Correlation between Real-Time BioSense Influenza Indicators and Data from the U.S. Influenza Sentinel Physicians Surveillance Network

Michelle N. Podgornik, M.P.H., Roseanne English, B.S., and Jerome I. Tokars, M.D., M.P.H.

Centers for Disease Control and Prevention, National Center for Public Health Informatics, Division of Emergency Preparedness and Response, Atlanta, GA 30333

Objective

The objective of this study was to determine which chief complaints and ICD-9-CM coded diagnoses from real-time BioSense hospital data correlate well with data from conventional influenza surveillance systems.

Background

Surveillance of influenza activity in the United States has traditionally centered on data gathered through the U.S. Influenza Surveillance System, which is maintained by the Centers for Disease Control and Prevention (CDC). A project to co-display these data with those data from the BioSense System was piloted during the 2006-2007 influenza season. These data and displays will be made available to users outside of CDC through the BioSense application Influenza Module in the 2007-2008 season. This module aims to improve the ability of public health to monitor influenza activity across the nation by summarizing major data sources in a unified, user-friendly format.

Methods

The study period was March 2006 to May 2007. During this period, the number of hospitals reporting data to BioSense increased from 30 to more than 360; most report emergency department chief complaint data only. We evaluated seven chief complaint or diagnosis indicators (i.e., bronchitis, cough, fever, influenza, otitis media, pneumonia, and upper respiratory infections) to determine their correlation with Sentinel Providers data. For chief complaints, these indicators were identified by text parsing for appropriate keywords (e.g., "flu" as a keyword for influenza). For final diagnoses, the indicators were identified by appropriate ICD-9-CM codes (e.g., 487 for influenza). These indicators were evaluated for three patient classes (emergency department patients (ED), inpatients, and outpatients) and two data types (chief complaint and final diagnosis). We calculated correlation coefficients between the percent of visits with the identified indicators in the BioSense data and the percent influenza-like illness (ILI) in the Sentinel Providers data. The data were aggregated by week at the national level.

Results

For ED chief complaint data, the percent of visits with the indicator varied from 5.6% for fever to 0.32% for influenza; strong correlations with Sentinel Provider ILI% were found for influenza ($r = 0.92$), fever ($r = 0.83$), and cough ($r = 0.82$). For ED diagnosis data, the indicators with the strongest correlations were influenza ($r = 0.92$), bronchitis ($r = 0.90$), and pneumonia ($r = 0.87$). For inpatient diagnosis data, the indicators with the strongest correlations were influenza ($r = 0.92$), bronchitis ($r = 0.86$), and otitis media ($r = 0.79$).

Conclusion

These analyses suggest that several indicators from BioSense hospital data correlate well with Sentinel Providers data and are likely to be useful for tracking influenza activity. However, the analyses must be repeated after stratifying by patient age and geographic area and adjusting for differing numbers of hospitals sending data during the study period. When selecting optimal indicators of influenza activity, it may also be useful to consider the latency of the data (chief complaints are reported in 1-2 days, while final diagnoses are reported in 1-4 weeks) as well as how commonly the indicator is reported (e.g., chief complaints of cough are much more common than those of influenza). In addition, since a larger percentage of hospitals report data for ED chief complaints than for any other data types, these may be the most useful BioSense data. Therefore, due to the strong correlations observed for cough, fever, and influenza, these three indicators will be included in the Influenza Module.