Utility of the ESSENCE Surveillance System in Monitoring the H1N1 Outbreak

Rekha S. Holtry[1], Lang M. Hung[1], Sheri H. Lewis[1]

[1]Johns Hopkins University, Applied Physics Laboratory

Abstract

The Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) enables health care practitioners to detect and monitor health indicators of public health importance. ESSENCE is used by public health departments in the National Capital Region (NCR); a cross-jurisdictional data sharing agreement has allowed cooperative health information sharing in the region since 2004. Emergency department visits for influenza-like illness (ILI) in the NCR from 2008 are compared to those of 2009. Important differences in the rates, timing, and demographic composition of ILI visits were found. By monitoring a regional surveillance system, public health practitioners had an increased ability to understand the magnitude and character of different ILI outbreaks. This increased ability provided crucial community-level information on which to base response and control measures for the novel 2009 H1N1 influenza outbreak. This report underscores the utility of automated surveillance systems in monitoring community-based outbreaks. There are several limitations in this study that are inherent with syndrome-based surveillance, including utilizing chief complaints versus confirmed laboratory data, discerning real disease versus those healthcare-seeking behaviors driven by panic, and reliance on visit counts versus visit rates.

Key Words: H1N1, swine flu, surveillance

Background

The Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) enables public health practitioners to observe abnormal behavior of health indicators across jurisdictions and view geographical spread of outbreaks that span across regions [1]. The Washington, DC Metropolitan area of the United States (U.S.), also referred to as the National Capital Region (NCR), encompasses two counties in Maryland, five counties in Virginia, and the District of Columbia. With its large population, high-profile establishments and events, the NCR draws visitors from all over the world, thereby leading to increased chances of introduction of emerging infectious diseases [2].

The NCR Disease Surveillance Network employs the ESSENCE system to maintain a regional surveillance capability. This network utilizes a variety of data sources collected both at a regional and jurisdictional level. Locally collected sources include Emergency Department (ED) chief complaints while regionally collected sources include over-the-counter (OTC) pharmaceutical sales and poison control call center data [3]. A cross-jurisdictional data sharing agreement established among the jurisdictions in 2004 allows cooperative sharing of health information across state/district boundaries for syndrome-based disease surveillance. Per the agreement, de-identified data from the secure local databases are transmitted to a central node for aggregation with data from other regional sources. These data are then made available to epidemiologists in the participating jurisdictions. Advanced visualization tools are used to organize the resulting wealth of information into a coherent view of population health status. Within the NCR, sharing of aggregated data via this distinctive architecture provides useful disease surveillance information to authorized public health users while complying with HIPAA privacy requirements.

In the spring of 2009, a novel influenza A (H1N1) virus resulted in increased influenza-like illness (ILI) activity in the U.S. throughout the summer and fall [4]. H1N1 was declared a pandemic, widespread human infection, by the World Health Organization (WHO) in early June 2009 [5]. As a result, public health departments across the U.S. were continually monitoring the health of their communities.

Methods

ED chief complaints, the primary source of clinical data in ESSENCE, were the data used for the analyses of this study. These data are typically provided as a free text field, and as a result, the data must go through a 13-step natural language parsing process to sanitize the text and determine syndrome grouping [6]. During the syndrome grouping step, ESSENCE bins chief complaints into two levels of progressively more sensitive groups. The first level groupings are called *subsyndromes* and the second, *syndromes*. For example, a chief complaint record containing “bad food” or “food” or “food poison” is assigned to the “food poisoning” subsyndrome, and the “food poisoning” subsyndrome, along with several other subsyndromes such as “diarrhea,” “vomiting,” and “gastroenteritis,” make up the Gastrointestinal (GI) syndrome. Furthermore, in addition to being defined by the baseline chief complaint terms, some subsyndromes can be defined by other subsyndromes. The hierarchical binning process of chief complaints to syndromes allows maximum sensitivity for capturing particular health conditions that can present in a variety of different ways while retaining the ability to narrow down health presentations by subsyndrome. Important to note is that the individual syndrome or subsyndrome definitions are determined by expert consensus and may be influenced by several factors such as data source characteristics, surveillance focus area, public health practitioner/agency priorities, etc [7]. In the NCR, the syndrome groups are: *Botulism-Like, Fever, Gastrointestinal, Hemorrhagic Illness, Localized Lesion, Lymphadenopathy, Neurological, Other, Rash,* *Respiratory,* and *Sudden-Illness/Death.* The OTC data source contains non-prescription medication and medical supply sales. For this data, ESSENCE bins applicable medications and supplies to the Fever, Gastrointestinal and the Respiratory syndrome only. Additionally, the user can query this data source by specific OTC category or OTC type.

For this study, ED chief complaint and OTC sales data in the NCR from 01 January 2008 through 31 December 2009 were used to compare the trends in ILI during the typical flu season and then after the discovery of the novel H1N1 strain. For ED data, to maximize specificity while maintaining sensitivity, time series were generated for the ILI subsyndrome made up of the Fever, Cough, and Sore Throat subsyndromes [8]. For the OTC data source, thermometer sales within OTC type were queried because thermometer sales were known to track closely with ILI trends [9].

Using the ESSENCE query portal, time series were generated for the ILI subsyndrome for all ages for 2008 and 2009. Then, additional time series were generated by age groups for the two years. OTC thermometer time series were also generated for each of the years. Thereafter, a time series for both years and a time series for each year by age group were plotted on a single graph using a specialized feature in ESSENCE. Lastly, ILI time series for 2008 and 2009 were plotted against OTC thermometer sales for those years.

Results

Figure 1 comparing the trends from January through late March for 2008 and 2009 show gradual seasonal elevation and subsequent decline in ILI counts for both years. However, for 2009 only, markedly elevated counts for ILI are seen around May, June, and October.

ED Visits Related to the ILI Subsyndrome in 2008 and 2009.

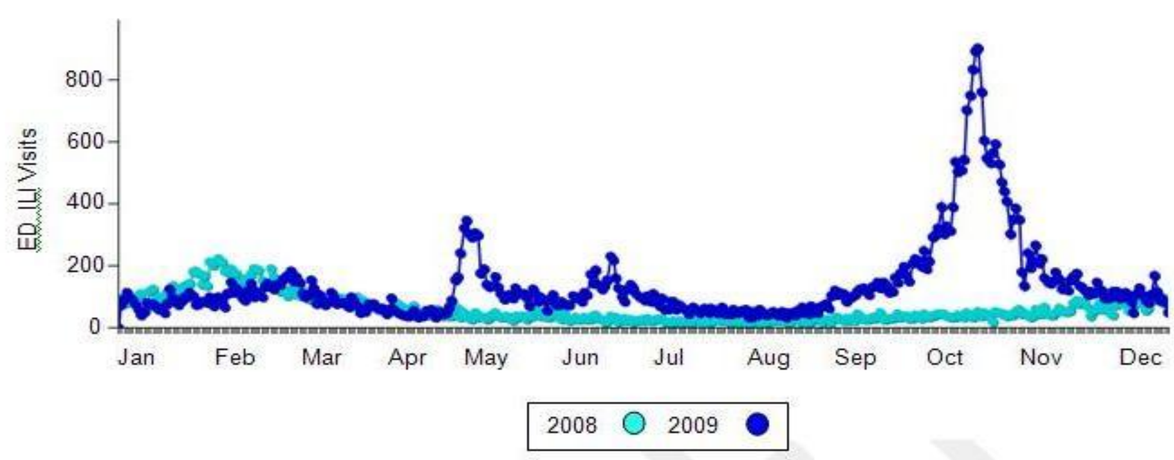


Figure 1. Time series of ILI-related chief complaints for 2008 and 2009.

In Figure 2, the time series graph by age group for 2008 shows seasonal elevation from the prior year into January and a gradual decline in counts for all age groups, with most counts in the 18-44 group. However, it is important to take into consideration that the 18-44 age group, having the largest age span, also encompasses the largest volume of counts and therefore may not necessarily reflect a true disproportion of disease. For 2009, remarkable findings include markedly high counts for all age groups with notably high counts for the 5-17 age group for each of the spikes around May, June, and October. Additionally in 2009, the 0-4 age group also closely follows this pattern; however, a relatively flat time series with unremarkable counts is seen for the 65+ age group.

ED Visits Related to the ILI Subsyndrome by Age Group in 2008 and 2009.

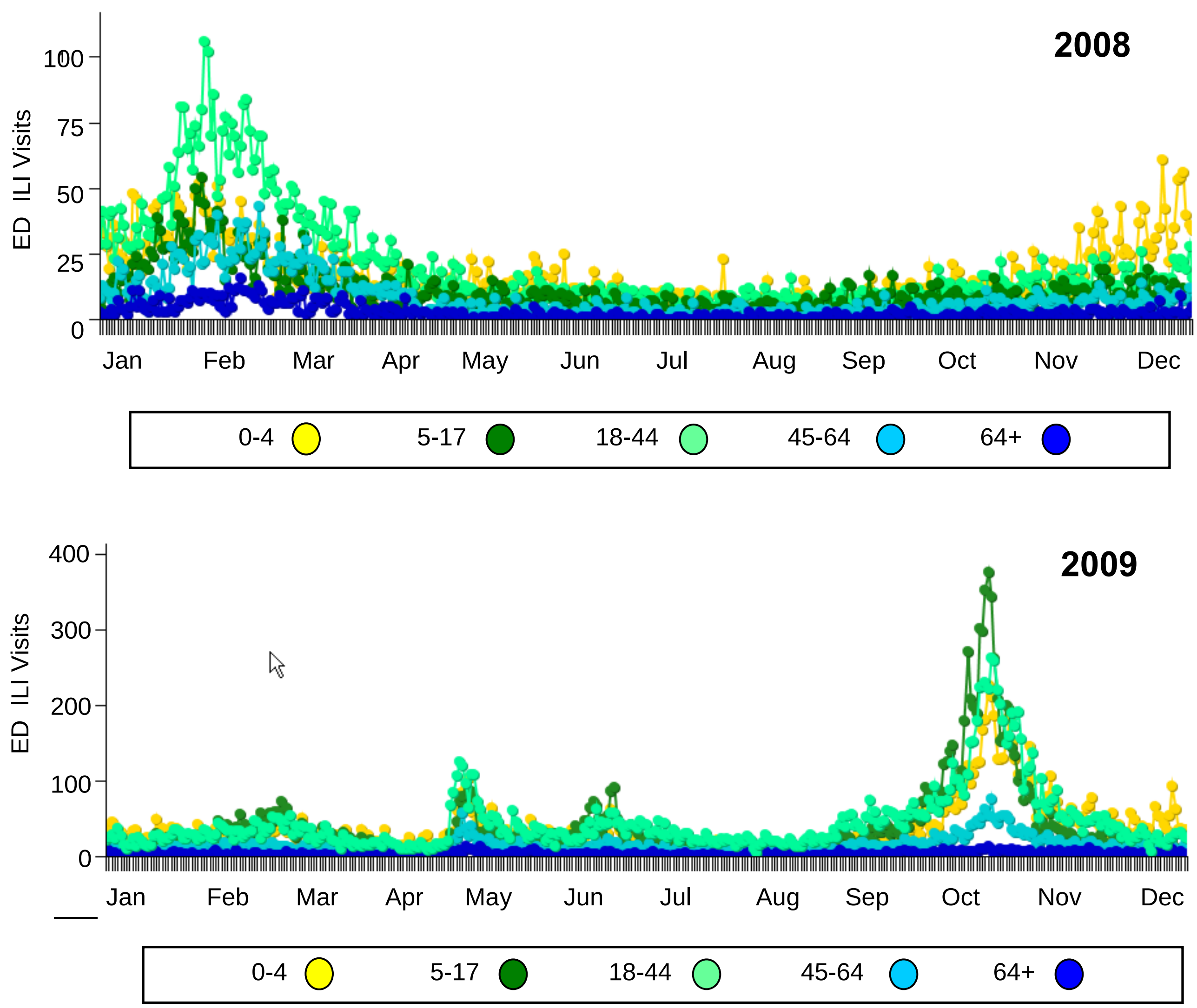


Figure 2. Time series of ILI chief complaints for 2008 and 2009 by Age Group.

Table 1 shows that age group 5-17 presented with the most dramatic increase in counts at 345% in 2009 as compared to 2008, followed by the 0-4 and 18-44 age groups (Table 1).

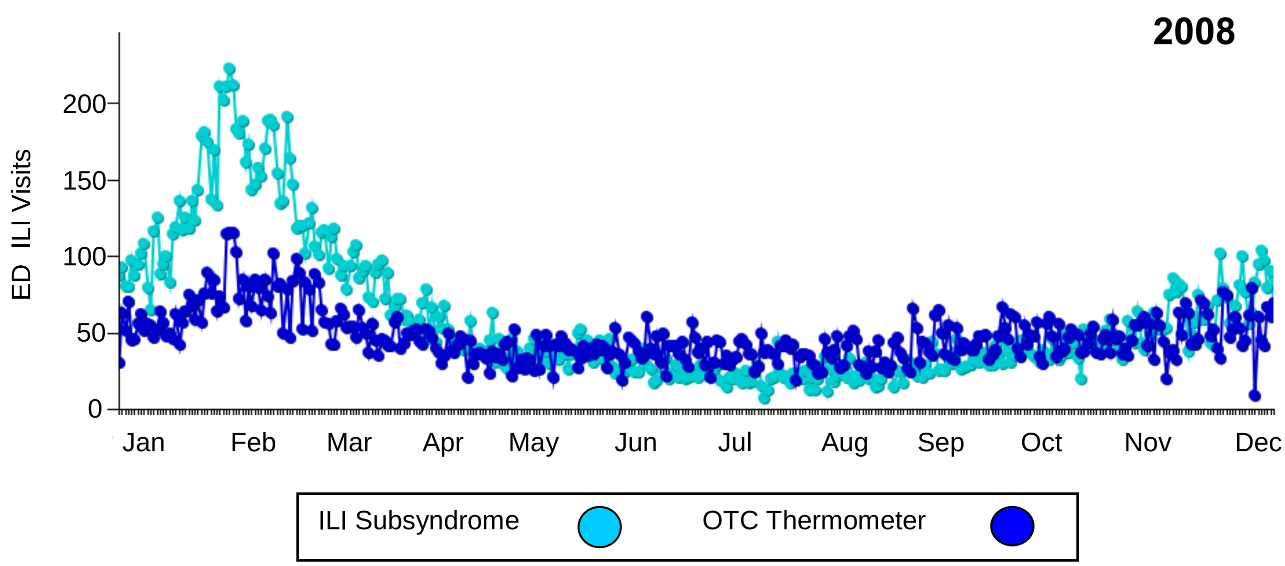
Table 1. ED visit counts of ILI-related chief complaints by age group for 2008 and 2009 with percent increase in cases for 2009.

Percent Increase in ED Visits Related to ILI Subsyndrome by Age Group in 2009 Compared to 2008.

AGE GROUP	2008 ILI COUNT	2009 ILI COUNT	PERCENT INCREASE IN CASES FOR 2009
0-4	6200	13574	218%
5-17	4065	14054	345%
18-44	7518	16398	218%
45-64	2944	5092	172%
65+	1005	1275	126%
TOTAL	**21,732**	**50,393**	**232%**

Figure 3 below shows a noticeable correlation between OTC thermometer sales and ED ILI visits for 2008 and 2009. Elevations in OTC thermometer sales can also be observed in May, June, and October time frames in 2009.

ED Visits Related to the ILI Subsyndrome and Corresponding OTC Thermometer Sales for All Ages in 2008 and 2009.



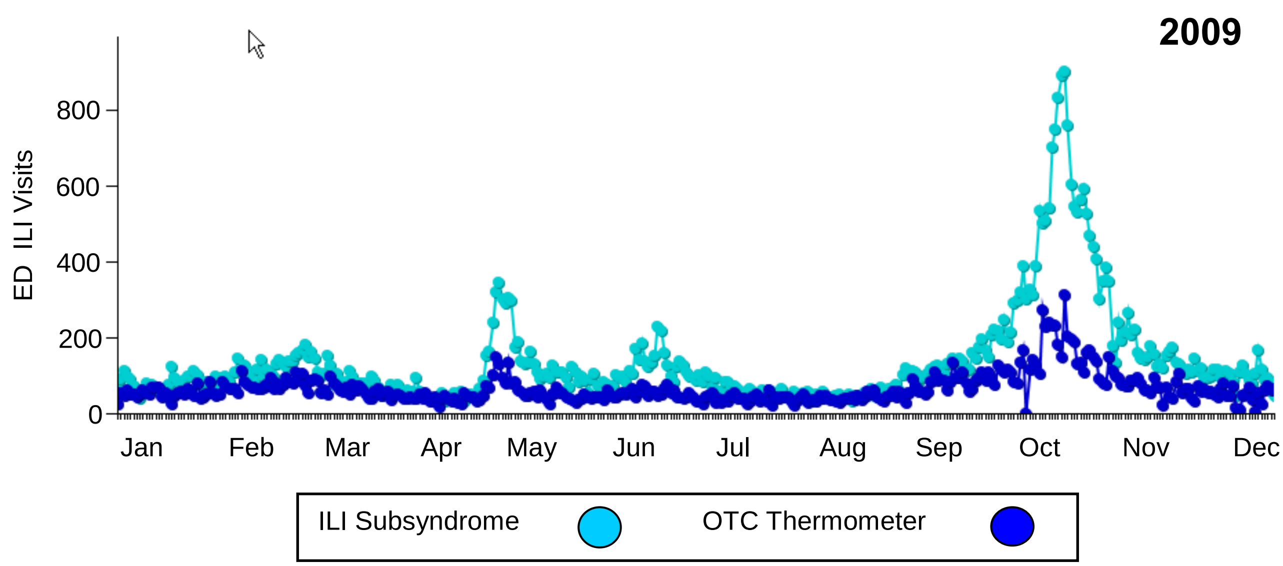


Figure 3. Time series of ILI-related chief complaints and OTC themometer sales for 2008 and 2009.

Discussion

The retrospective syndromic data show the now confirmed H1N1 influenza epidemics of 2009 (Figure 1) [10] with the first outbreak occurring in late spring of 2009 [11] and the second occurring in the fall of 2009 [12]. The two-year time frame highlights substantial differences in ILI activity after the introduction of H1N1 into the community. The frequency of ED visits more than doubled in 2009, thereby showing the effect of H1N1 in 2009 as compared to seasonal ILI in 2008. The 2009 H1N1 strain also displays an uncharacteristic early seasonal onset. This is starkly different from that of the seasonal flu, which typically peaks between late November and early March of a given year [13]. Furthermore, it appears that H1N1-related chief complaints tapered off in early November, which is historically considered the beginning of the seasonal flu [14].

When broken down by age groups, it is evident that H1N1 had a greater impact on the very young to middle-aged groups for the period of study (Figure 2, Table 1). While all the age groups had increased ILI presentations in 2009 as compared to 2008, ages 0-4, 5-17, and 18-44 appear to be most at risk (Figure 2, Table 1). This is in keeping with the clinical evidence that approximately 90% of hospitalizations and 88% of deaths from 2009 H1N1 occurred in people younger than 65 years old [15]. As ILI-related chief complaints increased, OTC purchases to remedy ILI symptoms also increased (Figure3). In 2008, both ED visits and OTC sales followed the typical seasonal flu trend. For 2009, while OTC thermometer sales for 2009 were again visibly correlated, atypical spikes were seen around May, June, and October periods with the onset, escalation, and decline of H1N1 outbreaks. These trends observed in the NCR are consistent with literature described by the Centers for Disease Control and Prevention (CDC) [9-11, 13]. The CDC noted that H1N1 appears to primarily affect the younger age groups, an important departure from seasonal flu, which impacts the vulnerable populations at both ends of the age spectrum [15]. These findings seen in ESSENCE in the NCR are important because they corroborate findings in other regions and trends seen nationally [16]. These findings support efforts by the CDC, as well as the state and local health department, whose efforts are focused on providing health messages and vaccinations to pre-school and school-aged children to slow down the spread of H1N1 [17]. Surveillance and monitoring will need to continue to determine whether the emergence of H1N1 has lasting effects on seasonal influenza patterns.

Study Limitations:

There are several limitations that are inherent with syndrome-based surveillance because of its reliance on health indicator data sources. Perhaps the most important among these to note is that ED patient visit counts, used in this study as the clinical data source, are based on patient chief complaints and triage entries rather than confirmed laboratory results. Hence, an implicit assumption in this study is that patients categorized into the ILI subsyndrome are those (at least the majority) infected with H1N1. This, however, is not an unreasonable assumption given that in August of 2009, the CDC, attested credibility to reporting unconfirmed cases based on ICD-9 syndromic diagnosies [18]. The CDC requested states to report both laboratory confirmed or "syndromic" cases of hospitaliztions and deaths. Because of the sheer enormity of outbreak, the CDC reasoned that testing could not be conducted on every individual seeking care and so "laboratory-confirmed data is a vast underestimate of the true number of cases and this bias would be exacerbated over the course of the pandemic as more and more people become ill" [18].

An associated limitation of using unconfirmed data is, discerning real disease from healthcare-seeking behavior driven by panic. In late April and early May of 2009, in the weeks after the recognition of cases in Mexico and the WHO reporting confirmed cases in California and other US states [19], fueled by media headlines, EDs across the US became inundated with record number of worried individuals seeking care. The vast majority presented with mild or no illness and did not require in-hospital care; furthermore, the increase in visits did not correspond to an increase in mortality [20]. However, due to the sensitive nature of syndrome-based surveillance, the "worried well" were captured within electronic disease surveillance systems in the NCR and in other states across the country [20]. In response, to increase specificity and enable extraction of cases warranting further scrutiny, NCR health departments searched incoming ED data for patients presenting with a fever, those who met specific chief complaints-based case definitions, and those who reported travel to affected areas. While these custom querying efforts partially aided in limiting the number of records for further investigation, it was still a laborious task for public health practitioners to isolate potential H1N1 cases within the large influx.

There are other limitations; systems such as ESSENCE rely on visit counts, and for several reasons do not typically divide by a denominator and report by rates of illness. Therefore, unreported (or unobserved) pre-processing data flow interruptions or delays can cause artifacts in time series that may be indistinguishable from real health patterns. Furthermore, by reporting counts and not rates, trends seen in time series may at first glance appear elevated simply because of the of nature of the data or the type of comparisons being made. An example within this study is the 18-44 age group presenting with the most counts on the time series. As discussed previously, this group represents a large span and proportion of the general population; high counts within this group may not necessarily mean disproportionate disease distributions. Yet, it is possible to overcome most of these limitations by remaining vigilant to incoming data streams, developing a good understanding of baseline counts typical for a particular region, understanding characteristics of the data and patient healthcare seeking behavior during large public health events, and employing common sense and judgment when interpreting the outputs from the system.

Conclusion

These findings document the emergence and spread of the 2009 H1N1 epidemics within specific health indicator data as seen in the NCR Disease Surveillance Network. The NCR has a unique network set up for health-indicator data collection and sharing that optimizes regional and inter-jurisdictional disease surveillance. In 2009, regular examination of the near real-time data sources through ESSENCE allowed public health practitioners in the NCR to better understand the full extent of illness in the community that may not have been captured by traditional sources. While there are important applications of electronic disease surveillance systems, there are known limitations that must be recognized prior to interpreting findings and drawing conclusions. Applied appropriately, supplementary information from systems such as ESSENCE can be invaluable and serve as a critical tool for public health decision-making.

In this study, the trends observed in representative ILI time series within ED and OTC data sources in the NCR provided important information on outbreak characteristics and corroborated with trends seen nationally within syndromic and traditional sources [21]. Specifically, comparing the health indicator data across 2008 and 2009 afforded the following observations in 2009: 1) H1N1 sustained through the spring and summer months unlike seasonal flu; 2) H1N1 spread rapidly through the community; and 3) H1N1 disproportionately affected the younger age groups. As was seen during the 2009 outbreaks of H1N1 in the NCR, information garnered from systems such as ESSENCE supported by a strong data-sharing architecture, serves a critical role in comprehensive disease surveillance. Such information can assist with narrowing focus to regions most in need of vaccines, triage clinics, disease prevention/management education, and other resources. The full potential of these systems may be realized by continuous monitoring and proper application by the end users.

Acknowledgements

This work was made possible through grant number 7UASI137-01 from the Department of Homeland Security's (DHS) Urban Area Security Initiative. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of DHS. The authors would like to acknowledge the exceptional cooperation of the Maryland Department of Health and Mental Hygiene (DHMH), the Virginia Department of Health, Division of Surveillance and Investigation (VDH/DSI), District of Columbia Department of Health, Bureau of Epidemiology and Health Risk Assessment (DCDOH), the eight local health departments represented in the NCR Enhanced Surveillance Operating Group (ESOG), and the Metropolitan Washington Council of Governments' (MWCOG) Health Officials Committee (HOC).

References

[1] Lombardo JS, Buckeridge DL, editors. Disease Surveillance: A Public Health Informatics Approach. Hoboken (NJ): John Wiley & Sons, Inc.; 2007.

[2] Ostroff SM, Kozarsky P. Emerging Infectious Diseases and Travel Medicine. Infect Dis Clin North Am. 1986;231-241.

[3] Lombardo JS. The ESSENCE II Disease Surveillance Test Bed for the National Capital Area. Johns Hopkins APL Tech Dig. 2003;24(4):327-334.

[4] Centers for Disease Control and Prevention. Update: Influenza Activity—United States, August 30, 2009-January 9, 2010. MMWR. 59(02):38-43. [cited 2010 Feb 9]. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5902a3.htm.

[5] World Health Organization. World Now at the Start of 2009 Influenza Pandemic. 2009 Jun 11 [cited 2010 Feb 10]. http://www.who.int/mediacentre/news/statements/2009/h1n1_pandemic_phase6_20090611/en/index.html.

[6] Sniegoski CA. Free Text Chief Complaint Text Processing. Laurel, MD: The Johns Hopkins University Applied Physics Laboratory; August 9, 2002. Internal Report STJ-02-009.

[7] Lombardo J, Burkom H, Elbert E, Magruder S, Lewis SH, Loschen W, et al. A Systems Overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). J Urban Health. 2003;80(2):i32-i42.

[8] Centers for Disease Control and Prevention. 2009 H1N1 Flu ("Swine Flu") and You. 2009 Sep 28 [cited on 2009 Sep 28]. http://www.cdc.gov/h1n1flu/qa.htm.

[9] Hogan WR, Wallstrom GL, Li R. Identifying Categories of Over-the-Counter Products with Superior Outbreak Detection Performance. The RODS Laboratory, Department of Biomedical Informatics, University of Pittsburgh. [cited 2010 Feb 12]. http://thci.org/_documents/temp/Hogan-IdentifyingCategories.doc.

[10] Centers for Disease Control and Prevention. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April-December 12, 2009. [cited 2010 Feb 9]. http://www.cdc.gov/h1n1flu/estimates_2009_h1n1.htm.

[11] Centers for Disease Control and Prevention. Update: Influenza Activity—United States, April-August 2009. [cited 2010 Feb 9]. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5836a6.html.

[12] Centers for Disease Control and Prevention. Update: Influenza Activity—United States, August 30, 2009-January 9, 2010. MMWR. 59(02):38-43. [cited 2010 Feb 9]. http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5902a3.htm.

[13] H1N1 Matches Seasonal Flu Peak Months Early: CDC. [cited 2010 Feb 12]. http://reuters.com/article/idUSTRE59J58H20091012.

[14] Centers for Disease Control and Prevention. 2009-2010 Influenza Season Week 4 ending January 30, 2010. [cited 2010 Feb 9]. http://www.cdc.gov/flu/weekly/.

[15] Centers for Disease Control and Prevention. 2009 H1N1 Flu ("Swine Flu") and You. [cited 2010 Feb 12]. http://www.cdc.gov/h1n1flu/qa.htm.

[16] Centers for Disease Control and Prevention. CDC Estimates of 2009 H1N1 Influenza Cases, Hospitalizations and Deaths in the United States, April 2009-January 16, 2010. [cited 2010 Feb 12]. http://www.cdc.gov/h1n1flu/estimates_2009_h1n1.htm.

[17] Centers for Disease Control and Prevention. Vaccine Against 2009 H1N1 Influenza Virus. [cited 2010 Feb 12]. http://www.cdc.gov/h1n1flu/vaccination/public/vaccination_qa_pub.htm.

[18] Centers for Disease Control and Prevention. Surveillance for Influenza and Pneumonia-Associated Hospitalizations and Deaths for the 2009-2010 Season. [cited 2010 Feb 12]. http://www.cdc.gov/h1n1flu/reportingqa.htm.

[19] World Health Organization. Influenza-like illness in the United States and Mexico. 2009 Apr 24 [cited 2010 Feb 12]. http://www.who.int/csr/don/2009_04_24/en/index.html.

[20] 2009 New York City Department of Health and Mental Hygiene Health Alert #15: H1N1 (Swine Origin) Update. 2009 May 1 [cited 2010 Feb 15]. www.nyc.gov/html/doh/downloads/pdf/cd/2009/09md15.pdf.

[21] Centers for Disease Control and Prevention. 2009 H1N1 Flu U.S. Situation Update. [cited 2010 Feb 15]. http://www.cdc.gov/h1n1flu/updates/us/.