整理 按两种方法筛选出的固定电话,与中国电信电话号码数据库提供的 3 级地址进行匹配,获取 3 级地址的地理定位信息作为该咨询电话的大致地理定位,因此该咨询电话在分析数据库中拥有 3 个分类变量属性:时间信息、地理信息及病名(或症状)信息,其中地理信息与大疫情系统中传染病实际发生的情况进行对照分析,以便于进一步开展聚集性统计分析。

2.2 预警方法

2.2.1 时间序列分析 主要是根据 12320 电话咨询中涉及突发公共卫生事件报告历史数据库,采用移动百分位数法动态计算咨询中相对应的历史基线,建立将当地当前观察周期(7 天)内咨询数与其相应历史基线实时进行比较的预警模型。当观察周期内发现的病例数达到预警阈值时,系统将在 24 小时内自动发出预警信号,见图 1。

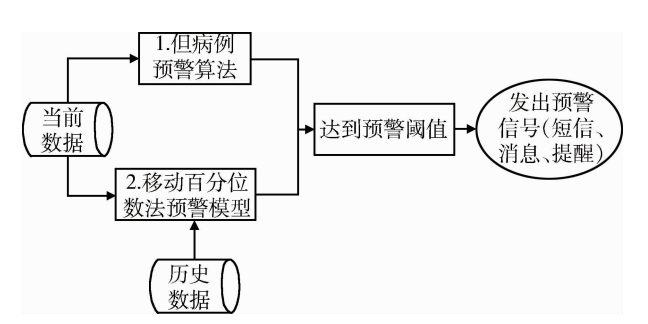


图 1 时间序列分析主要流程

2.2.2 空间扫描统计量 运用一系列扫描圆在研究区域监测出疾病空间聚集性。该方法在开始进行监测时,随机选取研究区域内某一病例点或小范围中心点(例如乡镇点),以其为圆心生成一系列扫描圆。这些扫描圆的半径由 0 到规定的上限逐步变化。当扫描圆半径达到规定的上限后,便又以区域内另外一个病例点为圆心,开始新一轮的圆形扫描。直到遍历完所有的病例点后结束,生成了无数个不同位置、大小不一的扫描圆。在扫描过程中,基于备择假设 H_1 : 至少存在一个扫描圆,其区域内发病率明显高于区域外,针对每个扫描圆,利用圆内外病例实际值和期望值计算一个似然比值。研究中根据病例数据概率分布情况,使用 Poisson 似然比值计算公式:

$$\lambda = \max_z \frac{\left(\frac{n_z}{\mu(z)}\right)^{n_z} \left(\frac{n_G - n_z}{\mu(G) - \mu(z)}\right)^{n_G - n_z}}{\left(\frac{n_G}{\mu(G)}\right)^{n_G}} \times I\left(\frac{n_z}{\mu(z)} > \frac{(n_G - n_z)}{(\mu(G) - \mu(z))}\right)$$

其中 λ 是似然比值, n_z 和 $\mu(z) = \frac{n_G}{m_G} * m_z$ 分别代表扫描圆 z 里的 7 天内实际病例数和根据零假设得到的病例数期望值,而 m_z 表示扫描圆 z 中人口数;根据下标类推, $m_G, n_G, \mu(G)$ 依次为县区 G 的人口数、发病人数和病例期望值, $n_G = \mu(G)$ 。 $I(\cdot)$ 是一个指示函数,当 $\frac{n_z}{\mu(z)} > \frac{(n_G - n_z)}{(\mu(G) - \mu(z))}$, 即搜索圆的发病率高于圆外该县区其他区域的发病率时等于 1, 反之为 0, 不运算。在扫描过程结束后,将所有扫描圆的似然比由大到小排序,选择排在前面的若干个作为疾病聚类备选区域进入 Monte Carlo 检验。通过检验的扫描圆便是最后监测到的疾病聚集高发区域。

2.3 预警流程

采用 Satscan 和地理信息系统实现整个预警过程：
(1) 通过咨询电话传染病归类和症状关键词的搜索，筛选出与突发公共卫生事件相关的服务请求数据，结

合固定电话装机地址建立空间数据库。(2) 调用咨询的地理数据，采用移动百分位数法和时空扫描统计量进行聚集性监测，对监测结果进行编辑，将结果返回数据库。(3) 利用地理信息系统软件对监测结果进行可视化展示，见图 2、图 3。

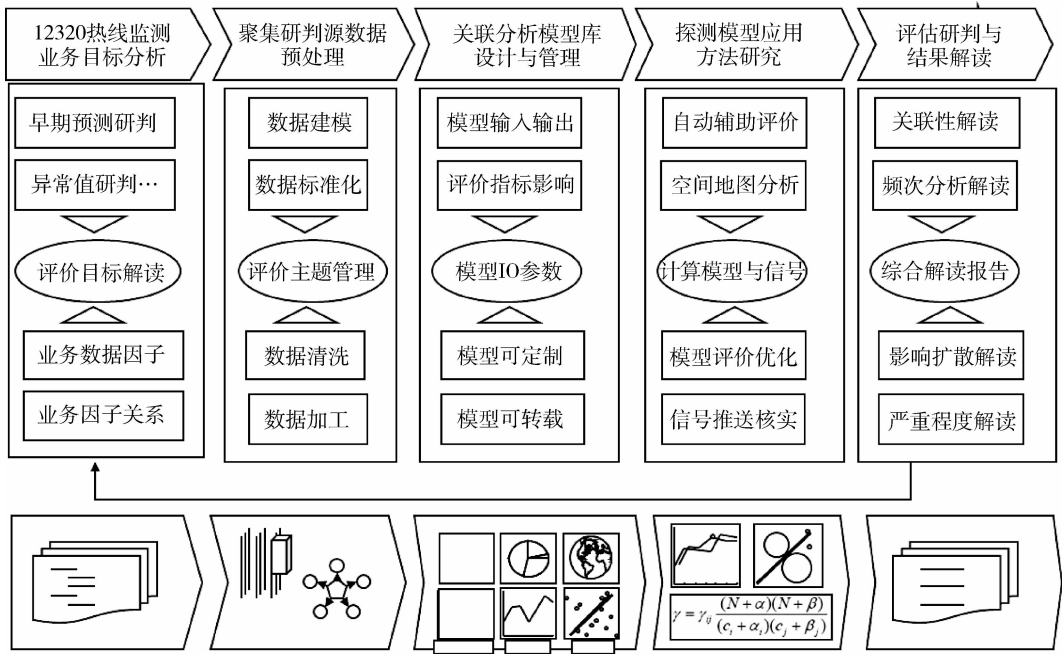


图 2 12320 卫生热点预警监测模型研究思路

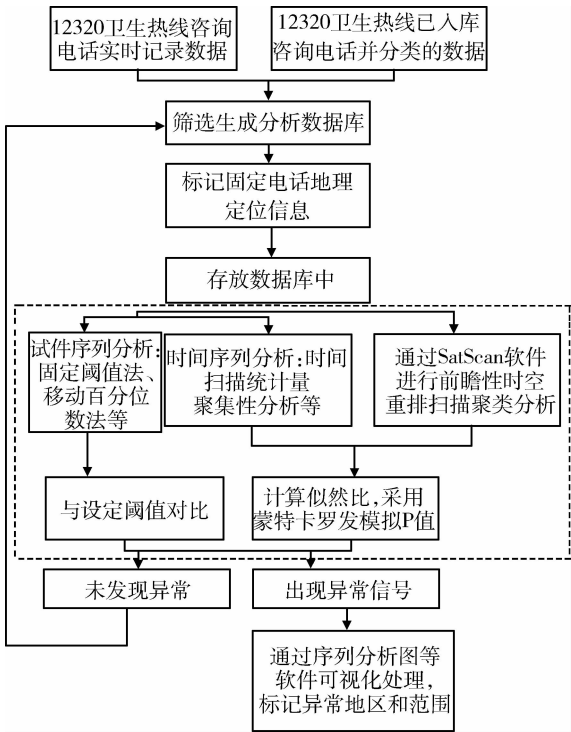


图 3 12320 预警监测系统信息处理流程

3 结果

3.1 实现信息分类及建立数据库信息

对所有的咨询电话进行数据结构分类统计，共分 5 大级别、7 个大类、180 个小类别，涉及突发公共卫生事件的事件报告、病种分类、临床表现及治疗等内容；通过中国电信的数据链接，建立 600 万条广州市固定电话匹配的地理信息数据库。

3.2 监测结果发布形式

在 12320 综合信息平台进行数据的实时更新和展示。该平台运用 JavaEE 架构，融合 R 语言、SatScan、ARCGIS 等分析工具实现了 12320 关联分析模型，用于疾病的早期聚集性监测，将发现的时间空间传染病或相关症状聚集性的地区以不同的符号或颜色标示出来，便于直观掌握当前监测预警情况。

3.3 监测系统的可接受性

目前广州 12320 每日咨询量约在 200 件次左右,经系统自动分类纳入统计约占 1/5,其中大部分工作由计算机自动化处理,提高获取信息的速度和质量。前期可先实现每日下载数据,随后手工完成数据的时空预警分析软件 SatScan 以及地理信息系统软件数据处理,再整合到 12320 卫生热线管理系统中,实施实时自动监测预警功能。

3.4 报告发布

2014-2015 年 11 月广州市 12320 共接到疑似突发公共卫生事件报告信息 406 件次,同期全市报告突发公共卫生事件 152 起。

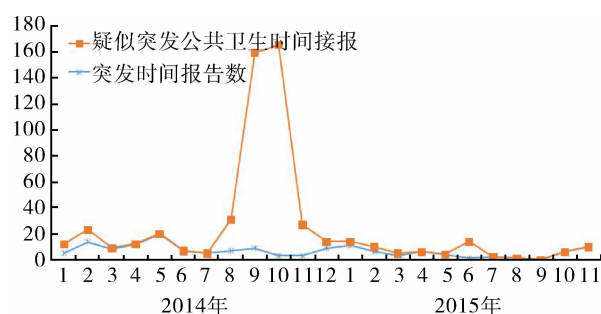


图 4 2014-2015 年广州 12320 疑似突发事件接报时间分布

4 讨论

4.1 国内外突发公共卫生事件早期预警的典型案例

近年来,对突发公共卫生事件的早期预警研究渐渐从利用医院等传统监测资料开展预警,扩展到互联网关键词搜索引擎、社交网站、电话热线咨询、药店销售、缺勤记录等非传统监测资料。国外学者曾利用网友在谷歌和雅虎搜索引擎上使用的检索关键词准确预测流感暴发^[6-7],而且谷歌搜索引擎是在政府公布信息之前 14 天准确预测到流感暴发。利用谷歌搜索引擎同样可以预测流感样疾病、胃肠炎及水痘等其他疾病^[8]。美国也有研究人员借助微博预测个体感染流感时间,准确率达 90%^[9]。英国学者通过分析国家医疗服务体系(National Health Service, NHS)热线中咨询“呕

吐”的数据,提前预测到院内诺瓦克病毒感染暴发疫情^[10],美国、加拿大学者也开展了公共卫生热线在突发公共卫生事件早期预警的应用研究^[11-12]。马晓薇等通过微博与传统网络媒体分析一起人禽流感疫情舆情监测特点^[13]。因此,基于公众电话热线咨询数据分析对于开展突发公共卫生事件的监测预警具有非常实用的价值,可以及时发现传染病疫情和暴发的苗头,将传染病疫情预警信息和舆情监测信息及早转给相关部门开展调查核实、应急处置和应对工作,将突发公共卫生事件对人民群众的影响降低到最小。

4.2 12320 热线突发公共卫生事件检测效果

12320 热线在接受公众信息咨询的同时成为获取疫情信息举报的一个途径,如南京市卫生局通过村民拨打 12320 第一时间得知一家 8 口不明原因发热(1 人死亡)的案例,及时开展了调查处置^[5];广州 12320 近两年来接听过包括登革热爆发、不明原因皮疹等多起疑似突发公共卫生事件报告。12320 在传染病疫情监测中虽起到一定的效果,但目前仅限于接受公众的疫情举报和时间分布曲线等浅层次的监测和汇总分析,及时性和敏感性都有待提高。此外,12320 接受公众咨询的信息量非常巨大,如 2015 年 9 月北京市 12320 热线受理咨询 40 184 件次,上海市受理 20 941 件次。如此庞大的数据量,蕴藏着大量传染病暴发或异常的苗头,但我国尚未建立一套监测预警系统,无法充分利用这部分信息进行数据挖掘。在本次研究设计中,利用电话热线咨询中含有的突发公共卫生事件异常或暴发信息的特点,叠加旧有的疫情监测网络数据,采用时空探测与匹配确认相结合的预警技术,开展突发公共卫生事件监测和症状监测,以便早期及时发出预警,同时也可以用于开展舆情监测。广州 12320 开通时间较晚,知晓率和使用率不高,目前的咨询量远比不上北京、上海、南京等地;部分数据信息不全,其中纳入数据统计分析的不到咨询总数的 1/10,目前已完成平台的建设工作,尚未对数据进行拟合分析,提示应继续加大宣传力度,在提供服务的同时做好大数据分析的先期准备工作;此

外,现阶段监测平台只能采集到固定电话地理信息,尚无法通过手机基站获取移动电话地理信息,有待以后进一步研究和完善。

5 结语

本系统通过分析公众卫生热线咨询数量的变化探测异常,是广州市突发公共卫生事件监测预警研究工作的有效补充和提升,同时有助于更快更全面的发现异常,利于早期开展干预和处置。

参考文献

- 1 胡明霞,赵科伙,吴思远. 12320 电话信息分析及其在突发公共卫生事件监测中应用 [J]. 中华疾病控制杂志, 2011, 15 (10): 8896 - 8980.
- 2 蒋燕,王超,侯培森. “三鹿牌婴幼儿奶粉事件”中 12320 公共卫生公益电话的作用及其启示 [J]. 中国健康教育, 2009, (4): 312 - 313.
- 3 冀永才,谭枫. 12320 热线奥运期间咨询业务数据分析 [J]. 中国健康教育, 2010, 2 (2): 158 - 159.
- 4 王超,谭枫. 12320 在甲型 H1N1 流感中风险沟通与舆情监测作用 [J]. 中国公共卫生, 2010, 11 (2): 1471 - 1472.
- 5 蒋燕,侯培森. 中国 12320 公共卫生公益电话试点建设及效果 [J]. 中国公共卫生, 2009, 12 (2): 1528 - 1529.
- 6 Iinsberg J, Mohebbi MH, Patel RS, et al. Detecting Influenza Epidemics Using Search Engine Query Data [J]. Nature, 2009, (457): 1012 - 1014.
- 7 Polgreen PM, Chen YM, Pennock DM, et al. Using Internet Searches for Influenza Surveillance [J]. Clin Infect Dis, 2008, (47): 1443 - 1448.
- 8 Pelat C, Turbelin C, Bar - Hen A, et al. More Diseases Tracked by Using Google Trends [J]. Emerg Infect Dis, 2009, (15): 1327 - 1328.
- 9 新华网. 美研究人员微博预测流感准确率高达 90% [EB/OL]. [2015 - 11 - 8]. http://news.xinhuanet.com/tech/2012-07/28/c_123483027.htm.
- 10 Loveridge P, Cooper D, Elliot AJ, et al. Vomiting Calls to NHS Direct Provide an Early Warning of Norovirus Outbreaks in Hospitals [J]. J Hosp Infect, 2010, 74 (4): 385 - 393.
- 11 Jane SR, Floyd F, Walter J. Using Nurse Hot Line Calls for Disease Surveillance [J]. Emerg Infect Dis, 1998, 4 (2): 329 - 332.
- 12 Rolland E, Moore KM, Robinson VA, et al. Using Ontario's "Telehealth" Health Telephone Helpline as an Early - warning System: a study protocol [J]. BMC Health Serv Res, 2006, 15 (6): 10.
- 13 马晓薇,刘艳慧,袁俊,等. 微博与传统网络媒体分析一起人禽流感疫情舆情监测特定分析 [J]. 疾病监测, 2013, (28): 61 - 63.

2016 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的医学信息学方面的国家级期刊。主管：国家卫生和计划生育委员会；主办：中国医学科学院；承办：中国医学科学院医学信息研究所。中国科技核心期刊（中国科技论文统计源期刊），RCCSE 中国核心学术期刊（武汉大学中国科学评价研究中心，Research Center for Chinese Science Evaluation），美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息技术，医学信息研究，医学信息组织与利用，医学信息教育，动态等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业的师生及广大医教研人员。

2016 年《医学信息学杂志》国内外公开发行，每册定价：15 元（月刊），全年 180 元。邮发代号：2 - 664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号（100020）医科院信息所《医学信息学杂志》编辑部；电话：010 - 52328673, 52328674, 52328671。

《医学信息学杂志》编辑部