

Validation of Syndromic ILI Data for Use in CDC's ILINet Surveillance, Pennsylvania

Sameh W. Bektor*¹, Kirsten Waller¹, Lenee Blanton² and Krista Kniss²

¹Epidemiology, Pennsylvania Department of Health, Harrisburg, PA, USA; ²Centers for Disease Control and Prevention, Atlanta, GA, USA

Objective

Discuss use of syndromic surveillance as a source for the state's ILI/Influenza surveillance

Discuss reliability of syndromic data and methods to address problems caused by data outliers and inconsistencies.

Introduction

ILINet is a CDC program that has been used for years for influenza-like illness (ILI) surveillance, using a network of outpatient providers who volunteer to track and report weekly the number of visits due to ILI and the total number of visits to their practice. Pennsylvania has a network of 95 providers and urgent care clinics that submit data to ILINet. However, ongoing challenges in recruiting and retaining providers, and inconsistent weekly reporting are barriers to receiving accurate, representative, and timely ILI surveillance data year-round. Syndromic surveillance data have been used to enhance outpatient ILI surveillance in a number of jurisdictions, including Pennsylvania. At present, 156 hospitals, or 90% of all Pennsylvania hospitals with emergency departments (EDs), send chief complaint and other information on their ED visits to the Department of Health's (PADOH) syndromic surveillance system. PADOH evaluated the consistency and reliability of ILI syndromic data as compared to ILINet data, to confirm that syndromic data were suitable for use in ILINet.

Methods

Pennsylvania ILINet data from the past 6 influenza seasons (2011-2012 to 2016-2017, or 314 weeks of data) were downloaded from the CDC's ILINet website. The statewide weekly percent of visits due to ILI in ILINet was used as the standard for comparisons. For syndromic surveillance, PADOH uses the Epicenter platform hosted by Health Monitoring Systems (HMS); visit-level data are also stored in SAS datasets at PADOH, and HMS forwards a subset of data to the National Syndromic Surveillance System Program. Using syndromic data from the same time period, the proportion of weeks with no syndromic data available was calculated for each facility. A state-developed ILI algorithm (very similar to the 2016 algorithm developed by the ISDS Syndrome Definitions Workgroup) was applied to ED visit chief complaint data to identify visits likely to be due to ILI. The algorithm flags the ER visit as ILI if chief complaint has any combinations of words for flu or fever plus either cough and sore throat or fever and both cough or sore throat. The percent of ED visits due to ILI per the syndromic algorithm (ILIsyn) was calculated for each week by hospital and state-wide. Facility ILIsyn trends were compared to the State level percent ILI data from ILINet by visually examining plots and by calculating Pearson correlation coefficients. Facilities that had ≥ 15 weeks where ILIsyn differed from percent ILI in ILINet by more than 5% were considered to be poorly correlated.

Results

A total of 156 hospitals were evaluated in the study. Twenty of the hospitals were excluded because they did not have syndromic data for at least 50% of the weeks in the study period, and an additional 20 were excluded because they had not agreed to have data forwarded to

CDC. Of the remaining 116 facilities, individual facility correlation coefficients between ILIsyn and ILINet trends ranged from 0.03 to 0.82 (examples are in Figure 1). Twenty-four hospitals (20.7%) were determined to be poorly correlated. When data from the remaining 92 hospitals were combined, the state ILINet and state-wide ILIsyn trends were strongly correlated statistically and graphically ($r=0.82$, $p < 0.0001$, Figure 2). Syndromic data from these 92 facilities were deemed acceptable for inclusion in ILINet.

Conclusions

Syndromic surveillance data are a valuable source for ILI surveillance. However, evaluation at the hospital-specific level revealed that useful information is not obtained from all facilities. This project demonstrated that validation of data at the facility level is crucial to obtaining reliable and meaningful information. More work is needed to understand which factors distinguish well-correlated from poorly-correlated facilities, and how to improve the quality of information obtained from poorly-correlated facilities.

Figure 1: Comparison of weekly percent of visits due to ILI in Pennsylvania ILINet and emergency department syndromic data, from a well-correlated hospital (A) and a poorly-correlated hospital (B).

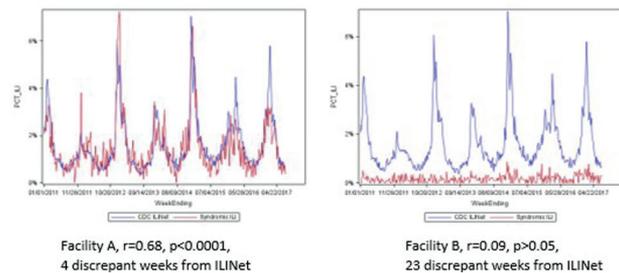
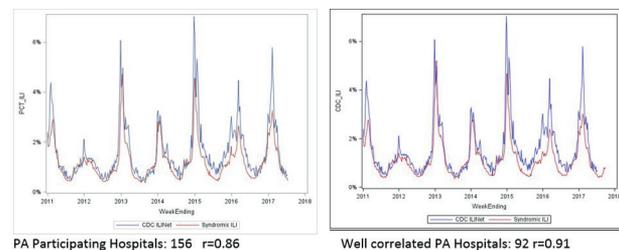


Figure 2: Percent of visits due to ILI, from Pennsylvania ILINet data and syndromic emergency department data from 92 well-correlated Pennsylvania hospitals.



Keywords

Influenza Like Illness; Syndromic Surveillance; data validation methods

Acknowledgments

Jonah M. Long, Pennsylvania Department of Health

*Sameh W. Bektor

E-mail: sbektor@pa.gov

