

基于互联网信息的公共卫生监测进展

赵永谦 马文军

510326 广州,暨南大学医学院(赵永谦); 511430 广州,广东省疾病预防控制中心,广东省公共卫生研究院(马文军)

通信作者:马文军, Email: mwj68@vip.tom.com

DOI: 10.3760/cma.j.issn.0254-6450.2017.02.027

【摘要】 互联网数据因其更新快和数据量丰富的特点被引入公共卫生监测领域。通过挖掘分析互联网中丰富的健康相关信息,研究者建立基于互联网的公共卫生监测系统,以快速估计人群中健康相关事件的分布。基于互联网的公共卫生监测系统根据信息来源不同可分为主动式监测系统和被动式监测系统。相较于传统监测系统,基于互联网的监测系统具有及时性佳和成本低的优点,还可在健康相关行为监测和卫生政策评价等传统监测力所不逮的领域发挥作用。但由于假阳性消息的存在,基于互联网的监测系统存在准确性不佳的缺点。本文介绍基于互联网的公共卫生监测进展及其在疾病预警中的应用,为我国开展相关工作提供参考。

【关键词】 公共卫生监测; 互联网

基金项目: 广东省省级科技计划(2014A040401041)

A review on the advancement of internet-based public health surveillance program Zhao Yongqian, Ma Wenjun

School of Medicine, Jinan University, Guangzhou 510326, China (Zhao YQ); Guangdong Provincial Center for Disease Control and Prevention, Guangdong Provincial Institute of Public Health, Guangzhou 511430, China (Ma WJ)

Corresponding author: Ma Wenjun, Email: mwj68@vip.tom.com

【Abstract】 Internet data is introduced into public health arena under the features of fast updating and tremendous volume. Mining and analyzing internet data, researchers can model the internet-based surveillance system to assess the distribution of health-related events. There are two main types of internet-based surveillance systems, i.e. active and passive, which are distinguished by the sources of information. Through passive surveillance system, information is collected from search engine and social media while the active system gathers information through provision of the volunteers. Except for serving as a real-time and convenient complementary approach to traditional disease, food safety and adverse drug reaction surveillance program, Internet-based surveillance system can also play a role in health-related behavior surveillance and policy evaluation. Although several techniques have been applied to filter information, the accuracy of internet-based surveillance system is still bothered by the false positive information. In this article, we have summarized the development and application of internet-based surveillance system in public health to provide reference for a better surveillance program in China.

【Key words】 Public health surveillance; Internet

Fund program: Science and Technology Planning Project of Guangdong Province (2014A040401041)

公共卫生监测是指长期、连续、系统地收集卫生事件相关数据,经过科学分析获得信息,并及时反馈给需要这些信息的人和机构,用以指导制定、完善和评价公共卫生干预措施与策略的过程^[1]。传统的公共卫生监测信息由基层医务人员提供,具有良好的准确性和完整性,提供的人口学和影响因素资料还可用于危险因素分析,但这样的监测系统存在局限性:①病例确认、信息传递和发布耗时长,影响及时性;②信息主要来源于医疗机构,未就医病例遗漏影响监测系统的敏感性;③对于新发疾病,尤其是新发传染病,由于缺乏可

靠的历史数据或无完善的监测网络,传统监测系统作用有限。近年来,为快速有效地获取健康相关事件信息,互联网被引入了公共卫生监测。本文通过文献综述,介绍基于互联网信息的公共卫生监测进展及其应用,为我国开展相关工作提供参考。

一、互联网公共卫生监测

随着互联网蓬勃发展,网民的互联网使用痕迹中蕴含丰富的健康相关信息,这为公共卫生监测带来了新机遇。互联网公共卫生监测区别于传统公共卫生监测之处在于以互联

网信息为信息来源,其通过长期、连续和系统的从互联网中收集分析疾病和健康事件相关信息,以早期发现和识别疾病和健康事件在人群中的分布,为制定公共卫生措施和政策提供依据^[2-3]。互联网公共卫生监测长期、连续和系统的方式收集疾病和健康相关信息,这种特点使其有别于短期的网络调查和收集公众情绪、态度和意见信息的舆情监测。

互联网公共卫生监测在信息采集、数据分析和应用领域均有发展。互联网公共卫生监测根据信息采集方式分为被动式与主动式;互联网被动监测主要基于搜索引擎和社交媒体连续收集数据^[4-12],信息来源于网民使用搜索引擎的痕迹和在社交媒体上发布的消息^[13-19];互联网主动监测为参与式监测,数据由志愿者周期性的通过电子邮件、网页或手机应用提供^[3, 20-25]。在数据分析方法上,被动式监测系统则以疾病或健康事件相关词为指示词,利用其数据量构建模型对总体人群做出估计;互联网主动监测则以志愿者中疾病和健康事件发生频次或频率估计总体人群情况。互联网监测的应用领域也在不断拓宽,以流感、登革热、肺结核、痢疾和埃博拉等传染性疾病预防为起点^[2-3, 26-27],逐步向食源性、糖尿病和肿瘤等非传染性疾病拓展^[28-32],现今在药物安全监测、人群不良行为监测和政策效果评价等领域亦有应用^[33-45]。尤其在传染病监测领域,已有监测系统成功应用^[46-50]。

二、互联网公共卫生被动监测

按照信息来源,互联网公共卫生被动监测分为基于搜索引擎的监测、基于社交媒体的监测和其他信息来源的监测。

1. 基于搜索引擎的被动监测:搜索引擎是网民获取健康信息的重要方式,研究者最先利用记录搜索行为的日志文件数据进行分析,以疾病和健康事件相关词为指示器,利用其搜索量数据建立模型,对人群中疾病和健康事件的分布做出估计。针对登革热和流感,研究者利用运营商提供的日志文件建立线性回归模型,估计其每日的活动情况^[4-5]。谷歌推出的可提供一定时间和空间内某词搜索热度(反映搜索量的间接指标)的谷歌趋势(Google Trend, GT)和谷歌搜索透视(Google Insight for Search)服务后,为研究者提供了新的数据来源^[6-7]。2008年谷歌推出谷歌流感趋势(Google Flu Trend, GFT)服务,以45个流感相关词为指示词构建线性模型估计流感活动,其后两度调整模型(将相关词扩充至160个、弱化媒体关注影响及使用弹性网络模型)以提高监测准确性^[8-9]。2009年后GFT服务范围覆盖了美国各州、主要城市以及20个国家,共38种语言^[10]。2011年,谷歌建立主要服务于南美和东南亚发展中国家的谷歌登革热趋势(Google Dengue Trends, GDT)。即使多次调整,GFT估计值远高于传统监测数据的问题未能解决,因此在2015年8月9日后,GFT与GDT均不再更新^[11]。不限于登革热和流感,研究表明搜索引擎数据可用于疟疾、梅毒和艾滋病等传染性疾病监测^[26-27],亦可应用于非传染性疾病、健康相关行为、药物安全和卫生政策效果评价等方面;通过分析相关词搜索量与疾病的关系,研究者发现搜索引擎数据可用于肿瘤、食源性疾病和精神卫生监测^[28-31];在药物安全领域,利用互联网搜索数

据可估计人群药物使用情况^[34]和药物不良反应^[35];在卫生政策或卫生干预措施实施后,可利用搜索引擎中相关词条搜索量的变化估计政策或干预措施效果,尤其在烟草控制方面,多国研究者利用搜索引擎数据估计提税政策、世界无烟日和禁烟令等政策对于吸烟行为的影响^[36-39]。

2. 基于社交媒体的被动监测:电子邮件、论坛、博客和即时通讯软件等社交媒体的兴起不仅为健康信息的传播提供有利条件,还为公共卫生监测提供了新的工具。同以疾病和健康事件相关词为指示词,研究者利用自然语言处理技术从社交媒体文本消息中获得指示词数据量,建立回归模型对人群中疾病和健康事件分布做出估计。传染病监测仍是探索最多的领域,尤其是流感监测^[13-16],研究者也将这项技术推广应用于登革热^[17]、霍乱^[18]和埃博拉^[19]等传染病的监测,并且致力于开发基于社交媒体的多病种监测系统。研究表明社交媒体同样可用于HIV相关高危行为、人群肥胖^[40-41]、吸烟等不良行为监测^[43]及药物不良反应监测^[42]等领域。即使在报告率低的食品卫生监测领域,社交媒体亦可帮助卫生部门发现本地食品卫生问题^[32-33]。

3. 其他信息来源的被动监测:互联网被动监测还可以新闻报告、官方通告、专家意见和健康网站作为信息来源。新发传染病监测系统(Program for Monitoring Emerging Diseases, ProMED-Mail)、全球公共卫生情报网络(The Global Public Health Intelligence Network, GPHIN)、医学智能系统(Medical Intelligence System, MedISys)和全球疾病预警地图(Healthmap)等基于互联网的监测工具主要信息来源于新闻媒体和公共卫生网站的通告,经人工或计算机技术分析判断是否发生疾病的暴发流行^[49]。这些系统运营数年并曾准确预报多次传染病疫情,在慢性非传染性疾病、食品卫生等其他公共卫生监测领域亦有运用。

三、互联网公共卫生主动监测

由于互联网被动监测信息完全依赖于网络使用者,具有许多不确定性。互联网主动监测(也称参与式监测)则是志愿者通过互联网主动提供健康相关信息。因为信息由志愿者主动提供,其可靠性、完整性和针对性均较被动式互联网监测系统有所提升。参与式监测在志愿者注册时进行背景调查,并对志愿者是否出现某些症状或行为进行连续调查,从而对一般人群中疾病或健康事件分布情况做出估计。目前,参与式监测系统主要用于传染性疾病监测,也用于药物副作用监测。

2003年9月,荷兰和比利时最先开发大流感调查(Great Influenza Survey, GIS)参与式监测系统用于流感监测,随后GIS系统在葡萄牙和意大利上线运行^[20]。2009年,英国流感调查(Flusurvey)系统上线,并加入欧洲多国联合构建的参与式监测网络——欧洲流感监测网(Influenzanet),至2014年欧洲流感监测网已覆盖10个欧洲国家或地区,拥有超过5万名注册志愿者^[2-3, 24]。2006年,澳大利亚流感追踪(Flutracking)开始在澳大利亚部分地区试运行,于次年推广至全境^[50]。2012年,身边的流感(Flu Near You, FNY)在美国和加拿大上

线,其友好的界面与功能帮助志愿者快速完成较高质量的调查^[21-22,25]。在执行流感监测任务的同时,英国流感调查和澳大利亚流感追踪还可提供疫苗覆盖率和保护效力等信息^[46,51]。参与式监测系统并不局限于流感,用于多种呼吸道疾病监测的墨西哥呼吸道疾病监测系统(Reporta)、用于登革热的监测的巴西登革热监测网络(Dengue na Web)和多种疾病监测的健康波多黎各(SaludBoricua)也相继上线运行^[3]。在药物安全监测领域,亦有在线社区开发了用于糖尿病相关设备引起的副作用监测的参与式监测系统^[44-45]。

四、互联网公共卫生监测的准确性

互联网来源数据是否可用于疾病监测的关键在于其是否符合实际情况,庞大的数据量和匮乏的背景信息使核实诊断十分困难,因此研究者通过分析估计值与传统监测数据相关性判断准确性。搜索引擎数据与传统监测数据的良好相关性在许多研究中均得到验证^[6-9,52-55],但背景信息匮乏使研究者难以利用技术手段剔除假阳性信息,因此提高基于搜索引擎系统的准确性主要依赖于指示词与实际健康相关事件的相关程度。韩国学者发现,随用户搜索行为的改变,关键词与健康相关事件的相关程度也会发生改变:同一个流感季内,GT 中不同流感关键词与传统监测数据的相关性不同;不同的流感季,同一流感关键词与传统监测数据的相关性亦会发生改变^[7]。此外,新闻报道和公众恐慌情绪均会促使大量假阳性信号产生,但基于搜索引擎的监测无法剔除此类信息,从而导致准确性降低^[5,7,9,28]。2012—2013 年甲型 H3N2 流感流行期间,由于媒体的过度报道,GFT 出现估计值远高于传统监测数据的情况^[8],即使在弱化媒体关注后,GFT 仍存在估计值过高的情况^[12]。不同于搜索引擎数据,社交媒体数据中蕴含了语义和情感信息,这些信息使基于社交媒体的监测可利用人工筛选^[13]、机器学习^[16]和情感分析^[17]等多种分析技术手段剔除部分假阳性信息。在针对推特(Twitter)的研究中,研究者在剔除包含统一资源定位符(Uniform Resource Locator, URL)的推文(Tweet)和转发推文后,与流感样病例监测数据及实验室确诊流感数据的相关性均升高^[16]。参与式监测直接从志愿者获取信息从根本上避免了部分假阳性信息,因此其准确性主要受志愿者数量和质量影响^[20,22,24]。由于缺少志愿者,多个参与式监测系统在运营初期表现不佳^[23];剔除了短期参与的志愿者后,参与式监测系统的准确性有所提高^[20,24,48]。

五、总结与展望

互联网发展促使以互联网信息为基石的互联网公共卫生监测不断革新信息采集和数据分析方法。在信息采集方法上,互联网主动监测不断改善以提高志愿者依从性,从最初周期性通过电子邮件发送问卷收集志愿者信息^[20],发展为仅完成网页中选项即可完成调查^[3,23-24,50],FNY 的用户仅通过手机应用即能完成调查^[21-22,25]。互联网被动监测系统则寻找最优信息源,最初数据来源于运营商提供的完整数据文件^[4-5];但由于可获得性差的原因,监测者转投向间接搜索量的 GT 和谷歌搜索透视搜索热度^[6-7];社交媒体信息中的

背景信息可用于噪音剔除,因此监测者将其作为新的数据来源,通过应用程序编辑接口(Application programming interface, API)或爬虫获取文本信息,再利用自然语言处理技术筛选后用于分析^[15-17]。在数据分析方法上,互联网被动监测不断尝试建立拟合度更高的模型以提高估计值的准确性,指示词的数量从单个增加到多个^[4,6,16-17],从简单线性模型转变为多元线性模型、负二项回归模型和弹性网络模型^[4,6,8,15]。在应用领域,亦从传染病监测拓展到慢性病、食源性疾病和人群不良行为监测^[26-45]。

相对于传统监测,互联网监测系统最大的优点是低成本和时效性好。互联网监测通过互联网获取数据,节省了人力、物力和财力。传统公共卫生监测系统,从发现病例到最终发布监测报告耗时长达数天或数周,而被动式监测可在次日发布结果。GFT 可在当日提供估计值,基于推特的监测系统的报告亦可先于传统监测数据约 2 周发布^[55]。除假阳性信息影响监测的准确性,互联网监测系统还有其他局限性:①匮乏的位置信息不利于开展空间上监测活动,基于搜索引擎的监测系统仅可通过 IP 地址推测位置信息,诸如推特等社交媒体虽然提供位置标签服务,但用户不添加地址信息或自定义地址的行为均不利用监测系统获取准确的位置信息^[4,19,56];②被动监测难以获取完整的原始数据,仅能获取运营商处理后数据(如 GT 提供的搜索热度)或部分原始数据(如利用 API 或爬虫获取 twitter 数据)^[8,15];③被动监测无法获得用户的知情同意,需要通过诸如地址粗略化和剔除特征化消息等方法保护用户隐私^[57];④互联网用户不能完全代表一般人群,在性别、年龄、居住地、文化程度和自我保健意识上均存在差异:社交媒体和搜索引擎用户以女性、年轻人和城镇居民居多,且文化程度较高;志愿者还具有自我保健意识较高的特点^[18,47-48,58]。

互联网数据与传统监测数据的高相关性说明其存在用于互联网监测的可能性,但 GFT 的失败也提示互联网数据用于公共卫生监测还需要扫清诸多障碍。其中最大的障碍是假阳性信息,研究者需要继续寻找剔除假阳性信息的方法:被动监测必须不断更新关键词及调整模型,以提高自身与人群中健康相关事件的相关程度^[7,9,13];基于社交媒体监测需要更加快速准确的数据获取和文本分析技术^[15,17],以快速提取和筛选社交媒体信息;主动式监测系统要求研究者采取多种方法扩展志愿者队伍并保证其依从性^[20,23-24,48]。因此,互联网公共卫生监测多作为辅助性工具提供参考性意见,尤其是对于新发疾病或传统监测力所不逮的领域,互联网公共卫生监测能及时提供信息^[19]。

目前,互联网公共卫生监测系统的信息来源主要为搜索引擎、内容分享型社交媒体和志愿者,但诸如美国社交媒体网站脸书(Facebook)留言板等即时通讯软件附带内容分享模块鲜有涉及,这是今后发展的一个重要方向^[59]。即时通讯软件在时效性上有极强的优越性,若能寻找到合适的信息获取技术即能在公共卫生监测领域发挥作用。社交媒体还提供了用户友好的界面与丰富的互动功能,参与式监测系统若

与社交媒体的结合,可提高资料收集的效率和志愿者的依从性。在克服假阳性信息障碍后,互联网监测系统将在公共卫生监测领域发挥重要作用,为疾病的早期预警和干预提供服务。

利益冲突 无

参 考 文 献

- [1] 詹思延. 流行病学[M]. 北京:人民卫生出版社,2012:189-201. Zhan SY. Epidemiology[M]. Beijing: People's Medical Publishing House,2012:189-201.
- [2] Milinovich GJ, Williams GM, Clements ACA, et al. Internet-based surveillance systems for monitoring emerging infectious diseases[J]. Lancet Infect Dis, 2014, 14(2): 160-168. DOI: 10.1016/s1473-3099(13)70244-5.
- [3] Wójcik OP, Brownstein JS, Chunara R, et al. Public health for the people: participatory infectious disease surveillance in the digital age[J]. Emerg Themes Epidemiol, 2014, 11(1): 7. DOI: 10.1186/1742-7622-11-7.
- [4] Polgreen PM, Chen YL, Pennock DM, et al. Using internet searches for influenza surveillance[J]. Clin Infect Dis, 2008, 47(11): 1443-1448. DOI: 10.1086/593098.
- [5] Chan EH, Sahai V, Conrad C, et al. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance[J]. PLoS Negl Trop Dis, 2011, 5(5): e1206. DOI: 10.1371/journal.pntd.0001206.
- [6] Althouse BM, Ng YY, Cummings DAT. Prediction of dengue incidence using search query surveillance[J]. PLoS Negl Trop Dis, 2011, 5(8): e1258. DOI: 10.1371/journal.pntd.0001258.
- [7] Cho S, Sohn CH, Jo MW, et al. Correlation between national influenza surveillance data and google trends in South Korea[J]. PLoS One, 2013, 8(12): e81422. DOI: 10.1371/journal.pone.0081422.
- [8] Olson DR, Konty KJ, Paladini M, et al. Reassessing google flu trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales[J]. PLoS Comput Biol, 2013, 9(10): e1003256. DOI: 10.1371/journal.pcbi.1003256.
- [9] Cook S, Conrad C, Fowlkes AL, et al. Assessing google flu trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic[J]. PLoS One, 2011, 6(8): e23610. DOI: 10.1371/journal.pone.0023610.
- [10] Eurosurveillance Editorial Team. Google Flu Trends includes 14 European countries[J]. Euro Surveill, 2009, 14(40): 19352.
- [11] Lazer D, Kennedy R, King G, et al. The parable of google flu: traps in big data analysis[J]. Science, 2014, 343(6176): 1203-1205. DOI: 10.1126/science.1248506.
- [12] 邹晓辉,朱闻斐,杨磊,等. 谷歌流感预测——大数据在公共卫生领域的尝试[J]. 中华预防医学杂志, 2015, 49(6): 581-584. DOI: 10.3760/cma.j.issn.0253-9624.2015.06.026. Zou XH, Zhu WF, Yang L, et al. Google Flu Trends—the initial application of big data in public health[J]. Chin J Prev Med, 2015, 49(6): 581-584. DOI: 10.3760/cma.j.issn.0253-9624.2015.06.026.
- [13] Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages[C]//Proceedings of the 1st workshop on social media analytics. New York, NY, USA: ACM, 2010: 115-122. DOI: 10.1145/1964858.1964874.
- [14] Lamos V, Cristianini N. Tracking the flu pandemic by monitoring the social web [C]//Proceedings of the 2nd international workshop on cognitive information processing. Elba: IEEE, 2010: 411-416. DOI: 10.1109/CIP.2010.5604088.
- [15] Achrekar H, Gandhe A, Lazarus R, et al. Predicting Flu Trends using Twitter data [C]//Proceedings of IEEE conference on computer communications workshops (INFOCOM WKSHPS). Shanghai, China: IEEE, 2011: 702-707. DOI: 10.1109/INFCOMW.2011.5928903.
- [16] Aslam AA, Tsou MH, Spitzberg BH, et al. The reliability of tweets as a supplementary method of seasonal influenza surveillance[J]. J Med Internet Res, 2014, 16(11): e250. DOI: 10.2196/jmir.3532.
- [17] Gomide J, Veloso A, Meira Jr W, et al. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter [C]//Proceedings of the 3rd international web science conference. New York, NY, USA: ACM, 2011: 3. DOI: 10.1145/2527031.2527049.
- [18] Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak[J]. Am J Trop Med Hyg, 2012, 86(1): 39-45. DOI: 10.4269/ajtmh.2012.11-0597.
- [19] Odium M, Yoon S. What can we learn about the Ebola outbreak from tweets? [J]. Am J Infect Control, 2015, 43(6): 563-571. DOI: 10.1016/j.ajic.2015.02.023.
- [20] Friesema IHM, Koppeschaar CE, Donker GA, et al. Internet-based monitoring of influenza-like illness in the general population: experience of five influenza seasons in the Netherlands [J]. Vaccine, 2009, 27(45): 6353-6357. DOI: 10.1016/j.vaccine.2009.05.042.
- [21] Chunara R, Goldstein E, Patterson-Lomba O, et al. Estimating influenza attack rates in the United States using a participatory cohort[J]. Sci Rep, 2015, 5: 9540. DOI: 10.1038/srep09540.
- [22] Smolinski MS, Crawley AW, Baltrusaitis K, et al. Flu Near You: Crowdsourced symptom reporting spanning 2 influenza seasons [J]. Am J Public Health, 2015, 105(10): 2124-2130. DOI: 10.2105/ajph.2015.302696.
- [23] Paolotti D, Carnahan A, Colizza V, et al. Web-based participatory surveillance of infectious diseases: the influenzanet participatory surveillance experience[J]. Clin Microbiol Infect, 2014, 20(1): 17-21. DOI: 10.1111/1469-0691.12477.
- [24] Tilston NL, Eames KTD, Paolotti D, et al. Internet-based surveillance of Influenza-like-illness in the UK during the 2009 H1N1 influenza pandemic[J]. BMC Public Health, 2010, 10(1): 650. DOI: 10.1186/1471-2458-10-650.
- [25] Chunara R, Aman S, Smolinski M, et al. Flu near you: an online self-reported influenza surveillance system in the USA [J]. Online J Public Health Inform, 2013, 5(1): e133. DOI: 10.5210/ojphi.v5i1.4456.
- [26] Ocampo AJ, Chunara R, Brownstein JS. Using search queries for malaria surveillance, Thailand [J]. Malar J, 2013, 12(1): 390. DOI: 10.1186/1475-2875-12-390.
- [27] Domnich A, Arbutova EK, Signori A, et al. Demand-based web surveillance of sexually transmitted infections in Russia [J]. Int J Public Health, 2014, 59(5): 841-849. DOI: 10.1007/s00038-014-0581-7.
- [28] Cooper CP, Mallon KP, Leadbetter S, et al. Cancer internet search activity on a major search engine, United States 2001-2003 [J]. J Med Internet Res, 2005, 7(3): e36. DOI: 10.2196/jmir.7.3.e36.
- [29] Bahk GJ, Kim YS, Park MS. Use of internet search queries to enhance surveillance of foodborne illness [J]. Emerg Infect Dis,

- 2015, 21(11):1906–1912. DOI:10.3201/eid2111.141834.
- [30] Ayers JW, Althouse BM, Allem JP, et al. Seasonality in seeking mental health information on google[J]. *Am J Prev Med*, 2013, 44(5):520–525. DOI:10.1016/j.amepre.2013.01.012.
- [31] Yang AC, Huang NE, Peng CK, et al. Do seasons have an influence on the incidence of depression? The use of an internet search engine query data as a proxy of human affect[J]. *PLoS One*, 2010, 5(10):e13728. DOI:10.1371/journal.pone.0013728.
- [32] Newkirk RW, Bender JB, Hedberg CW. The potential capability of social media as a component of food safety and food terrorism surveillance systems[J]. *Foodborne Pathog Dis*, 2012, 9(2):120–124. DOI:10.1089/fpd.2011.0990.
- [33] Harris JK, Mansour R, Choucair B, et al. Health department use of social media to identify foodborne illness-Chicago, Illinois, 2013–2014[J]. *MMWR Morb Mortal Wkly Rep*, 2014, 63(32):681–685.
- [34] Simmering JE, Polgreen LA, Polgreen PM. Web search query volume as a measure of pharmaceutical utilization and changes in prescribing patterns[J]. *Res Soc Adm Pharm*, 2014, 10(6):896–903. DOI:10.1016/j.sapharm.2014.01.003.
- [35] White RW, Tatonetti NP, Shah NH, et al. Web-scale pharmacovigilance: listening to signals from the crowd[J]. *J Am Med Inform Assoc*, 2013, 20(3):404–408. DOI:10.1136/amiajnl-2012-001482.
- [36] Huang JD, Zheng R, Emery S. Assessing the impact of the national smoking ban in indoor public places in China: evidence from quit smoking related online searches[J]. *PLoS One*, 2013, 8(6):e65577. DOI:10.1371/journal.pone.0065577.
- [37] Ayers JW, Althouse BM, Allem JP, et al. A novel evaluation of World No Tobacco day in Latin America[J]. *J Med Internet Res*, 2012, 14(3):e77. DOI:10.2196/jmir.2148.
- [38] Troelstra SA, Bosdriesz JR, de Boer MR, et al. Effect of tobacco control policies on information seeking for smoking cessation in the Netherlands: a google trends study[J]. *PLoS One*, 2016, 11(2):e0148489. DOI:10.1371/journal.pone.0148489.
- [39] Ayers JW, Althouse BM, Ribisl KM, et al. Digital detection for tobacco control: online reactions to the 2009 U.S. cigarette excise tax increase[J]. *Nicotine Tob Res*, 2014, 16(5):576–583. DOI:10.1093/ntr/ntt186.
- [40] Chunara R, Bouton L, Ayers JW, et al. Assessing the online social environment for surveillance of obesity prevalence[J]. *PLoS One*, 2013, 8(4):e61373. DOI:10.1371/journal.pone.0061373.
- [41] Young SD, Rivers C, Lewis B. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes[J]. *Prev Med*, 2014, 63:112–115. DOI:10.1016/j.pymed.2014.01.024.
- [42] Freifeld CC, Brownstein JS, Menone CM, et al. Digital drug safety surveillance: monitoring pharmaceutical products in Twitter[J]. *Drug Saf*, 2014, 37(5):343–350. DOI:10.1007/s40264-014-0155-x.
- [43] Myslín M, Zhu SH, Chapman W, et al. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products[J]. *J Med Internet Res*, 2013, 15(8):e174. DOI:10.2196/jmir.2534.
- [44] Mandl KD, Mcnabb M, Marks N, et al. Participatory surveillance of diabetes device safety: a social media-based complement to traditional FDA reporting[J]. *J Am Med Inform Assoc*, 2014, 21(4):687–691. DOI:10.1136/amiajnl-2013-002127.
- [45] Weitzman ER, Kelemen S, Quinn M, et al. Participatory surveillance of hypoglycemia and harms in an online social network[J]. *JAMA Intern Med*, 2013, 173(5):345–351. DOI:10.1001/jamainternmed.2013.2512.
- [46] Adler AJ, Eames KTD, Funk S, et al. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey[J]. *BMC Infect Dis*, 2014, 14(1):232. DOI:10.1186/1471-2334-14-232.
- [47] Cantarelli P, Debin M, Turbelin C, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance[J]. *BMC Public Health*, 2014, 14(1):984. DOI:10.1186/1471-2458-14-984.
- [48] van Noort SP, Muehlen M, Rebelo De Andrade H, et al. Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe[J]. *Euro Surveill*, 2007, 12(7):E5–6.
- [49] 苏雪梅, 陈强, 万明, 等. 基于网络媒体的疾病暴发与监测信息资源的现状与思考[J]. *医学信息学杂志*, 2009, 30(2):1–5. DOI:10.3969/j.issn.1673-6036.2009.02.001.
- Su XM, Chen Q, Wan M, et al. The present status and thought on network-media-based information resources of disease outbreak and surveillance[J]. *J Med Inform*, 2009, 30(2):1–5. DOI:10.3969/j.issn.1673-6036.2009.02.001.
- [50] Dalton C, Durrheim D, Fejsa J, et al. Flutracking: a weekly Australian community online survey of influenza-like illness in 2006, 2007 and 2008[J]. *Commun Dis Intell Q Rep*, 2009, 33(3):316–322.
- [51] Carlson SJ, Durrheim DN, Dalton CB. Flutracking provides a measure of field influenza vaccine effectiveness, Australia, 2007–2009[J]. *Vaccine*, 2010, 28(42):6809–6810. DOI:10.1016/j.vaccine.2010.08.051.
- [52] Gluskin RT, Johansson MA, Santillana M, et al. Evaluation of internet-based dengue query data: google dengue trends[J]. *PLoS Negl Trop Dis*, 2014, 8(2):e2713. DOI:10.1371/journal.pntd.0002713.
- [53] Ortiz JR, Zhou H, Shay DK, et al. Monitoring influenza activity in the United States: a comparison of traditional surveillance systems with google flu trends[J]. *PLoS One*, 2011, 6(4):e18687. DOI:10.1371/journal.pone.0018687.
- [54] Dugas AF, Hsieh YH, Levin SR, et al. Google flu trends: correlation with emergency department influenza rates and crowding metrics[J]. *Clin Infect Dis*, 2012, 54(4):463–469. DOI:10.1093/cid/cir883.
- [55] Ginsberg J, Mohebbi MH, Patel RS, et al. Detecting influenza epidemics using search engine query data[J]. *Nature*, 2009, 457(7232):1012–1014. DOI:10.1038/nature07634.
- [56] Burton SH, Tanner KW, Giraud-Carrier CG, et al. “Right time, right place” health communication on Twitter: value and accuracy of location information[J]. *J Med Internet Res*, 2012, 14(6):e156. DOI:10.2196/jmir.2121.
- [57] McKee R. Ethical issues in using social media for health and health care research[J]. *Health Policy*, 2013, 110(2/3):298–301. DOI:10.1016/j.healthpol.2013.02.006.
- [58] Gesualdo F, Stilo G, Agricola E, et al. Influenza-like illness surveillance on Twitter through automated learning of naive language[J]. *PLoS One*, 2013, 8(12):e82489. DOI:10.1371/journal.pone.0082489.
- [59] Gittelman S, Lange V, Gotway Crawford CA, et al. A new source of data for public health surveillance: facebook likes[J]. *J Med Internet Res*, 2015, 17(4):e98. DOI:10.2196/jmir.3970.

(收稿日期:2016-07-06)

(本文编辑:斗智)