# Classification of Emergency Department Syndromic Data for Seasonal Influenza Surveillance

**Atar Baer[1], PhD, Jacqueline Coberly[3], PhD, Lang Hung[3], MPH, Howard Burkom[3], PhD, Wayne Loschen,[3] MS, Joseph Lombardo[3], MS, Jeff Duchin[1,2], MD**

*Public Health – Seattle & King County (PHSKC) [1], University of Washington Department of Epidemiology [2], Johns Hopkins University – Applied Physics Laboratory [3]*

## OBJECTIVE

We evaluated several classifications of emergency department (ED) syndromic data to ascertain best syndrome classifications for ILI.

## BACKGROUND

There are no agreed-upon standards for defining influenza-like illness (ILI) using syndromic surveillance data. Some definitions restrict classification of ED chief complaints to specific mention of "flu", whereas others include "febrile illness" as a proxy for influenza. Furthermore, some systems define ILI based on chief complaint alone, whereas others incorporate diagnostic data into the definition. It is unknown how well these different classifications describe the "true" influenza season, or whether algorithm performance is definition-specific. Also, best practices for classification of syndromic data for tracking ILI may differ regionally, so previously published study results may not be generalizable.

## METHODS

The PHSKC syndromic surveillance system receives chief complaint data from 19 emergency departments (EDs) in King County. Diagnoses are provided by a subset of hospitals. Data for this evaluation were restricted to 10/7/2006 through 4/28/2007 and excluded one hospital that did not provide ED chief complaints. The weekly number of positive viral cultures submitted by influenza sentinel physicians ("ISP") and the weekly number of positive influenza rapid antigen test results ("RA") as reported by several local microbiology laboratories were used as the gold standard for comparison with ED data. We used 6 classifications of ILI ED visit data for this analysis: (1) A chief complaint with specific mention of flu or ICD-9 flu codes ("CCFLU"); (2) definition #1, but applied to either the chief complaint or diagnosis field ("ANYFLU") (3) a chief complaint with specific mention of febrile complaints or ICD-9 fever codes ("CCFEVER") (4) definition #3, but applied to either the chief complaint or diagnosis field ("ANYFEVER"); (5) definition #1, OR Definition #2 and a chief complaint with specific mention of cough, OR Definition #2 and a chief complaint with specific mention of sore throat, OR a chief complaint with mention of terms or ICD codes for sepsis, bronchiolitis, bacteremia, or pneumonia ("CCILI"); (6) definition #5, but applied to either the chief complaint or diagnosis field ("ANYILI"). We calculated SNR's[1] to compare the level of signals achieved by the ILI classifications with the level of background noise; the "true" signal was defined on the basis of ISP culture-positives. The association between positive viral specimens and other data sources was examined through correlation analysis; correlation coefficients were also calculated by lagging the ISP and RA positive specimens forward and backward in time by 1-week increments to examine the timeliness of the signals[1]. The number, timing, and concordance of alerts were compared for the 6 ILI classifications using several different algorithms (regression, EWMA, Poisson, and C1-C3).

## RESULTS

The SNR was highest for the "CCFLU" and "ANYFLU" categories; including data from the diagnosis, where available, did not consistently improve the SNR. The correlation with ED data was better for RA than for ISP results, and was highest for the "CCFEVER" and "ANYFEVER" categories. The ED data lagged behind the ISP signal by about 2-3 weeks, and tended to lag behind the RA signal by 1 week, with the exception of the "CCFEVER" and "ANYFEVER" categories, which did not lag behind RA. Using an algorithm that switches automatically between regression, EWMA, and Poisson depending on the data characteristics, the broader definitions of "ANYFLU", "ANYFEVER", and "ANYILI" had similar numbers of alerts within the defined flu season and alerted on the same days more than 80% of the time.

| | | ISP | | RA | |
|---|---|---|---|---|---|
| | SNR | Corr. | P-value | Corr. | P-value |
| **ANYFEVER** | 29.5 | 0.69 | <.0001 | 0.83 | <.0001 |
| **ANYFLU** | 48.1 | 0.55 | 0.0018 | 0.72 | <.0001 |
| **ANYILI** | 31.2 | 0.61 | 0.0003 | 0.76 | <.0001 |
| **CCFEVER** | 30.3 | 0.69 | <.0001 | 0.83 | <.0001 |
| **CCFLU** | 40.1 | 0.54 | 0.0023 | 0.69 | <.0001 |
| **CCILI** | 32.9 | 0.65 | 0.0001 | 0.77 | <.0001 |

## CONCLUSIONS

The febrile syndrome classification had the closest temporal correlation with the laboratory surveillance data, while the flu classification had the highest SNR; the use of diagnostic coding did not consistently improve these measures compared with using chief complaint alone. For classifications that were based on either the chief complaint or diagnosis fields, alerting characteristics were equally good.

Reference: (1) Marsden-Haug N, Foster VB, Gould PL, Elbert E, Wang H, Pavlin JA. Code-based syndromic surveillance for influenzalike illness by international classification of diseases, ninth revision. Emerging Infectious Diseases 13(2): 207-216.

Further Information: Atar Baer, atar.baer@kingcounty.gov