Original Paper

Automated Real-Time Collection of Pathogen-Specific Diagnostic Data: Syndromic Infectious Disease Epidemiology

Lindsay Meyers¹, BS; Christine C Ginocchio^{1,2,3}, PhD; Aimie N Faucett¹, MS; Frederick S Nolte⁴, PhD; Per H Gesteland⁵, MD, MS; Amy Leber⁶, PhD; Diane Janowiak⁷, BS, MT (ASCP); Virginia Donovan⁸, MD; Jennifer Dien Bard^{9,10}, PhD, D(ABMM); Silvia Spitzer¹¹, PhD; Kathleen A Stellrecht¹², PhD; Hossein Salimnia¹³, PhD; Rangaraj Selvarangan¹⁴, BVSc, PhD, D(ABMM); Stefan Juretschko¹⁵, PhD; Judy A Daly¹⁶, PhD; Jeremy C Wallentine¹⁷, MD; Kristy Lindsey¹⁸, BSc (Health); Franklin Moore¹⁸, MD; Sharon L Reed¹⁹, MD; Maria Aguero-Rosenfeld²⁰, MD; Paul D Fey²¹, BS, PhD, D(ABMM); Gregory A Storch²², MD; Steve J Melnick²³, MD, PhD; Christine C Robinson²⁴, PhD; Jennifer F Meredith²⁵, PhD; Camille V Cook¹, BS; Robert K Nelson¹, BS; Jay D Jones¹, MS; Samuel V Scarpino²⁶, PhD; Benjamin M Althouse^{27,28}, ScM, PhD; Kirk M Ririe²⁹, BS; Bradley A Malin³⁰, PhD; Mark A Poritz³¹, PhD

- ⁵Departments of Pediatrics and Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, UT, United States
- ⁶Laboratory of Microbiology and Immunoserology, Department of Laboratory Medicine, Nationwide Children's Hospital, Columbus, OH, United States
- ⁷Department of Lab Operations, South Bend Medical Foundation, South Bend, IN, United States
- ⁸Department of Pathology, New York University Winthrop Hospital, Mineola, NY, United States
- ⁹Clinical Microbiology and Virology Laboratory, Department of Pathology and Laboratory Medicine, Children's Hospital of Los Angeles, Los Angeles, CA, United States
- ¹⁰Keck School of Medicine, University of Southern California, Los Angeles, CA, United States
- ¹¹Molecular Genetics Laboratory, Stony Brook University Medical Center, Stony Brook, NY, United States
- ¹²Department of Pathology and Laboratory Medicine, Albany Medical Center, Albany, NY, United States
- ¹³Department of Pathology, Wayne State University School of Medicine, Detroit, MI, United States
- ¹⁴Clinical Microbiology, Virology and Molecular Infectious Diseases Laboratory, Department of Pathology and Laboratory Medicine, Children's Mercy Hospital, Kansas City, MO, United States
- ¹⁵Department of Pathology and Laboratory Medicine, Division of Infectious Disease Diagnostics, Northwell Health, Lake Success, NY, United States
- ¹⁶Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT, United States
- ¹⁷Department of Pathology, Intermountain Medical Center, Murray, UT, United States
- ¹⁸Laboratory of Microbiology, University of Massachusetts Medical School-Baystate, Springfield, MA, United States
- ¹⁹Department of Pathology and Medicine, Divisions of Clinical Pathology and Infectious Diseases, UC San Diego, San Diego, CA, United States
- ²⁰Department of Clinical Laboratories, New York University Langone Health, New York, NY, United States
- ²¹Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, United States
- ²²Department of Pediatrics, Washington University, St. Louis, MO, United States
- ²³Department of Pathology and Clinical Laboratories, Nicklaus Children's Hospital, Miami, FL, United States
- ²⁴Department of Pathology and Laboratory Medicine, Microbiology/Virology Laboratory Section, Children's Hospital Colorado, Aurora, CO, United States
- ²⁵Department of Laboratory Services, Microbiology Section, Greenville Health System, Greenville, SC, United States
- ²⁶Northeastern University, Boston, MA, United States
- ²⁷University of Washington, Seattle, WA, United States
- ²⁸New Mexico State University, Las Cruces, NM, United States
- ²⁹bioMérieux, Salt Lake City, UT, United States
- ³⁰Department of Biomedical Informatics, School of Medicine, Vanderbilt University, Nashville, TN, United States
- ³¹BioFire Defense, Salt Lake City, UT, United States

Corresponding Author:

http://publichealth.jmir.org/2018/3/e59/



¹BioFire Diagnostics, Salt Lake City, UT, United States

²bioMérieux USA, Durham, NC, United States

³Hofstra Northwell School of Medicine, Hempstead, NY, United States

⁴Department of Pathology and Laboratory Medicine, Medical University of South Carolina, Charleston, SC, United States

515 Colorow Drive Salt Lake City, UT, 84108 United States Phone: 1 8017366354 ext 365 Fax: 1 8015880507 Email: <u>lindsay.meyers@biofiredx.com</u>

Abstract

Background: Health care and public health professionals rely on accurate, real-time monitoring of infectious diseases for outbreak preparedness and response. Early detection of outbreaks is improved by systems that are comprehensive and specific with respect to the pathogen but are rapid in reporting the data. It has proven difficult to implement these requirements on a large scale while maintaining patient privacy.

Objective: The aim of this study was to demonstrate the automated export, aggregation, and analysis of infectious disease diagnostic test results from clinical laboratories across the United States in a manner that protects patient confidentiality. We hypothesized that such a system could aid in monitoring the seasonal occurrence of respiratory pathogens and may have advantages with regard to scope and ease of reporting compared with existing surveillance systems.

Methods: We describe a system, BioFire Syndromic Trends, for rapid disease reporting that is syndrome-based but pathogen-specific. Deidentified patient test results from the BioFire FilmArray multiplex molecular diagnostic system are sent directly to a cloud database. Summaries of these data are displayed in near real time on the Syndromic Trends public website. We studied this dataset for the prevalence, seasonality, and coinfections of the 20 respiratory pathogens detected in over 362,000 patient samples acquired as a standard-of-care testing over the last 4 years from 20 clinical laboratories in the United States.

Results: The majority of pathogens show influenza-like seasonality, rhinovirus has fall and spring peaks, and adenovirus and the bacterial pathogens show constant detection over the year. The dataset can also be considered in an ecological framework; the viruses and bacteria detected by this test are parasites of a host (the human patient). Interestingly, the rate of pathogen codetections, on average 7.94% (28,741/362,101), matches predictions based on the relative abundance of organisms present.

Conclusions: Syndromic Trends preserves patient privacy by removing or obfuscating patient identifiers while still collecting much useful information about the bacterial and viral pathogens that they harbor. Test results are uploaded to the database within a few hours of completion compared with delays of up to 10 days for other diagnostic-based reporting systems. This work shows that the barriers to establishing epidemiology systems are no longer scientific and technical but rather administrative, involving questions of patient privacy and data ownership. We have demonstrated here that these barriers can be overcome. This first look at the resulting data stream suggests that Syndromic Trends will be able to provide high-resolution analysis of circulating respiratory pathogens and may aid in the detection of new outbreaks.

(JMIR Public Health Surveill 2018;4(3):e59) doi: 10.2196/publichealth.9876

KEYWORDS

epidemiology; patients; privacy; communicable disease; internet; pathology, molecular

Introduction

Surveillance Landscape

p://publichealth.jmir.org/2018/3/e59/

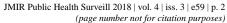
The availability of real-time surveillance data that can monitor the spread of infectious diseases benefits public health [1-3]. At present, tracking of respiratory or foodborne outbreaks relies on a variety of methods ranging from automated real-time electronic reporting to manual Web entry of test results. Systems such as the Centers for Disease Control and Prevention's (CDC) FluView [4], National Respiratory and Enteric Virus Surveillance Systems (NREVSS) [5], National Electronic Disease Surveillance System [6], Global Emerging Infections Surveillance (GEIS) [7], and others, although Web-based, still require manual entry of data from laboratories, resulting in data that are often incomplete or not current.

Syndrome-based surveillance systems [8-10] include BioSense (extraction of symptomatic data from electronic health records

[11]), Google Flu (tracking of internet search queries [12] but recently discontinued [13]), and Flu Near You (voluntary reporting [14]). Additionally, numerous next generation, syndromic surveillance systems, for example, pharmacy sales records [15,16], Twitter conversations [17,18], and Wikipedia hits [19,20] have come online in the past 5 years. However, these systems cannot report the specific pathogen causing an increase in a particular set of symptoms. Finally, there are more localized efforts such as GermWatch in Utah [21] and the Electronic Clinical Laboratory Reporting System (ECLRS) in New York [22] that draw from hospital information systems (HISs) and laboratory information systems (LISs). This disparity in technologies and data collection methods results in incomplete surveillance.

Comprehensive Testing

Comprehensive and uniform diagnostic test data will aid in the identification of potential outbreaks. A combination of broad respiratory pathogen testing and an internal electronic



surveillance system enabled the rapid dissemination of data across the largest health care system in New York, the North Shore-LIJ Health System (now Northwell Health), during the influenza A H1N1-2009 pandemic in the New York City area. Pathogen-specific molecular testing permitted rapid (1) notification to state epidemiologists, (2) tracking of the virus so that health care resources could be managed effectively, and (3) evaluation of influenza diagnostics [23,24]. Today, with the threat of emerging pathogens such as Middle East respiratory syndrome coronavirus (CoV), avian influenza, enterovirus (EV) D68, and Ebola virus, real-time surveillance programs are critical [25,26].

It is not always possible to accurately diagnose the causative agents of most infectious diseases from symptoms alone because of overlapping clinical presentation. Thus, to achieve maximal utility, infectious disease surveillance systems should move beyond syndrome-based reporting and be pathogen-specific and comprehensive, reporting on as many of the common pathogens for a particular syndrome as possible. Sensitive and specific automated molecular diagnostic systems that detect up to 4 different pathogens in a single sample have been available from in vitro diagnostic (IVD) manufacturers for some time [27,28]. However, adoption of IVD platforms with broad multiplexing capability has become widespread only in the last few years. Commercially available systems that can detect most of the known etiological agents for respiratory, gastrointestinal (GI), and other multipathogen syndromes [29-31] include the BioFire (Salt Lake City, UT) FilmArray System ([32]; Multimedia Appendix 1); the GenMark (Carlsbad, CA) eSensor XT-8 [33] and ePlex [34]; and the Luminex (Austin, TX) xTAG [35], nxTag [36], and Verigene systems [37].

Sharing of Patient Data

Multianalyte diagnostic tests provide the raw data needed for real-time pathogen-specific syndromic surveillance, but there remain a number of obstacles to sharing these results (reviewed in [38]). The obstacles largely center on information privacy and network security. A real-time surveillance system using diagnostic test results requires safeguards for protected health information (PHI). Medical records and devices have become attractive targets for cyber attackers in recent years [39], which has made hospitals and clinics reluctant to connect their local area networks (LANs) to the internet. Releasing patient test results requires the removal of PHI or authorization from the patient. Studies have shown that deidentification of patient data is not as simple as removing all specific identifiers because in the age of big data, under the right circumstances, it is possible to reassociate patients and their data using publicly available information [40-43].

We describe here the implementation of a real-time pathogen-specific surveillance system that overcomes the PHI concerns noted above. BioFire Syndromic Trends deidentifies, aggregates, and exports test results from FilmArray Instruments in use in US clinical laboratories [44]. Although data from all commercially available FilmArray panels [45] are exported to the Trend database, we focus here on the Respiratory Panel (RP) that can detect 17 viral (adenovirus, Adeno; coronavirus, CoV [OC43, 229E, NL63, HKU-1]; human metapneumovirus, hMPV;



human rhinovirus/enterovirus, HRV/EV; influenza A, Flu A [subtyping H1N1, 2009 H1N1, H3N2]; influenza B, Flu B; parainfluenza viruses, PIVs [1-4]; and respiratory syncytial virus, RSV) and three bacterial (*Bordetella pertussis, Chlamydia pneumoniae, and Mycoplasma pneumoniae*) pathogens [32,46,47].

With more than 362,000 patient results for the FilmArray RP test alone, the Trend database has many of the properties associated with big data as it applies to infectious disease [48]. After describing how the dataset can be cleaned of nonpatient tests, we make some observations on the seasonality of the different respiratory pathogens and the occurrence of codetection (more than one organism is detected in one test). Relatively little is known about rates of multiple concurrent respiratory infections and their overall impact on the health of the patient. Finally, we apply the ecological concept of species diversity [49] to observe a correlation between the abundance of each pathogen and the rate at which codetections (more than one positive result per test) occur in the tested population.

Methods

Origin of Syndromic Trends

FilmArray Trend was originally implemented to provide BioFire customers with an up-to-date view of the respiratory and GI pathogens circulating at their institution. From the perspective of an IVD manufacturer, the most uniform and thus the simplest method of accomplishing this is to follow a bottom-out approach to data export in which the FilmArray sends data to a cloud database managed by the manufacturer, and Web views of these data are available by clinicians at the hospital that generated the data (solid lines in Figure 1) rather than a top-out approach (dashed lines in Figure 1) in which the data are extracted from the hospital information system. This method provides the clinical institution with a tool to perform pathogen-specific surveillance for very little cost.

Patient Privacy When Exporting FilmArray Test Results

The Expert Determination study of the Trend data export algorithm (Multimedia Appendix 2) established that FilmArray patient results have been adequately deidentified. Therefore, a data use agreement (DUA), rather than business associates agreements (see Multimedia Appendix 2 for the difference between the two agreements) could be executed with each of the collaborating institutions (Multimedia Appendix 1). The DUAs define for the clinical laboratory how BioFire will manage and make use of the Trend data. The Trend client software residing on the FilmArray computer queries the FilmArray test result local database and exports the results to an Amazon Web Services database (Multimedia Appendix 1). The Trend client software performs deidentification on the FilmArray computer before export, as detailed in Multimedia Appendix 2. Health care providers (HCPs) are granted access to their institution's Trend data by the laboratory director. As Web access to view the data is restricted to the local site, deidentification of geographic indicators is not required. However, in the implementation of the public Trend website,

> JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 3 (page number not for citation purposes)

which presents FilmArray test results from around the United States, we have further aggregated the data with respect to geographic origin and obfuscated the date of the test (Multimedia Appendix 2). As only deidentified data are exported from the clinical institutions, no PHI is sent to or stored on the cloud server.

Test Utilization Rate and Pathogen Detection Rate

The FilmArray RP test utilization rate (TUR) metric is defined as the non-normalized number of RP patient test results generated each week across the Trend sites (computed as a centered 3-week moving average). To calculate the pathogen detection rate (as displayed in Figure 2 [second data view] and on the Trend website), we compute the rate for each organism at each institution as a centered 3-week moving average. To adjust for the capacity differences between sites, a national aggregate is calculated as the unweighted average of individual site rates. Only data from sites contributing more than 30 tests per week is included to avoid noise from small numbers of tests. Because the calculation of pathogen detection rate includes results from patients with multiple detections, the detection rate for all organisms can, in theory, add up to greater than one. In practice, this does not occur.

Comparison With the Centers for Disease Control and Prevention Influenza-Observed Rate of Detection

The CDC FluView rate of Flu A and Flu B detections, as well as the reported incidence of weighted influenza-like illness (ILI), are taken from the CDC website [4]. Only the CDC data from the Department of Health and Human Services regions that contained Trend pilot sites (Multimedia Appendix 1) were used for calculating the rate of influenza detections.

Calculation of Codetection Rates and Related Measures

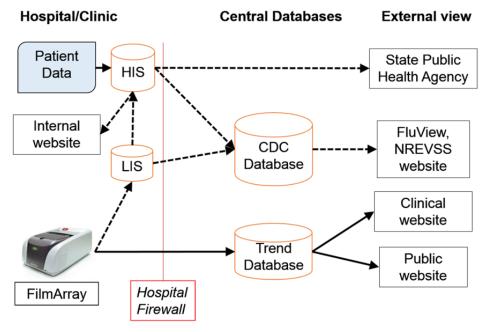
Pathogen codetections are defined as FilmArray tests in which two or three organisms are detected. We also calculated two other measures that relate to codetections: the circulating pathogen number and the measure of interspecific encounter (MIE). Both of these time series measures are calculated for each site and week, a centered 5-week moving average is computed, and then an unweighted average of all sites is used to create a national aggregate. The 5-week moving average is used to reduce noise because of small numbers of samples within a week at some sites.

More specifically, the circulating pathogen number is simply the count of the unique organisms detected at a site during a 1-week period. MIE is calculated from the frequencies of each organism at a site (number of positive test results for an organism divided by the number of FilmArray tests performed at that site). To reduce noise, we only include site data if more than 10 FilmArray tests were performed in that week. If $P_1...P_N$ are the percentage detection of the N different organisms circulating at a single site over a single week, then MIE is defined as shown in equation 1:

(1)
$$MIE = \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} P_i P_j$$

Conceptually, MIE is an attempt to estimate the likelihood that a patient infected with one organism may be infected with another unique organism circulating in the population at a given period in time, resulting in a coinfection.

Figure 1. Schema for export of in vitro diagnostic (IVD) test results to an external database. Bottom-Out and Top-Out approaches for data export are indicated by solid and dashed lines, respectively. Some institutions have developed their own systems for aggregating and displaying infectious disease data (indicated by internal website). HIS: hospital information system; LIS: laboratory information system; CDC: Centers for Disease Control and Prevention; NREVSS: National Respiratory and Enteric Virus Surveillance Systems.





JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 4 (page number not for citation purposes)

Figure 2. Detection of respiratory panel (RP) organisms over time across all sites. Detection of FilmArray RP pathogens in the Trend dataset displayed as stacked area graphs. All data views have the same time period (July 2013 through July 2017). (First data view) Count of each organism. The test utilization rate (TUR) metric (purple line, units are FilmArray RP tests performed) and count of FilmArray RP tests that are negative (white are between pathogen count and TUR) are indicated. The y-axis values are not indicated as this is considered proprietary information. (Second data view) Pathogen detection rates for all organisms. (Third data view) Pathogen detection rates for the subset of organisms that show seasonality (see Results and the legend for the list of organisms). (Fourth data view) Human rhinovirus (HRV) or enterovirus (EV) detection rates. The CDC weighted influenza-like illness (ILI; scaled up tenfold to be visible against the pathogen data) is indicated (black line) in the third and fourth data views. Organisms follow the same color scheme in all panels; the order of organisms in the legend (down then across) matches that of the stacked area graph top to bottom.



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 5 (page number not for citation purposes)

Results

Sending FilmArray Data Directly to the Cloud

The most general and efficient way to aggregate test results from the FilmArray instrument in a clinical laboratory is to follow a bottom-out approach to data export (Figure 1; Multimedia Appendix 1). In this scheme, the FilmArray instrument (at the bottom of the information hierarchy) directly sends data via the internet to a single cloud database where it can be viewed by HCPs at the originating institution. This data export pathway contrasts with a top-out approach (Figure 1) in which diagnostic test results are pushed from the instrument up through the LIS, to the HIS (at the top of the information hierarchy) and, finally, a subset of this information is forwarded to cloud-based databases.

Initial testing of the Trend export mechanism was performed in collaboration with the clinical laboratories of the Medical University of South Carolina. This trial allowed us to develop and test auto-export functions and deidentification protocols for the Trend software. The deidentification requirement of the Health Insurance Portability and Accountability Act (HIPAA) of 1996, specifically the Safe Harbor provision, requires the removal of 18 enumerated variables that could directly or indirectly identify an individual [50]. In accord with this requirement, the first stage study did not export test identifiers or free-form text fields and only returned the year of the test. The initial dataset provided low-resolution information but was a useful platform to evaluate the proposed system. Further development to enable export of higher resolution data required the design of routines that would adhere to an alternative HIPAA deidentification strategy, namely, the Expert Determination approach, which requires a risk assessment demonstrating that the chance of reidentifying an individual is sufficiently small [51]. The Expert Determination process identified and made recommendations for fields that could facilitate disclosure of PHI (Multimedia Appendix 2). A summary of the Expert Determination results detailing the risk of Trend data in regard to replicability, availability, and distinguishability is shown in Multimedia Appendix 2.

All sites (Multimedia Appendix 1) submitted the Trend project for review by their local institutional review board; all but one of the 20 review boards deemed the project exempt because of the absence of PHI export. Thus, the security requirements for the database and the controls necessary for storage and transport of deidentified data are significantly reduced.

Following the protocol established by Expert Determination review, the Trend software delays the export of results until the number of tests queued for export exceeds a minimum threshold for each type of FilmArray panel. In practice, this results in an average time to export of less than 2 hours from each site that has multiple instruments. A total of 99.11% (74,912/75,585) of the test results exported automatically occurred within 24 hours of test completion.

Characteristics of the FilmArray Sites Used in the Trend Pilot Study

The 20 sites contributing to the Trend pilot project (Multimedia Appendix 1) have the same average number of instruments; six (range: 1-22) as for all US FilmArray customers. The Trend pilot sites have been using the FilmArray RP test for an average of 3.8 years (range: 1-6) before June 2017. The size of the institutions participating ranges from 300 to 6400 beds, with the majority being large hospitals, and health care networks with an average of 1100 beds. Six (30%, 6/20) sites are pediatric hospitals, and one is a reference laboratory. Fifteen (75%, 15/20) of the sites have uploaded archived FilmArray RP test results to the Trend database, with eight (40%, 8/20) reporting results dating back to 2012. Unless stated otherwise, the data presented here cover the period from July 2013 to July 2017.

The algorithm used to diagnose the cause of respiratory disease varies by site. More than half of the Trend sites do not enforce an institutional respiratory testing protocol and, even within sites that have a required protocol, some discretionary use of FilmArray RP is allowed. Without detailed records from each institution's HIS, it is not possible to determine whether the FilmArray RP was used as a front line test or as a reflex test (typically following a negative result for influenza and RSV).

Cleaning Nonpatient Test Results From the Trend Database

To determine the prevalence of respiratory pathogens, we needed to expunge the Trend database of test results that are not derived from clinical patient samples. Nonpatient results come from a variety of sources including verification testing, routine quality control (QC), and proficiency testing (PT; Multimedia Appendix 3). Despite this complexity, the majority of nonpatient test results can be identified and distinguished from the patient-derived data because of the high number of positive organism calls in a single test and because of the temporal aspects of verification and control testing (Multimedia Appendix 3 shows one such identification method). QC tests are estimated to account for half of all FilmArray RP results in which more than three organisms are detected. In addition to the exclusion of tests temporally associated with validation events, all results with four or more positives were removed from further analysis (approximately 1% of the filtered total). This includes the small fraction of test results with exactly four organisms (Multimedia Appendix 3, Tests after event removal column) because the minority are derived from patient testing.

Detection of Respiratory Pathogens in Trend Samples From 2013 to 2017

The detection counts and pathogen detection rates derived from the Trend dataset for each organism in the FilmArray RP are shown in Figure 2. Other views of these data, including percent detection of individual organisms or combinations of organisms, are available on the BioFire Syndromic Trends public website [44]. The FilmArray RP TUR (see Methods) and the individual organism detection counts increased over this period because the Trend clinical sites increased their utilization of the FilmArray RP tests (Figure 2, first data view). Seasonal fluctuations can also be seen within this growth pattern, with



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 6 (page number not for citation purposes)

use increasing up to four-fold each winter when compared with the previous summer. HRV/EV, the most common pathogen detected group, is identified in approximately one-fourth of all samples tested each year (Multimedia Appendix 4). Other pathogens detected in approximately one-tenth of the samples include RSV, the PIVs, ADV, influenza, and hMPV. *M pneumoniae*, *C pneumonia*, and *B pertussis* are detected in a small percentage (one-fiftieth) of all samples. The average percentage of each organism is relatively constant over the 4 years of data in the Trend database (Multimedia Appendix 5).

The pathogens' seasonal variability measured by percent detection can be classified into at least three groups. Group 1: the majority of organisms follow the classical respiratory season (October-March) and increase by more than ten-fold above their baseline detection rate (Figure 2, third data view). These include the CoVs, Flu A, Flu B, hMPV, the PIVs, and RSV (PIV3 is a slight exception to this rule in that it peaks in the summer months and has a winter peak that is only detected regionally; data not shown). Within this group, all but five viruses demonstrate significant fluctuations from year to year; Flu B, hMPV, OC43, and PIV3 and RSV experience relatively consistent annual peaks. Group 2: HRV/EV is in a class by itself in that it is detected in a high percentage of tests over time (minimum of one-tenth of tests in winter) and experiences moderate peaks of two- to three-fold outside the respiratory season baseline in the early fall and spring (Figure 2, fourth data view). Group 3: the bacteria and Adeno are present at a relatively constant rate (Multimedia Appendix 6). The CDC FluView reported rate of ILI tracks moderately well with the group 1 organisms (cross-correlation of 0.85) and not with HRV/EV or with Adeno and the bacteria.

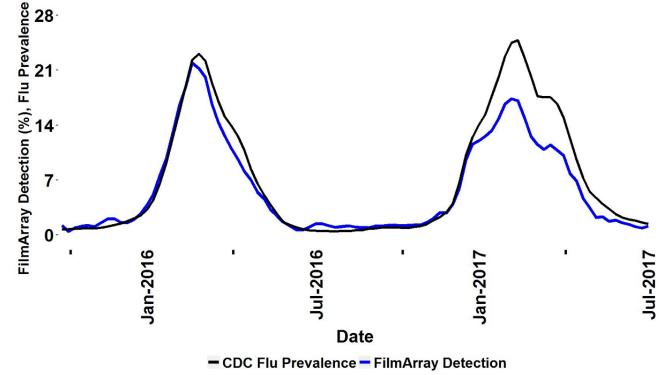
Comparison of Trend With Centers for Disease Control and Prevention Measures of Influenza

The CDC FluView network [4] gathers information about influenza prevalence from a large number of public health and clinical laboratories in the United States. FluView is considered the gold standard for these measures. We compared the Trend detection rates for Flu A (all subtypes) plus Flu B with the FluView Influenza (A and B) from September 2015 to July 2017 (Figure 3). The analysis was restricted to this time period because of a change in the CDC's reporting of flu prevalence in the fall of 2015. A cross-correlation of 0.974 was observed between the Trend Flu A or B percent detection and FluView reported influenza prevalence. Notably, the onset, peak, and duration of the influenza season coincide between the two measures.

Respiratory Panel Codetections

We found that approximately 38,000 FilmArray RP tests in the Trend dataset had two or three codetections. The most common codetections observed are those involving HRV/EV, which is the pathogen with the overall highest rate of detections (Figure 4, first data view). The codetection rate within each organism varies widely (from one-tenth to one-half; Figure 4, second data view). Although an additional pathogen was detected in half of the Adeno and CoV positive samples, codetections were observed in only one-tenth of the samples positive for either Flu A or Flu B (Figure 4, second data view).

Figure 3. Trend influenza detection rate compared with Centers for Disease and Prevention's (CDC) influenza activity. Percent of combined FilmArray Flu A (all subtypes) and Flu B detections (blue line) and CDC-reported influenza prevalence (black lines). CDC data are aggregated only from regions with participating Trend sites.



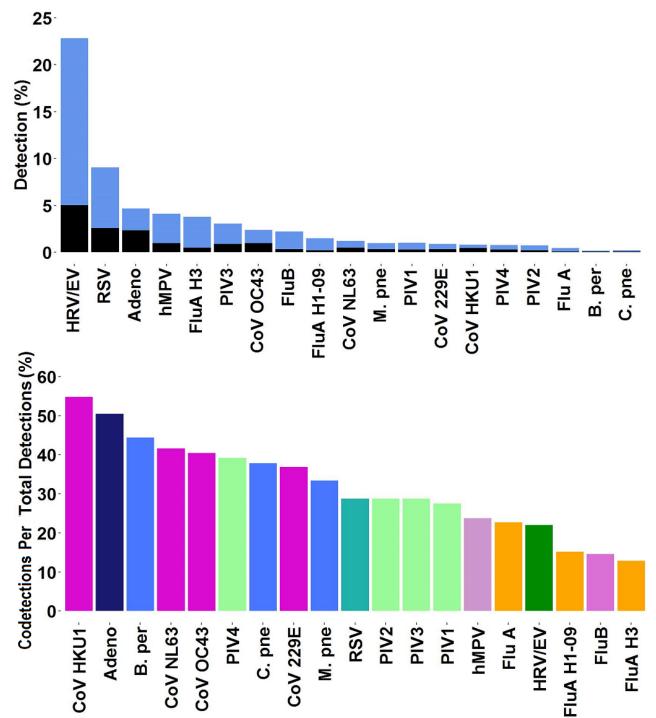


JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 7 (page number not for citation purposes)

http://publichealth.jmir.org/2018/

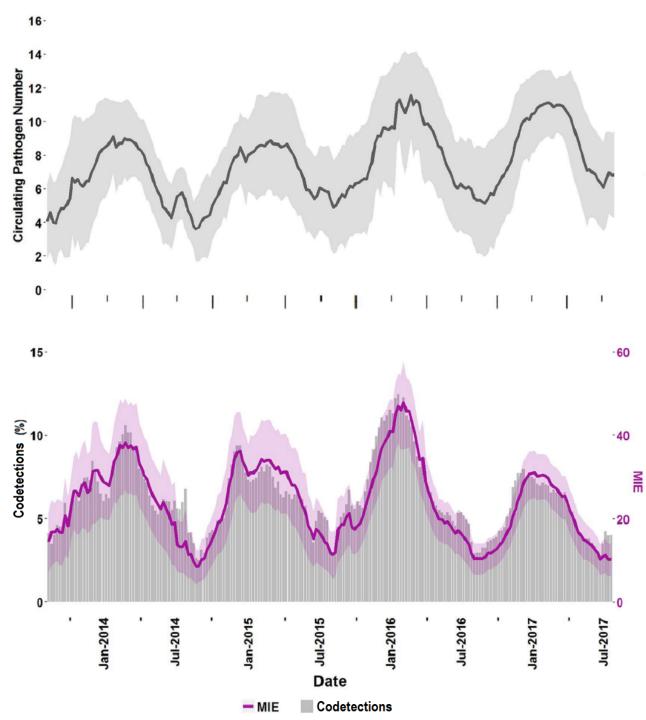
3/e59/

Figure 4. Detection rates for all organisms compared with codetections. Percent total positive detections for each organism in the respiratory panel (RP) Trend dataset is presented in stacked bars, showing the rate of detection of a single organism (first data view, blue) and those involved in a codetection (first data view, black). Data are calculated for each site during the period from July 2013 to July 2017, when available, and then aggregated. (Second data view) Percentage of each organism involved in a codetection is shown. Bars are colored by pathogen family (CoV, purple; bacteria, blue; PIVs, green; Flu A, yellow).



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 8 (page number not for citation purposes)

Figure 5. Seasonal variation in pathogen diversity and codetections. (First data view) Average circulating pathogen number (black line) and one SD computed across all Trend sites (gray area). (Second data view) Rate of codetections in the respiratory panel (RP) Trend dataset (gray bars, left axis), the measure of interspecific encounter (MIE) index (purple line, right axis), and MIE CIs (shaded purple area).



Trend data have high temporal, spatial, and organism-specific resolution. These three properties allow for a novel evaluation of codetections. The observed rates of codetections should be influenced by the number of circulating pathogens detected by the FilmArray RP test at a particular site. Figure 5, first data view, shows the average number of unique organisms detected at each site in a given week (see Methods: Calculation of codetection rates). This number fluctuates from a summer low of four to a winter high of 11 pathogens. Figure 5, second data view (gray bars), shows that the total rate of organism

codetections in the Trend dataset fluctuates annually, with peak rates occurring in the winter months. The average rates have been as high as one in 8 tests in the winter of 2016 and as low as one in 50 in the summer of 2014.

From the Trend data, an MIE can be calculated as the probability of a codetection, weighted by the prevalence of each circulating pathogen at a site. Although the value of the MIE metric is higher than the actual codetection rate, it correlates well (Figure 5, second data view, purple line compared with the gray bars has a cross-correlation of 0.9488 at a lag of 0). The magnitude



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 9 (page number not for citation purposes)

adjustment between MIE and the observed codetections is calculated by the slope of the linear regression of the two metrics (Multimedia Appendix 7) and has a value of $4.05 \ (R^2 = .9003)$.

Discussion

Properties of Trend Data

This study describes BioFire Syndromic Trends, a new system for real-time reporting of widespread pathogen-specific syndromic data. Even in its pilot phase, the Trend database already has many of the features that characterize big data [48]. The Vs of big data—volume (amount), velocity (speed of acquisition), veracity (accuracy), variety (diversity of information), and value (utility)—should be kept in mind as we consider the properties of Trend in clinical and public health settings.

The Trend RP dataset is growing at an average rate of >400,000 pathogen test results per month (>20,000 patient tests with 20 pathogens). Connecting the first 20 clinical sites has provided insight into the principal concerns that will be raised by the legal, information technology, and administrative departments of the HCPs that house FilmArray instruments. It should be possible, therefore, to expand the Trend installed base by 10to 20-fold over the next few years. Similarly, the existence of Trend should enable other IVD manufacturers to build their own Trend-like systems with greater acceptance on the part of their customers, thereby allowing a more global and comprehensive surveillance perspective.

The data in Figure 2 are similar to previous demonstrations of the seasonality associated with different respiratory viruses [52-55]. What is novel is that these data are generated automatically, on site, and in close to real time compared with other surveillance systems. Nearly all of the test results are exported to the Trend database within 24 hours of being generated. As part of the deidentification protocol, sequential FilmArray RP tests of the same type are put into the same time bin. This has the effect that test results are exported faster during periods of peak use, such as during the peak of the respiratory season or during an outbreak. Trend should be instrumental at a local level to determine the start of a respiratory season; many hospitals make significant changes to their operations based on this event; however, at present, data collection to track the respiratory season is often slow and manual, or semiautomated at best.

The key to implementing Trend clinical sites was to demonstrate that FilmArray test results can be exported without the risk of breaching PHI confidentiality either directly or through some combination of the data that were exported. Trend successfully used the Expert Determination process as prescribed by the HIPAA guidelines (see Multimedia Appendix 2), which greatly simplified the data sharing agreement between BioFire Diagnostics and the clinical site and allowed HCPs to use Trend without risk of inadvertently disclosing PHI.

The software architecture underlying the Trend system is both simple and secure: (1) no changes to the institutional firewall or LAN are needed; (2) the Trend database cannot reach back and query the FilmArray computer because of the institutional



firewall, which is set to outbound data only; and (3) Trend software can only submit data to the cloud database and cannot query the database (Multimedia Appendix 1). Yet, despite this security, authorized users of the Trend database can mine the deidentified data to look for novel patterns in respiratory pathogen epidemiology.

The Costs and Benefits of Bottom-Out Data Export System

The goal of an epidemiological surveillance network is to infer which infectious diseases are circulating in the general population based on testing a sample of patients [56]. Different surveillance systems have different biases in their data; biases that perturb the ability to predict true population prevalence.

Although the removal of all PHI has great benefits in terms of implementation, it also has several shortcomings that complicate interpretation of the data. First, Trend cannot account for the variability in the diagnostic testing algorithms applied to the selection of samples to be tested by the FilmArray instruments. During the respiratory season, HCPs may prescreen patients with other diagnostic tests including rapid antigen or molecular assays for influenza and RSV or commercial and laboratory-developed molecular tests for a mix of other respiratory pathogens. Depending upon the sensitivity of these upstream tests, more than half of influenza and RSV for the subset of the patients screened would be excluded from the Trend dataset if the front line test is positive. This testing protocol may skew the actual prevalence of not only influenza and RSV but all other individual respiratory pathogens and coinfections detected by the FilmArray. In some institutions, testing is reserved for hospitalized patients and others at risk for developing complications of respiratory tract infections, including the very young, very old, and immunocompromised patients. So Trend data may represent a less healthy patient population and not necessarily general community prevalence. Conversely, there are sites that perform a significant number of tests for the outpatient setting. This may create variability among the clinical sites' percent positivity and introduces a challenge to comparing pathogen intensity between sites.

The uncertainties surrounding the testing algorithm and the precise patient population tested should not interfere with determining the onset, peak, and duration of the pathogen season at each institution. These limitations on the data are likely to be common among almost all current surveillance systems for similar reasons. Given these concerns, the agreement between the percent positivity of Flu A or B as determined by Trend and the percent positivity reported by CDC FluView Influenza is striking (Figure 3), supporting the validity and utility of the Trend data.

The second source of concern in the Trend dataset is a consequence of the removal of sample identification such that we cannot directly determine whether the sample was from a patient or was a nonclinical sample (verification test, QC, or PT) and should be removed from further epidemiological analysis. We estimate that nonpatient testing makes up approximately one-fiftieth of the total FilmArray RP tests. Automated detection algorithms remove roughly one in 25 of the total RP tests, including approximately half of the nonclinical

JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 10 (page number not for citation purposes)

samples. With the exception of the four positive tests, the clinical samples removed by filtering should be a random sampling of all patient tests. The remaining nominal fraction of nonpatient tests has essentially no impact on the Trend evaluation of pathogen prevalence, but they do make it more difficult to perform high-resolution analysis of pathogen codetections. This is especially true for codetections of low prevalence organisms where QC positives are likely to be more common than real positives. Future updates to the FilmArray software will simplify the process by which the instrument operator can tag tests of nonpatient samples, thereby largely eliminating the need to filter such test results from the Trend database before analysis.

The Seasonality and Coinfections of Respiratory Pathogens

The total positivity rate of the FilmArray RP test varies from a low of approximately one-third of tests in the summer months to a high of three-fourths of the tests in December and January. Figure 5, second data view, shows that the average number of different circulating pathogens at a single institution can vary from eight up to 11 during the winter months. Even during the peak periods of ILI, many respiratory infections are due to other viruses (Figure 2, third data view) that can present clinically in a similar fashion [57,58]. Therefore, the presumption of an influenza infection based on reported influenza percent positivity, without diagnostic testing for the virus, can lead to the inappropriate use of antiviral agents [59]. Conversely, without comprehensive testing, a negative influenza or RSV test can lead to the prescription of an unnecessary antibiotic. Trend data can be a valuable aid for antimicrobial stewardship programs because it provides real-time information regarding the causes of respiratory infections and highlights the prevalence of viral infections.

As previously observed [55], the viruses that share the winter seasonality of influenza demonstrate annual or biennial behavior. It is possible that the viruses that share an influenza-like seasonality but do not show a two-year cycle (RSV and hMPV) are actually alternating strains, but the FilmArray RP Test does not detect this difference (eg, the FilmArray RP does not differentiate between RSV A and RSV B). Adeno and the bacteria show constant occurrence through the year; HRV is in a unique class with peaks in the fall and spring.

Detection of multiple respiratory viruses in the same patient has been reported before. In the Trend dataset, the rate of dual triple codetections was approximately and 7.94% (28,741/362,101), with HRV/EV as the organism most commonly observed in a codetection. Some viruses such as ADVs and the CoVs are detected in the presence of another organism approximately half of the time (Figure 4). In principle, a FilmArray RP positive result may represent detection of residual pathogen nucleic acid from a previous infection that has resolved. However, several studies suggest that coinfections are associated with more severe disease [60-62] (see also discussion in [63]). In such cases, information about multiple detections can provide infection control practitioners with data that can assist in bed management and in the assessment of risk for nosocomial infections in a patient population that has been



segregated by the occurrence of a common pathogen. Such information can prevent the introduction of a new pathogen associated with cohorting patients during busy respiratory seasons [64-66].

The question of whether different respiratory pathogens interfere with, or facilitate, growth in a human host is of some interest and not well understood. With the right data, it can be studied at the population [67], individual [68], and cellular level [63]. Because the Trend data still include some nonpatient tests, we have chosen not to analyze every possible dual or triple infection individually. Rather, we have taken a global approach and compared the overall rate of observed codetections with MIE, which is a measure of the diversity of viruses circulating in a specific region and time period. MIE is similar, but not identical, to Probability of Interspecific Encounter (PIE [69]), also referred to as the Gini-Simpson index (1-D, where D is the Simpson's index), which is used in ecology as a measure of the species diversity of a region. Similarly, the circulating pathogen number of Figure 5, first data view, is identical to the Species Richness measure of ecology. We calculate MIE using frequencies (P_i) of pathogen positivity per FilmArray test and note that the sum of all pathogen frequencies can add up to more than unity because of codetections or be less than unity because of the presence of negative tests. In this regard, MIE differs from PIE because it is not a probability measure.

Figure 5, second data view, shows that the observed rate of codetections is a constant fraction of MIE (approximately one-quarter as indicated by the linear regression of Multimedia Appendix 7). This observation suggests that, in the aggregate, respiratory pathogens are appearing in coinfections at a rate that can be predicted by their observed abundance. The data, however, may be biased by the patient population tested and the type of respiratory disease. The data also does not rule out that there are particular respiratory pathogens that occur more or less often in mixed infections than predicted by their individual percent positivity rates [63,70]. As we improve our ability to remove nonpatient test results from the Trend dataset, we will be able to characterize specific virus codetection rates and their significance [54,55,67,68,71,72].

Applications of Trend Data

As with weather forecasting, there is both a theoretical and a practical interest in predicting the next few weeks or months of the respiratory season [73-76]. Trend contributes to infectious disease forecasting efforts because the data are timely and comprehensive. As the number of sites participating in Trend increases, it will be possible to localize the reported infections to smaller geographical regions. At a high enough density of Trend sites, patterns of movement of respiratory pathogens across the United States will become visible in a way that has not been easily observed before now.

The Trend RP data show the percentage contribution of each pathogen to what is currently being detected by FilmArray RP testing (Figure 2, second data view) [44]. This analysis does not take into account changes in the rate of testing over a given season; information that should provide additional data regarding disease intensity and severity. In contrast, the simple metric, TUR, describes the non-normalized rate of FilmArray test usage

JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 11 (page number not for citation purposes)

and serves as a surrogate for the level of syndromic disease that HCPs observe (Figure 2, first data view).

TUR suffers from two defects. First, it is closely linked to the sales of the FilmArray test and thus is proprietary data that BioFire does not share (Google took a similar position in regard to releasing the search queries used by Google Flu Trends [12]). Second, TUR is driven by both the demand for testing and the growth in FilmArray product adoption and increasing acceptance and usage by HCPs. A useful step beyond TUR would be a normalized rate that can adjust for the underlying growth of testing unrelated to the intensity and duration of the respiratory disease season. An increase in a normalized TUR metric may indicate the prevalence of circulating respiratory viruses and the intensity of respiratory disease overall. Likewise, an increase in negative tests, may indicate the occurrence of an outbreak caused by an emerging pathogen.

Public health agencies, which include local and state health departments and the CDC, are specifically exempt under a HIPAA provision that allows clinical laboratories to disclose PHI to the agencies for specified public health purposes [77]. The exemption includes follow-up studies on reportable infectious diseases. Real-time pathogen-specific syndromic surveillance systems such as Trend will allow state health departments to more rapidly identify, acquire, and test residual samples from potential outbreaks. Conversely, perceived outbreaks may actually be coincidental multi-organism seasonal surges, and rapid analysis by Trend-like systems could prevent timely and costly outbreak investigation.

Given the movement in health care technology toward greater vertical integration of a hospital's data, the bottom-out approach exemplified by Trend will face more competition from top-out approaches (Figure 1, see, eg, GermWatch in Utah, [21]) because these systems can capture patient information (eg, age, gender, and patient address) that is critical for more detailed epidemiological analysis. However, combining PHI with the diagnostic test result in the top-out approach makes these systems more complex and difficult to implement and may limit participation by health care institutions. Ironically, bottom-out data export systems have a role to play in the development of top-out systems because bottom-out export provides a rapid and efficient means to quality check the data flowing from top-out systems. Trend data could also be combined with data derived from other automated diagnostic platforms [78,79]. This work might best be accomplished by a third party that is viewed as independent and impartial. For example, in the case of data originating in the United States, a federal institution or a private foundation could host a database to which IVD manufacturers would contribute their different syndromic test results. The benefits of a more complete view of circulating pathogens should outweigh the complexities of combining data from different platforms.

Future Outlook

Syndromic Trends is a novel surveillance tool for simultaneously monitoring multiple syndromic diseases that has demonstrated promise in expanding our knowledge of the epidemiology of infectious diseases. Indeed, the close correlation of seasonal respiratory viruses tracked by Trend with reported CDC ILI highlights the major contributory role of multiple respiratory pathogens beyond influenza to ILI. The national and global expansion of Trend will provide a comprehensive tool to study the impact of coinfections, understand the role of previously underappreciated pathogens, and clarify true disease epidemiology. Finally, systems such as Trend will be essential for the rapid identification of disease anomalies indicating potential emergent outbreaks, thereby providing an independent tool for public health surveillance.

Acknowledgments

This work was partially supported by NIH grant 5U01AI074419 (LM, KMR, and MAP). The authors would like to thank Chris Thurston and Spencer Rose (BioFire Defense) for building the Trend public website; Andrew Wallin (BioFire Defense) for reviewing the MIE data analysis; Anna Hoffee (BioFire Diagnostics) for assistance with the figures; Mark Pallansch (CDC), Kirsten St. George (New York State Department of Health), and Allyn Nakashima (Utah Department of Health) for useful discussions; Anne Blaschke and colleagues at BioFire Diagnostics and BioFire Defense for reviewing the manuscript.

Conflicts of Interest

LM, ANF, RKN, CVC, JDJ, KMR, CCG, and MAP are present or former employees of bioMérieux, Inc, or its subsidiaries. bioMérieux markets the FilmArray System and Trend. FSN, PHG, DJ, VD, AL, JDB, SS, KAS, HS, RS, SJ, JAD, JCW, KL, FM, SLR, MA-R, PDF, GAS, SJM, SVS, and BMA are research contractors of BioFire Diagnostics for the development of the BioFire Syndromic Trends system. CCR and JFM are members of the Syndromic Trends Working Group. BAM is a paid consultant of BioFire Diagnostics.

Multimedia Appendix 1

BioFire Syndromic Trends System.

[PDF File (Adobe PDF File), 132KB-Multimedia Appendix 1]

Multimedia Appendix 2

Deidentification of Patient data for Infectious Disease Epidemiology.



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 12 (page number not for citation purposes)

[PDF File (Adobe PDF File), 127KB-Multimedia Appendix 2]

Multimedia Appendix 3

Cleaning Trend Data.

[PDF File (Adobe PDF File), 141KB-Multimedia Appendix 3]

Multimedia Appendix 4

Detection of FilmArray RP Organisms by Type.

[PDF File (Adobe PDF File), 196KB-Multimedia Appendix 4]

Multimedia Appendix 5

Detection of FilmArray RP Organisms by Year.

[PDF File (Adobe PDF File), 402KB-Multimedia Appendix 5]

Multimedia Appendix 6

Detection of Adenovirus and the Three Bacteria.

[PDF File (Adobe PDF File), 216KB-Multimedia Appendix 6]

Multimedia Appendix 7

Linear Regression of MIE and Observed Codetections.

[PDF File (Adobe PDF File), 207KB-Multimedia Appendix 7]

References

- Yang W, Karspeck A, Shaman J. Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. PLoS Comput Biol 2014 Apr;10(4):e1003583 [FREE Full text] [doi: 10.1371/journal.pcbi.1003583] [Medline: 24762780]
- 2. van den Wijngaard CC, van Pelt W, Nagelkerke NJ, Kretzschmar M, Koopmans MP. Evaluation of syndromic surveillance in the Netherlands: its added value and recommendations for implementation. Euro Surveill 2011 Mar 3;16(9) [FREE Full text] [Medline: 21392486]
- 3. Bonačić Marinović A, Swaan C, van Steenbergen J, Kretzschmar M. Quantifying reporting timeliness to improve outbreak control. Emerg Infect Dis 2015 Feb;21(2):209-216 [FREE Full text] [doi: 10.3201/eid2102.130504] [Medline: 25625374]
- 4. Centers for Disease Control and Prevention. 2016 Feb 18. Overview of Influenza Surveillance in the United States URL: https://www.cdc.gov/flu/weekly/overview.htm[WebCite Cache ID 6zL9BQRyR]
- 5. Centers for Disease Control and Prevention. 2018 May 9. National Respiratory and Enteric Virus Surveillance System (NREVSS) URL: <u>http://www.cdc.gov/surveillance/nrevss/[WebCite Cache ID 6wY28o3ZP]</u>
- Lamb E, Satre J, Hurd-Kundeti G, Liscek B, Hall CJ, Pinner RW, Centers for Disease Control and Prevention (CDC). Update on progress in electronic reporting of laboratory results to public health agencies - United States, 2014. MMWR Morb Mortal Wkly Rep 2015 Apr 3;64(12):328-330 [FREE Full text] [Medline: 25837244]
- Reaves EJ, Valle R, Chandrasekera RM, Soto G, Burke RL, Cummings JF, et al. Use of bibliometric analysis to assess the scientific productivity and impact of the global emerging infections surveillance and response system program, 2006-2012. Mil Med 2017 May;182(5):e1749-e1756. [doi: <u>10.7205/MILMED-D-16-00276</u>] [Medline: <u>29087920</u>]
- Nelson NP, Yang L, Reilly AR, Hardin JE, Hartley DM. Event-based internet biosurveillance: relation to epidemiological observation. Emerg Themes Epidemiol 2012 Jun 18;9(1):4 [FREE Full text] [doi: 10.1186/1742-7622-9-4] [Medline: 22709988]
- Cheng CK, Lau EH, Ip DK, Yeung AS, Ho LM, Cowling BJ. A profile of the online dissemination of national influenza surveillance data. BMC Public Health 2009 Sep 16;9:339 [FREE Full text] [doi: 10.1186/1471-2458-9-339] [Medline: 19754978]
- Hulth A, Rydevik G. GET WELL: an automated surveillance system for gaining new epidemiological knowledge. BMC Public Health 2011 Apr 21;11:252 [FREE Full text] [doi: 10.1186/1471-2458-11-252] [Medline: 21510860]
- Benoit SR, Burkom H, McIntyre AF, Kniss K, Brammer L, Finelli L, et al. Pneumonia in US hospitalized patients with influenza-like illness: BioSense, 2007-2010. Epidemiol Infect 2013 Apr;141(4):805-815. [doi: <u>10.1017/S0950268812001549</u>] [Medline: <u>22800659</u>]



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 13 (page number not for citation purposes)

- 12. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
- 13. GoogleFlu. 2015 Aug 20. The Next Chapter for Flu Trends URL: <u>https://ai.googleblog.com/2015/08/</u> <u>the-next-chapter-for-flu-trends.html[WebCite Cache ID 6wY2CxwLm]</u>
- 14. Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, Wójcik O, et al. Flu Near You: crowdsourced symptom reporting spanning 2 influenza seasons. Am J Public Health 2015 Oct;105(10):2124-2130. [doi: 10.2105/AJPH.2015.302696] [Medline: 26270299]
- Das D, Metzger K, Heffernan R, Balter S, Weiss D, Mostashari F, New York City Department of Health and Mental Hygiene. Monitoring over-the-counter medication sales for early detection of disease outbreaks--New York City. MMWR Suppl 2005 Aug 26;54:41-46. [Medline: <u>16177692</u>]
- Patwardhan A, Bilkovski R. Comparison: flu prescription sales data from a retail pharmacy in the US with Google Flu trends and US ILINet (CDC) data as flu activity indicator. PLoS One 2012;7(8):e43611 [FREE Full text] [doi: 10.1371/journal.pone.0043611] [Medline: 22952719]
- Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013;8(12):e83672 [FREE Full text] [doi: <u>10.1371/journal.pone.0083672</u>] [Medline: <u>24349542</u>]
- Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011 May 4;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
- Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. PLoS Comput Biol 2014 Nov;10(11):e1003892 [FREE Full text] [doi: 10.1371/journal.pcbi.1003892] [Medline: 25392913]
- McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. PLoS Comput Biol 2014 Apr;10(4):e1003581 [FREE Full text] [doi: <u>10.1371/journal.pcbi.1003581</u>] [Medline: <u>24743682</u>]
- 21. Gesteland PH, Samore MH, Pavia AT, Srivastava R, Korgenski K, Gerber K, et al. Informing the front line about common respiratory viral epidemics. AMIA Annu Symp Proc 2007 Oct 11:274-278 [FREE Full text] [Medline: 18693841]
- 22. Nguyen TQ, Thorpe L, Makki HA, Mostashari F. Benefits and barriers to electronic laboratory results reporting for notifiable diseases: the New York City Department of Health and Mental Hygiene experience. Am J Public Health 2007 Apr;97 Suppl 1:S142-S145. [doi: 10.2105/AJPH.2006.098996] [Medline: 17413058]
- Crawford JM, Stallone R, Zhang F, Gerolimatos M, Korologos DD, Sweetapple C, et al. Laboratory surge response to pandemic (H1N1) 2009 outbreak, New York City metropolitan area, USA. Emerg Infect Dis 2010 Jan;16(1):8-13 [FREE Full text] [doi: 10.3201/eid1601.091167]
- 24. Ginocchio CC, Zhang F, Manji R, Arora S, Bornfreund M, Falk L, et al. Evaluation of multiple test methods for the detection of the novel 2009 influenza A (H1N1) during the New York City outbreak. J Clin Virol 2009 Jul;45(3):191-195. [doi: 10.1016/j.jcv.2009.06.005] [Medline: 19540158]
- 25. Al-Tawfiq JA, Zumla A, Gautret P, Gray GC, Hui DS, Al-Rabeeah AA, et al. Surveillance for emerging respiratory viruses. Lancet Infect Dis 2014 Oct;14(10):992-1000. [doi: 10.1016/S1473-3099(14)70840-0] [Medline: 25189347]
- 26. Gautret P, Gray GC, Charrel RN, Odezulu NG, Al-Tawfiq JA, Zumla A, et al. Emerging viral respiratory tract infections--environmental risk factors and transmission. Lancet Infect Dis 2014 Nov;14(11):1113-1122. [doi: 10.1016/S1473-3099(14)70831-X] [Medline: 25189350]
- 27. Salez N, Nougairede A, Ninove L, Zandotti C, de Lamballerie X, Charrel RN. Prospective and retrospective evaluation of the Cepheid Xpert® Flu/RSV XC assay for rapid detection of influenza A, influenza B, and respiratory syncytial virus. Diagn Microbiol Infect Dis 2015 Apr;81(4):256-258. [doi: 10.1016/j.diagmicrobio.2015.01.008] [Medline: 25662018]
- Nolte FS, Gauld L, Barrett SB. Direct comparison of Alere i and cobas Liat Influenza A and B tests for rapid detection of influenza virus infection. J Clin Microbiol 2016 Dec;54(11):2763-2766 [FREE Full text] [doi: 10.1128/JCM.01586-16] [Medline: 27582513]
- 29. Salez N, Vabret A, Leruez-Ville M, Andreoletti L, Carrat F, Renois F, et al. Evaluation of four commercial multiplex molecular tests for the diagnosis of acute respiratory infections. PLoS One 2015;10(6):e0130378 [FREE Full text] [doi: 10.1371/journal.pone.0130378] [Medline: 26107509]
- 30. Zhang H, Morrison S, Tang YW. Multiplex polymerase chain reaction tests for detection of pathogens associated with gastroenteritis. Clin Lab Med 2015 Jun;35(2):461-486 [FREE Full text] [doi: 10.1016/j.cll.2015.02.006] [Medline: 26004652]
- Buchan BW, Ledeboer NA. Emerging technologies for the clinical microbiology laboratory. Clin Microbiol Rev 2014 Oct;27(4):783-822 [FREE Full text] [doi: 10.1128/CMR.00003-14] [Medline: 25278575]
- 32. Poritz MA, Blaschke AJ, Byington CL, Meyers L, Nilsson K, Jones DE, et al. FilmArray, an automated nested multiplex PCR system for multi-pathogen detection: development and application to respiratory tract infection. PLoS One 2011;6(10):e26047 [FREE Full text] [doi: 10.1371/journal.pone.0026047] [Medline: 22039434]



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 14 (page number not for citation purposes)

- Pierce VM, Hodinka RL. Comparison of the GenMark Diagnostics eSensor respiratory viral panel to real-time PCR for detection of respiratory viruses in children. J Clin Microbiol 2012 Nov;50(11):3458-3465 [FREE Full text] [doi: 10.1128/JCM.01384-12] [Medline: 22875893]
- Nijhuis RHT, Guerendiain D, Claas ECJ, Templeton KE. Comparison of ePlex respiratory pathogen panel with laboratory-developed real-time PCR assays for detection of respiratory pathogens. J Clin Microbiol 2017 Jun;55(6):1938-1945 [FREE Full text] [doi: 10.1128/JCM.00221-17] [Medline: 28404682]
- 35. Gadsby NJ, Hardie A, Claas ECJ, Templeton KE. Comparison of the Luminex Respiratory Virus Panel fast assay with in-house real-time PCR for respiratory viral infection diagnosis. J Clin Microbiol 2010 Jun;48(6):2213-2216 [FREE Full text] [doi: 10.1128/JCM.02446-09] [Medline: 20357215]
- 36. Lee CK, Lee HK, Ng CW, Chiu L, Tang JW, Loh TP, et al. Comparison of Luminex NxTAG Respiratory Pathogen Panel and xTAG Respiratory Viral Panel FAST Version 2 for the detection of respiratory viruses. Ann Lab Med 2017 May;37(3):267-271 [FREE Full text] [doi: 10.3343/alm.2017.37.3.267] [Medline: 28224774]
- 37. Butt SA, Maceira VP, McCallen ME, Stellrecht KA. Comparison of three commercial RT-PCR systems for the detection of respiratory viruses. J Clin Virol 2014 Nov;61(3):406-410. [doi: 10.1016/j.jcv.2014.08.010] [Medline: 25183359]
- 38. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. BMC Public Health 2014 Nov 5;14:1144 [FREE Full text] [doi: 10.1186/1471-2458-14-1144] [Medline: 25377061]
- Williams PA, Woodward AJ. Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. Med Devices (Auckl) 2015;8:305-316 [FREE Full text] [doi: 10.2147/MDER.S50048] [Medline: 26229513]
- 40. Benitez K, Malin B. Evaluating re-identification risks with respect to the HIPAA privacy rule. J Am Med Inform Assoc 2010;17(2):169-177 [FREE Full text] [doi: 10.1136/jamia.2009.000026] [Medline: 20190059]
- 41. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PLoS One 2011;6(12):e28071 [FREE Full text] [doi: 10.1371/journal.pone.0028071] [Medline: 22164229]
- 42. El Emam K, Rodgers S, Malin B. Anonymising and sharing individual patient data. BMJ 2015 Mar 20;350:h1139 [FREE Full text] [Medline: 25794882]
- 43. Sarpatwari A, Kesselheim AS, Malin BA, Gagne JJ, Schneeweiss S. Ensuring patient privacy in data sharing for postapproval research. N Engl J Med 2014 Oct 23;371(17):1644-1649. [doi: <u>10.1056/NEJMsb1405487</u>] [Medline: <u>25337755</u>]
- 44. Syndromictrends. Respiratory Pathogen Trends URL: <u>https://syndromictrends.com/metric/panel/respiratory/percent_positivity/</u> organism/main [accessed 2018-06-06] [WebCite Cache ID 6zy2491xP]
- 45. BioFire Diagnostics. FilmArray® Panels URL: <u>https://www.biofiredx.com/products/the-filmarray-panels/[WebCite Cache ID 6wY2LgLDk]</u>
- 46. Hayden RT, Gu Z, Rodriguez A, Tanioka L, Ying C, Morgenstern M, et al. Comparison of two broadly multiplexed PCR systems for viral detection in clinical respiratory tract specimens from immunocompromised children. J Clin Virol 2012 Apr;53(4):308-313. [doi: 10.1016/j.jcv.2011.12.020] [Medline: 22296791]
- Popowitch EB, O'Neill SS, Miller MB. Comparison of the Biofire FilmArray RP, Genmark eSensor RVP, Luminex xTAG RVPv1, and Luminex xTAG RVP fast multiplex assays for detection of respiratory viruses. J Clin Microbiol 2013 May;51(5):1528-1533 [FREE Full text] [doi: 10.1128/JCM.03368-12] [Medline: 23486707]
- 48. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. J Infect Dis 2016 Dec 1;214(suppl_4):S375-S379 [FREE Full text] [doi: 10.1093/infdis/jiw400] [Medline: 28830113]
- 49. Magurran AE. Measuring Biological Diversity. Malden, MA: Wiley-Blackwell; 2003.
- 50. Office for Civil Rights. U.S. Department of Health & Human Services. 2012 Nov 26. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule URL: <u>https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/</u> De-identification/hhs_deid_guidance.pdf[WebCite Cache ID 6zLMpiEkk]
- Malin B, Karp D, Scheuermann RH. Technical and policy approaches to balancing patient privacy and data sharing in clinical and translational research. J Investig Med 2010 Jan;58(1):11-18 [FREE Full text] [doi: 10.2310/JIM.0b013e3181c9b2ea] [Medline: 20051768]
- 52. Stockmann C, Ampofo K, Hersh AL, Carleton ST, Korgenski K, Sheng X, et al. Seasonality of acute otitis media and the role of respiratory viral activity in children. Pediatr Infect Dis J 2013 Apr;32(4):314-319 [FREE Full text] [doi: 10.1097/INF.0b013e31827d104e] [Medline: 23249910]
- Radin JM, Hawksworth AW, Kammerer PE, Balansay M, Raman R, Lindsay SP, et al. Epidemiology of pathogen-specific respiratory infections among three US populations. PLoS One 2014;9(12):e114871 [FREE Full text] [doi: <u>10.1371/journal.pone.0114871</u>] [Medline: <u>25549089</u>]
- 54. Brittain-Long R, Andersson LM, Olofsson S, Lindh M, Westin J. Seasonal variations of 15 respiratory agents illustrated by the application of a multiplex polymerase chain reaction assay. Scand J Infect Dis 2012 Jan;44(1):9-17. [doi: 10.3109/00365548.2011.598876] [Medline: 21867470]
- 55. Olofsson S, Brittain-Long R, Andersson LM, Westin J, Lindh M. PCR for detection of respiratory viruses: seasonal variations of virus infections. Expert Rev Anti Infect Ther 2011 Aug;9(8):615-626. [doi: <u>10.1586/eri.11.75</u>] [Medline: <u>21819328</u>]



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 15 (page number not for citation purposes)

- Reed C, Chaves SS, Daily Kirley P, Emerson R, Aragon D, Hancock EB, et al. Estimating influenza disease burden from population-based surveillance data in the United States. PLoS One 2015;10(3):e0118369 [FREE Full text] [doi: 10.1371/journal.pone.0118369] [Medline: 25738736]
- 57. Ruuskanen O, Lahti E, Jennings LC, Murdoch DR. Viral pneumonia. Lancet 2011 Apr 9;377(9773):1264-1275. [doi: 10.1016/S0140-6736(10)61459-6] [Medline: 21435708]
- 58. Pavia AT. Viral infections of the lower respiratory tract: old viruses, new viruses, and the role of diagnosis. Clin Infect Dis 2011 May;52 Suppl 4:S284-S289 [FREE Full text] [doi: 10.1093/cid/cir043] [Medline: 21460286]
- 59. Brendish NJ, Malachira AK, Armstrong L, Houghton R, Aitken S, Nyimbili E, et al. Routine molecular point-of-care testing for respiratory viruses in adults presenting to hospital with acute respiratory illness (ResPOC): a pragmatic, open-label, randomised controlled trial. Lancet Respir Med 2017 May;5(5):401-411. [doi: <u>10.1016/S2213-2600(17)30120-0</u>] [Medline: <u>28392237</u>]
- 60. Goka EA, Vallely PJ, Mutton KJ, Klapper PE. Single, dual and multiple respiratory virus infections and risk of hospitalization and mortality. Epidemiol Infect 2015 Jan;143(1):37-47. [doi: 10.1017/S0950268814000302] [Medline: 24568719]
- 61. Martin ET, Fairchok MP, Stednick ZJ, Kuypers J, Englund JA. Epidemiology of multiple respiratory viruses in childcare attendees. J Infect Dis 2013 Mar 15;207(6):982-989. [doi: <u>10.1093/infdis/jis934</u>] [Medline: <u>23288925</u>]
- Rhedin S, Lindstrand A, Rotzén-Östlund M, Tolfvenstam T, Ohrmalm L, Rinder MR, et al. Clinical utility of PCR for common viruses in acute respiratory illness. Pediatrics 2014 Mar;133(3):e538-e545 [FREE Full text] [doi: 10.1542/peds.2013-3042] [Medline: 24567027]
- 63. Pinky L, Dobrovolny HM. Coinfections of the respiratory tract: viral competition for resources. PLoS One 2016;11(5):e0155589 [FREE Full text] [doi: 10.1371/journal.pone.0155589] [Medline: 27196110]
- 64. Thorburn K, Kerr S, Taylor N, van Saene HK. RSV outbreak in a paediatric intensive care unit. J Hosp Infect 2004 Jul;57(3):194-201. [doi: 10.1016/j.jhin.2004.03.013] [Medline: 15236847]
- 65. Gooskens J, Jonges M, Claas EC, Meijer A, van den Broek PJ, Kroes AM. Morbidity and mortality associated with nosocomial transmission of oseltamivir-resistant influenza A(H1N1) virus. J Am Med Assoc 2009 Mar 11;301(10):1042-1046. [doi: 10.1001/jama.2009.297] [Medline: 19255111]
- 66. Te Wierik MJ, Nguyen DT, Beersma MF, Thijsen SF, Heemstra KA. An outbreak of severe respiratory tract infection caused by human metapneumovirus in a residential care facility for elderly in Utrecht, the Netherlands, January to March 2010. Euro Surveill 2012 Mar 29;17(13) [FREE Full text] [Medline: 22490384]
- 67. Casalegno JS, Ottmann M, Duchamp MB, Escuret V, Billaud G, Frobert E, et al. Rhinoviruses delayed the circulation of the pandemic influenza A (H1N1) 2009 virus in France. Clin Microbiol Infect 2010 Apr;16(4):326-329 [FREE Full text] [doi: 10.1111/j.1469-0691.2010.03167.x] [Medline: 20121829]
- Karppinen S, Toivonen L, Schuez-Havupalo L, Waris M, Peltola V. Interference between respiratory syncytial virus and rhinovirus in respiratory tract infections in children. Clin Microbiol Infect 2016 Feb;22(2):208.e1-208.e6 [FREE Full text] [doi: 10.1016/j.cmi.2015.10.002] [Medline: 26482269]
- 69. Hurlbert SH. The nonconcept of species diversity: a critique and alternative parameters. Ecology 1971 Jul;52(4):577-586. [doi: <u>10.2307/1934145</u>] [Medline: <u>28973811</u>]
- Essa S, Owayed A, Altawalah H, Khadadah M, Behbehani N, Al-Nakib W. Mixed viral infections circulating in hospitalized patients with respiratory tract infections in kuwait. Adv Virol 2015;2015:714062 [FREE Full text] [doi: 10.1155/2015/714062] [Medline: 25983755]
- Bhattacharyya S, Gesteland PH, Korgenski K, Bjørnstad ON, Adler FR. Cross-immunity between strains explains the dynamical pattern of paramyxoviruses. Proc Natl Acad Sci U S A 2015 Oct 27;112(43):13396-13400 [FREE Full text] [doi: 10.1073/pnas.1516698112] [Medline: 26460003]
- 72. Achten NB, Wu P, Bont L, Blanken MO, Gebretsadik T, Chappell JD, et al. Interference between respiratory syncytial virus and human rhinovirus infection in infancy. J Infect Dis 2017 Apr 1;215(7):1102-1106 [FREE Full text] [doi: 10.1093/infdis/jix031]
- Nsoesie EO, Brownstein JS, Ramakrishnan N, Marathe MV. A systematic review of studies on forecasting the dynamics of influenza outbreaks. Influenza Other Respir Viruses 2014 May;8(3):309-316 [FREE Full text] [doi: <u>10.1111/irv.12226</u>] [Medline: <u>24373466</u>]
- 74. Chretien J, George D, Shaman J, Chitale RA, McKenzie FE. Influenza forecasting in human populations: a scoping review. PLoS One 2014;9(4):e94130 [FREE Full text] [doi: 10.1371/journal.pone.0094130] [Medline: 24714027]
- 75. Biggerstaff M, Alper D, Dredze M, Fox S, Fung IC, Hickmann KS, Influenza Forecasting Contest Working Group. Results from the centers for disease control and prevention's predict the 2013-2014 Influenza Season Challenge. BMC Infect Dis 2016 Jul 22;16:357 [FREE Full text] [doi: 10.1186/s12879-016-1669-x] [Medline: 27449080]
- 76. Moran KR, Fairchild G, Generous N, Hickmann K, Osthus D, Priedhorsky R, et al. Epidemic forecasting is messier than weather forecasting: the role of human behavior and internet data streams in epidemic forecast. J Infect Dis 2016;214(suppl 4):S404-S408. [doi: 10.1093/infdis/jiw375] [Medline: 28830111]
- Office for Civil Rights. United States Department of Health & Human Services. 2003 Apr 30. Disclosures For Public Health Activities [45 CFR 164.512(b)] URL: <u>https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/special/</u> <u>publichealth/publichealth.pdf[WebCite</u> Cache ID 6wY2zvbOH]



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 16 (page number not for citation purposes)

- 78. Shekalaghe S, Cancino M, Mavere C, Juma O, Mohammed A, Abdulla S, et al. Clinical performance of an automated reader in interpreting malaria rapid diagnostic tests in Tanzania. Malar J 2013 Apr 24;12:141 [FREE Full text] [doi: 10.1186/1475-2875-12-141] [Medline: 23617722]
- 79. Andre E, Isaacs C, Affolabi D, Alagna R, Brockmann D, de Jong BC, et al. Connectivity of diagnostic technologies: improving surveillance and accelerating tuberculosis elimination. Int J Tuberc Lung Dis 2016 Dec;20(8):999-1003 [FREE Full text] [doi: 10.5588/ijtld.16.0015] [Medline: 27393530]

Abbreviations

Adeno: adenovirus **CDC:** Centers for Disease Control and Prevention CoV: coronaviruses **DUA:** data use agreements ECLRS: Electronic Clinical Laboratory Reporting System Flu: influenza **GEIS:** Global Emerging Infections Surveillance HCP: health care provider HIPAA: Health Insurance Portability and Accountability Act HIS: hospital information system **hMPV:** human metapneumovirus HRV/EV: human rhinovirus/enterovirus **ILI:** influenza-like illness **IVD:** in vitro diagnostic LAN: local area network LIS: laboratory information system MIE: measure of interspecific encounter NREVSS: National Respiratory and Enteric Virus Surveillance Systems PHI: protected health information PIE: probability of interspecific encounter PIV: parainfluenza virus PT: proficiency testing **QC:** quality control **RP:** respiratory panel **RSV:** respiratory syncytial virus

TUR: test utilization rate

Edited by H Wu; submitted 25.01.18; peer-reviewed by J Olsen, D Broniatowski, M Paul, C Chen; comments to author 18.02.18; revised version received 29.03.18; accepted 12.04.18; published 06.07.18

Please cite as:

Meyers L, Ginocchio CC, Faucett AN, Nolte FS, Gesteland PH, Leber A, Janowiak D, Donovan V, Dien Bard J, Spitzer S, Stellrecht KA, Salimnia H, Selvarangan R, Juretschko S, Daly JA, Wallentine JC, Lindsey K, Moore F, Reed SL, Aguero-Rosenfeld M, Fey PD, Storch GA, Melnick SJ, Robinson CC, Meredith JF, Cook CV, Nelson RK, Jones JD, Scarpino SV, Althouse BM, Ririe KM, Malin BA, Poritz MA

Automated Real-Time Collection of Pathogen-Specific Diagnostic Data: Syndromic Infectious Disease Epidemiology JMIR Public Health Surveill 2018;4(3):e59 URL: <u>http://publichealth.jmir.org/2018/3/e59/</u>

doi: <u>10.2196/publichealth.9876</u> PMID: <u>29980501</u>

©Lindsay Meyers, Christine C Ginocchio, Aimie N Faucett, Frederick S Nolte, Per H Gesteland, Amy Leber, Diane Janowiak, Virginia Donovan, Jennifer Dien Bard, Silvia Spitzer, Kathleen A Stellrecht, Hossein Salimnia, Rangaraj Selvarangan, Stefan Juretschko, Judy A Daly, Jeremy C Wallentine, Kristy Lindsey, Franklin Moore, Sharon L Reed, Maria Aguero-Rosenfeld, Paul D Fey, Gregory A Storch, Steve J Melnick, Christine C Robinson, Jennifer F Meredith, Camille V Cook, Robert K Nelson, Jay D Jones, Samuel V Scarpino, Benjamin M Althouse, Kirk M Ririe, Bradley A Malin, Mark A Poritz. Originally published in JMIR Public Health and Surveillance (http://publichealth.jmir.org), 06.07.2018. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 17 (page number not for citation purposes)

Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on http://publichealth.jmir.org, as well as this copyright and license information must be included.



JMIR Public Health Surveill 2018 | vol. 4 | iss. 3 | e59 | p. 18 (page number not for citation purposes)

© 2018. This work is licensed under https://creativecommons.org/licenses/by/4.0/ (the "License"). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.

