

# 利用网络数据分析我国活禽市场与人感染H7N9禽流感病例的地理关系

于伟文 杜鹏程 陈晨 卢珊 阙飙 杜孝平 徐建国

**【摘要】** 目的 探讨利用互联网数据分析活禽市场与人感染H7N9禽流感病例之间的地理关系。方法 利用网络大数据的方法,对人感染H7N9禽流感病例附近的活禽市场开展基于地理位置和可及性研究,搜索全部有病例报告地区的活禽市场,分析不同城市间活禽市场密度、人群密度、禽类生产密度等因素。结果 对42个报告病例的地级以上城市10 615个活禽市场与病例所在地的距离分析表明,发病例数与常住人口密度、市场数量及其密度呈正相关性;在已报告的13个H7N9禽流感病毒检测阳性活禽市场中,除广东佛山市的3个市场外,其余都在5 km内发现病例,表明活禽交易与病例的居住地分布密切相关;共识别出13个病例较集中的热点区域,分布在杭州、深圳等市。结论 H7N9禽流感病毒检测阳性活禽市场所在城市的报告病例数明显高于其他城市,且影响范围多在5 km范围内,应重新审视农贸活禽市场在城市中的布局 and 设置。

**【关键词】** H7N9禽流感病毒;禽类交易市场;风险预测;实时网络信息技术

**Investigation of geographic relationship between farmer's market with live birds and human infections with avian influenza A (H7N9) virus based on internet data analysis** Yu Weiwen<sup>1,2</sup>, Du Pengcheng<sup>1</sup>, Chen Chen<sup>1</sup>, Lu Shan<sup>1</sup>, Kan Biao<sup>1</sup>, Du Xiaoping<sup>2</sup>, Xu Jianguo<sup>1</sup>. 1 State Key Laboratory for Infectious Diseases Prevention and Control, National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2 College of Software Engineering, Beijing University of Aeronautics & Astronautics  
Corresponding author: Xu Jianguo, Email: xujianguo@icdc.cn

This work was supported by grants from the National Natural Science Foundation of China (No. 81290345) and the National Key Program of Mega Infectious Diseases (No. 2011ZX10004-001).

**【Abstract】 Objective** Farmer's markets with live birds (FMLB) are key sites where human infections by influenza A virus subtype H7N9 happened. Approximately 80% cases have exposed to FMLB. This study is to investigate the geographic relationship between FMLB and human cases based on analysis of internet data of their geographic locations. **Methods** Using big data from internet, we searched all FMLB in the cities where the human cases have been reported, then analyzed geographic relations, and evaluated the possibility of visits of the patients to the FMLB around them. The densities of FMLB, population and live poultries were also analyzed. **Results** Forty-two cities and 10 615 markets were included in the study. It is indicated that the number of human cases has positive correlations with the population density, the number and density of markets. Except three markets in Foshan, human cases have been reported within 5 km of 10 of 13 markets, which shows that the live bird trading is highly relevant with the distribution of cases. We identified 13 hot spots in the cities including Hangzhou, Shenzhen, et al, where clustered cases have emerged. **Conclusion** The numbers of human cases are significantly high in cities where FMLB are detected positive for H7N9 virus. These virus positive markets usually affect the people's residence within 5km area. The number and location of FMLB in cities should be re-evaluated and re-planned for healthy city where the risk of residents infecting avian influenza virus is greatly reduced or eliminated.

**【Key words】** Avian influenza A (H7N9) virus; Farmer's markets with live birds; Risk forecast; Real-time information technology

DOI: 10.3760/cma.j.issn.0254-6450.2014.03.011

基金项目:国家自然科学基金(81290345); 国家科技重大专项(2011ZX10004-001)

作者单位:102206 北京,中国疾病预防控制中心传染病预防控制所 传染病预防控制国家重点实验室 感染性疾病诊治协同创新中心(于伟文、杜鹏程、陈晨、卢珊、阙飙、徐建国); 北京航空航天大学软件学院(于伟文、杜孝平)

于伟文、杜鹏程同为第一作者

通信作者:徐建国, Email: xujianguo@icdc.cn

人感染H7N9禽流感病例约80%具有明确的活禽市场暴露史。在流行地区、流行期间关闭活禽市场是控制疫情的有效手段<sup>[1,2]</sup>。但禽类感染H7N9禽流感病毒并不发生死亡,也无明显的临床症状<sup>[3,4]</sup>。该特征使预防和控制人感染H7N9禽流感病毒的工作异常艰难。目前很难使用常规技术获得活禽市场与人感染H7N9禽流感病毒相关的地理距离、范围、强度等特点数据。实时网络信息技术可实时获取网络产生的海量数据,通过筛选并对其真伪进行统计性分析进而得出相应结论,与传统信息学技术相比,其获取信息更可靠、全面、快速。谷歌公司已利用该技术建立了流感监测实时信息系统,比常规监测方法提前1周发现流感暴发<sup>[5]</sup>。有研究利用航空介导的流行病全球性扩散过程,重新定义了有效移动距离作为流行病学意义上的地理距离,并成功用于解释2009年甲型H1N1流感和2003年SARS暴发传播过程<sup>[6,7]</sup>。为此本研究旨在利用人感染H7N9禽流感病例与相关禽类交易市场信息,结合使用实时网络信息技术中获得的市場信息、禽类交易信息和人群信息发现H7N9禽流感病例的空间特征,对重新定位、查找可能携带H7N9禽流感病毒的禽类和市场,及控制传播和扩散具有重要参考意义。

### 资料与方法

1. 资料来源:人感染H7N9禽流感病例信息数据来自“中国疾病预防控制中心信息系统”中的“传染病报告信息管理系统”数据库,包含截至2014年2月10日作为乙类传染病进行网络直报的所有病例。H7N9禽流感病毒感染相关市场包括病历报告中记录病例工作单位或住址的市场、网络新闻中报道的经过流行病学调查证实病例发病前去过的市场及家禽和环境样本H7N9禽流感病毒检测阳性的市场。在百度地图中(百度公司,中国北京)通过关键词“地名+农贸市场”进行搜索和查找,获取符合条件的市场名称、地址、经纬度坐标等信息,通过手工方式进行筛选,去除与活禽交易无关的场所。出现病例城市的人口、地理、经济数据来自各地区统计年鉴。

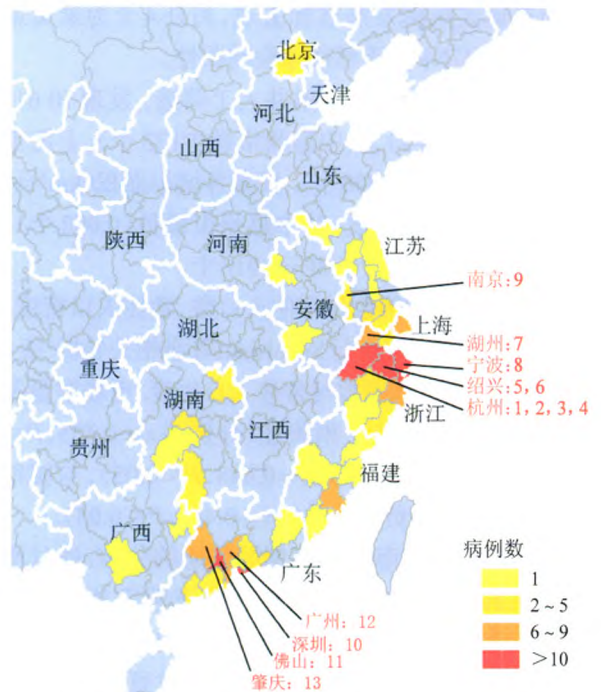
2. 分析方法及定义:调用谷歌地图(谷歌公司,美国加利福尼亚州圣克拉拉)相应功能,利用经纬度坐标计算病例居住地和市場间的两两最小公路距离作为有效地理距离,用以定义病例和市場间的距离。通过对已知H7N9禽流感病毒感染相关市場和病例之间的距离进行统计分析,确定病例与相关市場距离的分布特征,结合我国城市人群出行特征研

究数据,确定距离界值用以定义热点区域。凡出现病例区域的任何3个病例,若在以上界值范围内存在三者重合的市場,则该3个病例被假设为热点暴发病例,与之相邻的市場被定义为热点市場。天然地理障碍(如河流、山脉等)需经人工修正。

3. 统计学分析:使用SPSS 16.0软件,对各城市发病例数与城市人口、面积、市場分布等因素进行Spearman等级相关分析。

### 结果

1. 病例、市場地域分布及相关因素:186例报告病例(源自“传染病报告信息管理系统”数据库)共涉及42个地级以上城市,分布于北京(2例)、上海(8例)、江苏(12例)、浙江(82例)、安徽(2例)、福建(15例)、湖南(8例)、广东(55例)、广西(2例)共9个省、自治区和直辖市(图1)。其中浙江、广东两省病例数占报告病例数的73.7%。有24个城市发病例数<3例;浙江省杭州、宁波、绍兴和广东省深圳、佛山报告病例数均>10例,占总报告病例数的44.6%。这些城市为人口集中、经济较为发达的大中型城市(图2和表1)。



注:红色为热点区域编号及所在城市

图1 2013年12月以来我国发生人感染H7N9禽流感病例的城市和热点区域分布

来自7个[广州(2)、佛山(5)、深圳(1)、杭州(1)、台州(2)、绍兴(1)、湖州(1)]有报告病例城市共搜索到13个与H7N9禽流感病毒感染相关的市場;

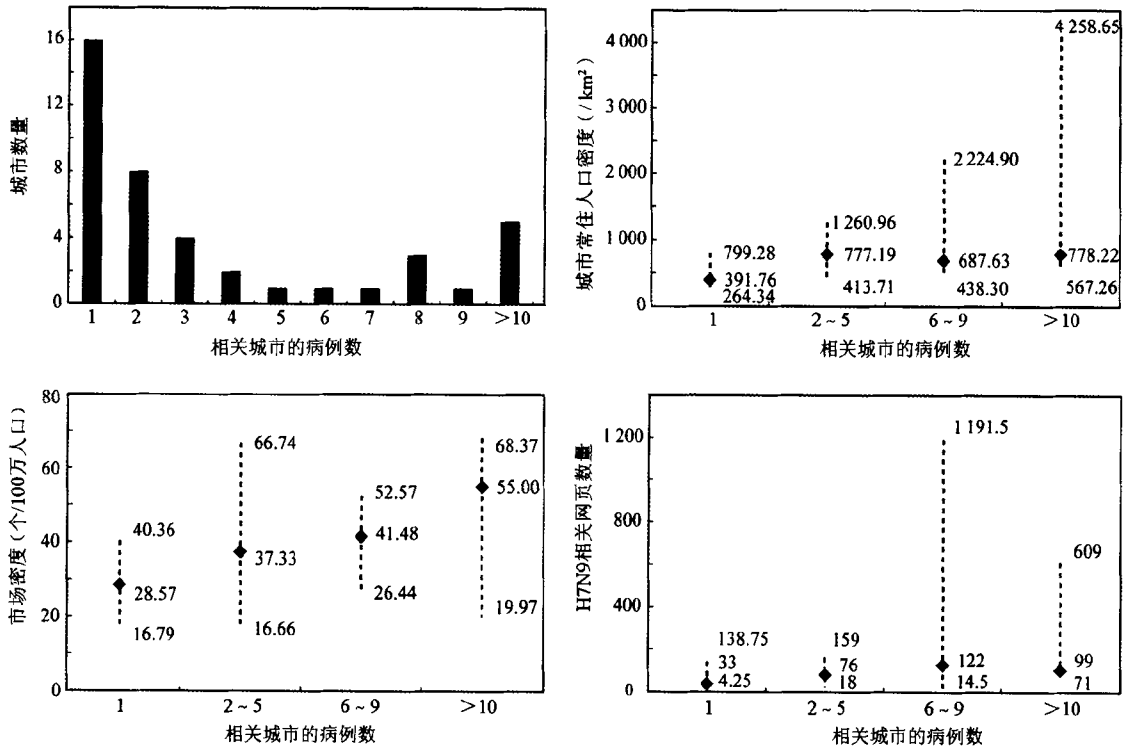


图2 人感染 H7N9 禽流感病例与城市数量、人口密度、市场密度及相关网页数量的关系(M和Q)

其中3个为病例工作单位或居住地点,6个经流行病学调查证实病例发病前曾去过,另外4个为家禽或环境样本H7N9禽流感病毒检测阳性。

42个城市经自动搜索并人工筛选,发现10615个农贸市场(表1),其中惠州市最少(23个),上海市最多(679个),M=200个。根据各城市面积(km<sup>2</sup>)计算得到农贸市场密度,其中惠州市最低(0.002个),东莞市最高(0.183个),M=0.018个。

将各城市病例数与常住人口数、常住人口密度、市场数量及其密度进行Spearman等级相关分析(表1和图2),显示发病例数与常住人口密度、市场数量及其密度具有一定的正相关性(分别为 $r_s=0.455, P<0.05; r_s=0.398, P<0.05; r_s=0.448, P<0.05$ ),与城市面积具有一定的负相关性( $r_s=-0.446, P<0.05$ )。

2. 相关活禽市场与病例居住地的分布:通过网络收集到的13个H7N9禽流感病毒感染相关市场(表2),主要分布于广州、佛山、深圳、杭州、台州、绍兴和湖州,其中广州、佛山、深圳的农贸市场H7N9禽流感病毒检测阳性,且报告病例数显著高于其他城市。有研究发现,以出发地至目的地计算,我国52.9%的城市居民出行距离分为3~5 km和5~8 km两层次<sup>[8-10]</sup>。因此本研究将5 km作为报告病例与市场距离的界值。发现除佛山3个市场外,其余所有H7N9禽流感病毒感染相关市场周边(<5 km)至少

有1名病例。其中距离病例最近的市场发现在台州,其距离仅为0.07 km;而广州的1个市场距离病例最远,为4.89 km。深圳一个H7N9禽流感病毒检测阳性市场周边出现病例最多(6例),且与病例距离较近(1.57 km)。186例中有172例(92.47%)居住地周边5 km内可搜索到活禽农贸市场,表明活禽交易市场与病例居住地分布密切相关。

3. 热点区域和病例居住地与活禽市场距离的分布:以5 km为限定区域,若该区域内任意3名病例存在与至少1个市场重合,则该3名病例及其5 km范围内均作为热点区域,该城市则作为有可能发生疫情暴发的热点城市。在杭州、绍兴、湖州、宁波、广州、深圳、佛山、肇庆和南京9个城市中识别到13个热点区域(图1和表3),其中杭州4个、绍兴2个,其余城市均为1个。共涉及病例73人,占总病例数的39.3%,其中深圳热点区域SZ1的相关病例数最多,为15例。在病例数>10人的城市中,热点区域聚集病例所占比例(热点区域病例数/总病例数)分别为杭州(19/28)、绍兴(11/12)、宁波(9/15)、深圳(15/17)、佛山(3/11),其中佛山热点区域涉及病例比例最低,表现为散发病例居多,而其他城市则为聚集病例居多。

热点区域内共涉及活禽农贸市场1003个,占9个城市全部市场总数的32.3%,其中广州热点区域GZ1涉及市场最多,为255个。提示这些市场在相

表1 我国186例人感染H7N9禽流感病例与相关城市人口、地理及经济数据的关系

省(区、市)	城市	人口数(万)	面积(km <sup>2</sup> )	病例数	人口密度(/km <sup>2</sup> )	市场个数	市场密度(个/100万人口)	家禽年出栏(万只)
北京	北京	2 069.30	16 410.54	2	1 260.96	647	31.27	11 774.32 <sup>a</sup>
上海	上海	2 380.43	6 340.50	3	3 754.33	679	28.52	4 083.76 <sup>a</sup>
江苏	南京	816.10	6 597.00	4	1 237.08	519	63.60	3 248.85 <sup>b</sup>
	无锡	638.00	4 787.60	2	1 332.61	464	72.73	2 960.00 <sup>b</sup>
	徐州	976.66	11 258.00	1	867.53	214	21.91	20 454.80 <sup>c</sup>
	苏州	1 065.40	8 488.00	2	1 255.18	490	45.99	7 311.40 <sup>c</sup>
	淮安	480.34	10 072.00	1	476.91	186	38.72	5 758.50 <sup>c</sup>
	盐城	726.02	17 000.00	1	427.07	297	40.91	20 411.20 <sup>c</sup>
	泰州	507.00	5 787.00	1	876.10	160	31.56	4 592.20 <sup>c</sup>
浙江	杭州	880.02	16 596.00	28	530.26	484	55.00	4 425.81 <sup>c</sup>
	宁波	763.90	9 816.00	15	778.22	120	15.71	2 994.65 <sup>c</sup>
	温州	916.00	11 786.00	5	777.19	409	44.65	2 337.77 <sup>d</sup>
	嘉兴	454.40	3 915.00	4	1 160.66	497	109.38	4 204.36 <sup>e</sup>
	湖州	289.35	5 818.00	9	497.34	140	48.38	4 977.16 <sup>b</sup>
	绍兴	500.26	8 279.00	12	604.25	351	70.16	2 252.11 <sup>c</sup>
	金华	536.16	10 941.00	2	490.05	97	18.09	2 571.16 <sup>c</sup>
	台州	600.50	9 411.00	6	638.08	391	65.11	3 268.74 <sup>c</sup>
	丽水	211.70	17 298.00	1	122.38	69	32.59	1 315.76 <sup>b</sup>
安徽	安庆	618.65	15 398.00	1	401.77	93	15.03	269.00 <sup>c</sup>
	阜阳	1 014.00	9 775.00	1	1 037.34	69	6.80	1 048.03 <sup>c</sup>
福建	福州	711.54	11 968.00	1	594.54	225	31.62	2 379.28 <sup>c</sup>
	厦门	367.00	1 699.39	3	2 159.60	137	37.33	-
	三明	266.72	22 928.80	1	116.33	60	22.50	-
	泉州	812.00	11 015.00	8	737.18	164	20.20	2 309.26 <sup>d</sup>
	漳州	480.99	12 600.00	1	381.74	123	25.57	-
	宁德	339.03	13 452.00	1	252.03	54	15.93	-
湖南	邵阳	707.00	20 829.00	1	339.43	58	8.20	-
	岳阳	552.31	15 019.20	2	367.74	92	16.66	-
	永州	609.00	22 441.00	2	271.38	61	10.02	-
	娄底	418.40	8 117.60	3	515.42	33	7.89	-
广东	广州	1 275.00	7 434.00	7	1 715.09	537	42.12	11 136.00 <sup>b</sup>
	佛山	719.43	3 868.00	11	1 859.95	479	66.58	9 065.47 <sup>b</sup>
	深圳	1 300.18	1 953.00	17	6 657.35	315	24.23	456.06 <sup>c</sup>
	江门	445.00	9 541.00	2	466.41	297	66.74	9 780.90 <sup>c</sup>
	肇庆	391.80	15 000.00	8	261.20	160	40.84	11 380.00 <sup>c</sup>
	惠州	463.36	11 200.00	3	413.71	23	4.96	-
	梅州	521.35	15 925.00	1	327.38	101	19.37	-
	阳江	271.00	7 822.00	2	346.46	93	34.32	-
	东莞	825.48	2 465.00	1	3 348.80	453	54.88	-
	中山	312.09	1 783.67	3	1 749.71	261	83.63	-
广西	南宁	666.16	22 112.00	1	301.27	293	43.98	2 442.40 <sup>d</sup>
	贺州	223.00	11 855.00	1	188.11	220	98.65	-

注:<sup>a</sup> 2012年;<sup>b</sup> 2011年;<sup>c</sup> 2010年;<sup>d</sup> 2009年;<sup>e</sup> 2006年;- 缺少数据

应热点区域内的危险度较大,在流行病学调查和实施疫情防控措施时应予以重点关注。

## 讨 论

人感染H7N9禽流感主要是通过暴露的活禽市场传播,为一种典型的行为生态型新发传染病<sup>[11]</sup>。

但是,如果没有禽流感病毒,活禽市场对人并不构成威胁。在目前活禽市场被污染、市场活禽带病毒及从运输到市场交易环节的活禽带病毒不明情况下,研究活禽市场可能传播禽流感病毒的特点,对预防禽流感具有重要意义。

与传统流行病学和病原学方法不同,本研究通过实时网络信息技术和大数据分析技术,收集、分析海量网络数据,获取发生人感染H7N9禽流感病例地区的农贸市场分布及其相关信息,研究城市活禽市场与感染病例间的地理关系。在2013年人感染H7N9禽流感疫情的溯源分析中,本研究组曾使用“人肉搜索”和“网络爬虫技术”获得活禽交易相关的海量网络数据,分析全国活禽交易网络,对各地发病风险进行预测并推测出不同城市可能传播的路径<sup>[4]</sup>。2014年疫情的发展证实该预测具有较高的准确度。谷歌公司开发的流感、登革热等疾病监测实时信息系统也证明了此类技术应用用于公共卫生领域的可行性,能够为传染病的预测、预警提供新的有效信息来源<sup>[5]</sup>。同时基于网络信息分析的研究也表明,网络信息可作为评估传染病发病风险的参考。但如何将这些信息用于构建明确的疾病传播风险指数,目前尚无准确、可应用的数学模型,因此还需要更多的研究予以验证。

本研究发现,H7N9禽流感病毒检测阳性的农贸市场所在城市的报告病例数远远高于其他城市。研究中显示除佛山的3个市场外,所有H7N9禽流感病毒检测阳性的农贸市场周边(<5 km范围)至少有1名H7N9禽流感病例;92.47%的报告病例其居住地5 km范围内,至少有1个农贸市场。表明人感染H7N9禽流感与接触活禽市场的确存在密切关系,同时也证

实在应对2013年人感染H7N9禽流感疫情中,在特定地区实施关闭活禽市场、暂停活禽交易等措施是正确的。因此,我国公共卫生策略应从长计议,在城市特别是大型城市重新考虑农贸市场的布局 and 设置。

(感谢北京易贸创想软件有限公司宋华文、刘庆肖在网络数据采集工作中给予的技术支持)



表 2 H7N9 禽流感病毒感染相关市场与病例的距离及其分布

城市	病例数	市场名称	与 H7N9 禽流感病毒感染关系 <sup>a</sup>	与最近病例距离 (km)	5 km 范围内病例数
广州	7	天河区长湴综合市场	1	4.89	1
		天河区沙东市场	3	4.64	1
佛山	11	北滘合诚农贸市场	1	12.54	0
		三洲农产品交易中心	1	21.56	0
		黄岐永盛市场	3	8.94	0
		南海区大沥桂江三鸟批发市场	3	1.98	1
		盐步环球水产市场	2	4.72	1
深圳	17	龙岗区南湾街道康桥肉菜市场	1	1.57	6
杭州	28	益农镇农贸市场	2	0.47	3
台州	6	三甲街道菜市场	3	0.37	1
		尤溪镇菜市场	3	0.07	1
绍兴	12	朱公湖农贸市场	3	1.35	3
湖州	9	德清新发地(原泰源)农贸市场	2	0.22	3

注: <sup>a</sup> 1: 禽类或环境样本 H7N9 禽流感病毒检测阳性; 2: 病例的单位或居住地; 3: 流行病学调查明确显示病例发病前曾去过该市场

表 3 2014 年我国 9 个城市 13 个人感染 H7N9 禽流感热点区域分布及其特征

热点区域编号	城市	病例数	热点区域市场个数
1	杭州	8	78
2	杭州	3	11
3	杭州	5	112
4	杭州	3	5
5	绍兴	3	16
6	绍兴	8	79
7	湖州	3	6
8	宁波	9	44
9	南京	3	185
10	深圳	15	129
11	佛山	3	87
12	广州	6	255
13	肇庆	4	11

参 考 文 献

[1] Gao R, Cao B, Hu Y, et al. Human infection with a novel avian-origin influenza A (H7N9) virus [J]. N Engl J Med, 2013, 368 (20):1888-1897.

[2] Xu J, Lu S, Wang H, et al. Reducing exposure to avian influenza H7N9 [J]. Lancet, 2013, 381(9880):1815-1816.

[3] Zhou J, Wang D, Gao R, et al. Biological features of novel avian influenza A (H7N9) virus [J]. Nature, 2013, 499(7459): 500-503.

[4] Chen C, Lu S, Du P, et al. Silent geographical spread of the H7N9 virus by online knowledge analysis of the live bird trade

with a distributed focused crawler [J]. Emerg Microb Infect, 2013, 2(12):e89.

[5] Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks [J]. Clin Infect Dis, 2009, 49(10):1557-1564.

[6] McLean AR. Coming to an airport near you [J]. Science, 2013, 342(6164):1330-1331.

[7] Brockmann D, Helbing D. The hidden geometry of complex, network-driven contagion phenomena [J]. Science, 2013, 342(6164):1337-1342.

[8] [http://news.xinhuanet.com/mrdx/2009-11/02/content\\_12372481.htm](http://news.xinhuanet.com/mrdx/2009-11/02/content_12372481.htm).

[9] Huang SS, Song R, Tao Y. Behavior of urban residents travel mode choosing and influencing factors: taking Beijing as an example [J]. Commun Stand, 2008, 9:124-128. (in Chinese) 黄树森, 宋瑞, 陶媛. 大城市居民出行方式选择行为及影响因素研究——以北京市为例 [J]. 交通标准化, 2008, 9:124-128.

[10] Chen Z, Liu YS, Shi F. Study on distance diversion curve of inhabitant trips modal [J]. Commun Stand, 2006, 8:157-159. (in Chinese) 陈征, 刘英舜, 石飞. 城市居民出行方式距离曲线的研究与应用 [J]. 交通标准化, 2006, 8:157-159.

[11] Xu JG. Behavioral and ecological infectious diseases: from SARS to H7N9 avian influenza outbreak in China [J]. Chin J Epidemiol, 2013, 34(5):417-418. (in Chinese) 徐建国. 从 SARS 到人感染 H7N9 禽流感——行为生态型传染病 [J]. 中华流行病学杂志, 2013, 34(5):417-418.

(收稿日期: 2014-02-25)

(本文编辑: 张林东)