Artificial Intelligence in Medicine 59 (2013) 169–174

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

Comparing the accuracy of syndrome surveillance systems in detecting influenza-like illness: GUARDIAN vs. RODS vs. electronic medical record reports

Julio C. Silva[a], Shital C. Shah[a], Dino P. Rumoro[a], Jamil D. Bayram[b], Marilyn M. Hallock[a], Gillian S. Gibbs[a,*], Michael J. Waddell[c]

[a] Rush University, 1653 West Congress Parkway, Chicago, IL 60612, USA
[b] Johns Hopkins School of Medicine, 5801 Smith Avenue, Davis Building Suite 3220, Baltimore, MD 21209, USA
[c] Pangaea Information Technologies, 219 West Chicago Avenue, 2nd Floor, Chicago, IL 60654, USA

ARTICLE INFO

Article history:
Received 15 April 2013
Received in revised form
19 September 2013
Accepted 24 September 2013

Keywords:
Biosurveillance
Syndromic surveillance systems
Influenza-like illness
Public health informatics
GUARDIAN
RODS

ABSTRACT

Background: A highly sensitive real-time syndrome surveillance system is critical to detect, monitor, and control infectious disease outbreaks, such as influenza. Direct comparisons of diagnostic accuracy of various surveillance systems are scarce.
Objective: To statistically compare sensitivity and specificity of multiple proprietary and open source syndrome surveillance systems to detect influenza-like illness (ILI).
Methods: A retrospective, cross-sectional study was conducted utilizing data from 1122 patients seen during November 1–7, 2009 in the emergency department of a single urban academic medical center. The study compared the Geographic Utilization of Artificial Intelligence in Real-time for Disease Identification and Alert Notification (GUARDIAN) system to the Complaint Coder (CoCo) of the Real-time Outbreak Detection System (RODS), the Symptom Coder (SyCo) of RODS, and to a standardized report generated via a proprietary electronic medical record (EMR) system. Sensitivity, specificity, and accuracy of each classifier's ability to identify ILI cases were calculated and compared to a manual review by a board-certified emergency physician. Chi-square and McNemar's tests were used to evaluate the statistical difference between the various surveillance systems.
Results: The performance of GUARDIAN in detecting ILI in terms of sensitivity, specificity, and accuracy, as compared to a physician chart review, was 95.5%, 97.6%, and 97.1%, respectively. The EMR-generated reports were the next best system at identifying disease activity with a sensitivity, specificity, and accuracy of 36.7%, 99.3%, and 83.2%, respectively. RODS (CoCo and SyCo) had similar sensitivity (35.3%) but slightly different specificity (CoCo = 98.9%; SyCo = 99.3%). The GUARDIAN surveillance system with its multiple data sources performed significantly better compared to CoCo ($\chi^2 = 130.6$, $p < 0.05$), SyCo ($\chi^2 = 125.2$, $p < 0.05$), and EMR-based reports ($\chi^2 = 121.3$, $p < 0.05$). In addition, similar significant improvements in the accuracy (>12%) and sensitivity (>47%) were observed for GUARDIAN with only chief complaint data as compared to RODS (CoCo and SyCo) and EMR-based reports.
Conclusion: In our study population, the GUARDIAN surveillance system, with its ability to utilize multiple data sources from patient encounters and real-time automaticity, demonstrated a more robust performance when compared to standard EMR-based reports and the RODS systems in detecting ILI. More large-scale studies are needed to validate the study findings, and to compare the performance of GUARDIAN in detecting other infectious diseases.

1. Background

The principle goal of syndrome surveillance is early recognition of disease trends, allowing for advanced mobilization of public health and medical resources, thereby reducing transmissibility and subsequently mitigating morbidity and mortality [1]. Surveillance is critical for identifying and controlling infectious disease outbreaks, such as influenza [2]. For surveillance of influenza-like

* Corresponding author at: Department of Emergency Medicine, Rush University, 1653 West Congress Parkway, Suite 108 Kellogg, Chicago, IL 60612, USA.
Tel.: +1 312 563 3820; fax: +1 312 942 4021.
E-mail address: Gillian_Gibbs@rush.edu (G.S. Gibbs).

0933-3657/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.artmed.2013.09.001

illness (ILI), public health departments receive data from multiple sources, including outpatient clinics and emergency departments.

Healthcare facilities utilize a variety of reporting mechanisms, ranging from individual staff manually reviewing paper-based patient records to the generation of standard reports from electronic medical records (EMRs). Miller et al. [3] reports that typical disease surveillance systems such as EMR-based reports are too slow and insensitive to detect large-scale disease outbreaks. In order to fill this gap, more sophisticated electronic syndrome surveillance systems have been developed at the national level, such as: BioSense [4], and Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE) [5], and at the state and local level such as Real-time Outbreak Detection System (RODS) [6], North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) [7], New York City Department of Health and Mental Hygiene Syndromic surveillance system [8], Public Health – Seattle and King County Syndromic surveillance system [9], and Geographic Utilization of Artificial Intelligence in Real-Time for Disease Identification and Alert Notification (GUARDIAN) [10]. While some surveillance systems, such as RODS, are freely available as open-source projects, most are closed-source projects which are only available for use by public health departments and are not made available to individual hospitals or researchers.

The Centers for Disease Control and Prevention (CDC) published guidelines for evaluating public health surveillance systems for early detection of outbreaks, which include system description, timely and valid outbreak detection, system experience measures, and functionality. These guidelines provide conclusions and recommendations for use and improvement of systems for early outbreak detection [11]. A surveillance system with high sensitivity and specificity ideally should be able to balance the need for early detection of outbreaks and the cost associated with unnecessary responses to false alarms. Although many disease surveillance systems are currently available and evaluated in the published literature, studies with head-to-head comparisons of these various systems are scant [12–22].

The goal of this study was to compare the accuracy, based on sensitivity and specificity, of a few selected disease surveillance systems. The influenza-like illness (ILI)[1] case definition used by the CDC (fever >100 °F [>37.8 °C] plus cough and/or sore throat in the absence of other confirmed diagnoses) [28] was selected for comparing GUARDIAN to RODS [Complaint Coder (CoCo) Batch and Symptom Coder (SyCo)], and to standard EMR-based reports. A brief description of each of the surveillance systems utilized in this article is provided in the methods section. RODS surveillance system was chosen due to its open-source code and standard EMR-based reports were chosen due to their ubiquitous use. Other systems, such as BioSense and ESSENCE, were not included in the comparison due to the inability of the authors to obtain access to those systems for research purposes, due to their proprietary and restricted nature.

[1] Importance of influenza surveillance: the World Health Organization (WHO) estimates that every year there are between three and five million severe cases of influenza causing 250,000–500,000 deaths worldwide [23]. The Centers for Disease Control and Prevention (CDC) estimate that between April 1, 2009 and March 13, 2010, the United States had between 43 and 88 million cases of the 2009 H1N1 influenza strain, leading to 12,270 deaths and 270,000 hospitalizations [24]. The severity and pervasiveness of influenza throughout the world results in a significant economic and social burden. In a typical flu season, the years of life lost for a mean age of 75.7 is 594,000. During the 2009 H1N1 pandemic flu season it was a staggering 1.9 million years of life lost [25,26]. The estimated direct hospital impact of an annual influenza burden is estimated to be 3.1 million hospitalized days, at a cost of $6 billion [27].

2. Methods

A retrospective cross-sectional study was conducted utilizing data from 1122 patients evaluated between November 1, 2009 and November 7, 2009 in the emergency department of an urban tertiary academic medical center. During the week of November 1–7, 2009 (part of the 2009 H1N1 influenza outbreak), the medical center saw a significant spike in influenza as well as influenza-like illness cases. We evaluated the following five existing surveillance systems and their appropriate subcomponents: GUARDIAN, GUARDIAN with chief complaint only, EMR-based reports from a commercial proprietary system, CoCo (RODS), and SyCo (RODS).

2.1. RODS (CoCo & SyCo)

Real-time Outbreak and Disease Surveillance (RODS) is open-source software, developed in 1999, with the intention to collect and analyze disease surveillance data in real-time, for public health purposes. RODS uses pluggable text classifiers, e.g., Complaint Coder (CoCo) Batch and Symptom Coder (SyCo) [6]. Hospitals send real-time data to RODS using Health Level 7 (HL7) message protocols. RODS automatically classifies the registration chief complaint, using Bayesian classifiers, into one of seven syndrome categories with CoCo and into one of 89 categories using SyCo. If the chief complaint does not fit into any of these categories, it is classified as "Other". RODS stores the data in a relational database, aggregates the data for analysis using data warehousing techniques, applies univariate and multivariate statistical detection algorithms to the data, and alerts users when an anomaly in syndrome counts is detected [29].

2.2. Electronic medical records

Electronic medical record (EMR)-based surveillance is the most common method utilized to report surveillance data. EMR-based surveillance utilizes standardized reports generated directly through a hospital's EMR system that captures some data elements such as chief complaints, or discreet data such as International Classification of Diseases (ICD) codes. The chief complaints are the most widely used ED data element for syndromic surveillance [30–32]. Depending on the sophistication of the reporting mechanisms, a clinical full-time employee manually reviews either a database of patient records or a standardized report generated based on keyword matches. In our study, we used a common proprietary commercial EMR system.

2.3. Geographic Utilization of Artificial Intelligence in Real-Time for Disease Identification and Alert Notification (GUARDIAN)

The GUARDIAN surveillance system is funded by the United States Department of Defense and is a propriety beta research project currently being deployed and tested at various hospitals in the Chicago metropolitan area. Upon successful implementation, the GUARDIAN surveillance system will be commercially available for both hospitals and public health departments. The GUARDIAN system receives patient data from the existing HL7 interfaces used to communicate throughout a hospital's existing systems. The GUARDIAN system does not enforce any requirements on the incoming data beyond that of the underlying HL7 standard (e.g., the Public Health Information Network messaging standard [33]) in order to make the integration of GUARDIAN with existing systems as seamless as possible. All incoming patient data is parsed, cross-linked to other data related to the same patient, and stored in a relational database management system. GUARDIAN then applies Natural Language Processing (NLP) algorithms to extract known