

# Expanding the Functionality of Syndromic Surveillance Systems: Data Mining and Query Development

Elizabeth Hibler, M.P.H., Dave Trepanier, M.S., Michael Wade, M.P.H.

*Epidemiology Resource Center, Indiana State Department of Health*

## OBJECTIVE

This paper describes the development of targeted query tools and processes designed to maximize the extraction of information from, and improve the quality of, the hospital emergency department chief complaint data stream utilized by the Indiana State Department of Health (ISDH) for syndromic surveillance.

## BACKGROUND

The Indiana Public Health Emergency Surveillance System (PHESS) currently receives approximately 5,000 near real-time chief complaint messages from 55 hospital emergency departments daily. The ISDH partners with the Regenstrief Institute to process, batch, and transmit data every three hours. The Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE) tool is utilized to analyze these chief complaint data and visualize generated alerts.<sup>1</sup>

The ISDH syndromic surveillance team discovered that certain chief complaints of interest were coded into the "other" syndrome and not visible in typical daily alert data. Staff determined that even a single chief complaint containing keywords related to specific reportable diseases could be of significant public health value and should be made available to investigating epidemiologists.<sup>2</sup>

In addition, data quality is critical to the success of the program and must be evaluated to ensure optimal system performance. Metrics related to data flow and completeness were identified to serve as indicators of hospital connectivity or coding problems. These measures included the percent change in daily admits and the proportion of chief complaints missing the patient address.

## METHODS

Initially, epidemiologists used ESSENCE to develop queries searching for chief complaints of interest. The process was inefficient, however, because a separate query was necessary for each reportable condition and could only be saved using the "bookmark" feature in ESSENCE. An automated and less labor intensive method was desired. A batch file was developed to execute a SQL query script and email the results in html format to staff. For this initial report, the queries were written to search the inbound PHESS database, which contained the raw data stream.

The queries were later modified to search the ESSENCE database in order to take advantage of the

chief complaint parser. This improved efficiency by allowing for more precise filtering of chief complaints with unrelated keywords. The SQL queries were then re-coded using the SAS system, which allowed for enhanced data management and analysis. Currently, the SAS code is used to produce reports in both EXCEL and PDF formats. All batch files are executed daily as a Windows Scheduled Task.

## RESULTS

This recently implemented process has already documented improvement. Originally, the report indicated 23 facilities sending more than 10% of chief complaints with a blank address field. This issue was quickly resolved for 21 facilities and currently only one hospital group is missing greater than 10% of the address data. In addition, an example of qualitative success was use of the query to identify the chief complaint "exposure to meningitis," which facilitated contact tracing and provision of prophylaxis to individuals at risk from exposure to bacterial meningitis. These data were also used to initiate follow-up on mumps-related chief complaints and to ensure that an outbreak was not occurring in Indiana.

## CONCLUSIONS

A future challenge for sustaining the PHESS is developing innovative methods to enhance the system's utility. The queries developed to produce the daily PHESS report add much value by providing users with otherwise hidden data of interest and a monitoring tool to help improve system reliability. At ISDH, customized queries have been developed and distributed to specific program areas such as rabies, tuberculosis, trauma, and asthma. In addition, the Marion County Health Department, located in the county with the state's greatest population density, currently receives a custom report of chief complaints of interest from hospitals within the county. Although evaluation is ongoing, the process of query development and automation has already broadened the application of the chief complaint data and increased effectiveness of the PHESS.

## REFERENCES

[1] Lombardo, JS, The ESSENCE II Disease Surveillance Test Bed for the National Capital Area. Johns Hopkins APL Technical Digest, vol 24(4):2003:327-34.

[2] 410 IAC 1-2.3.

Further Information:

Michael Wade, [mwade@isdh.IN.gov](mailto:mwade@isdh.IN.gov)