Current Infectious Disease Reports (2019) 21:32
https://doi.org/10.1007/s11908-019-0694-5

TECHNOLOGY AND INFECTIOUS DISEASE (C HEBERT, SECTION EDITOR)

Electronic Health Record Use in Public Health Infectious Disease Surveillance, USA, 2018–2019

Sarah J. Willis[1] · Noelle M. Cocoros[1] · Liisa M. Randall[2] · Aileen M. Ochoa[1] · Gillian Haney[2] · Katherine K. Hsu[2] · Alfred DeMaria Jr[2] · Michael Klompas[1,3]

Abstract

Purpose of Review Electronic health records (EHRs) are an excellent source of data for disease symptoms, laboratory results, and medical treatments. Thus, EHR data may improve the completeness of notifiable disease case reporting and enable longitudinal collection of disease data. The purpose of this review is to examine the current state of EHR use in public health infectious disease surveillance in the USA.

Recent Findings A wide variety of EHR data is used in infectious disease surveillance. EHR data were used to assess the incidence of Lyme disease and identify newly diagnosed HIV infections. EHR disease detection algorithms combined laboratory reports, diagnosis codes, and medication orders to identify cases and, in the case of Lyme disease, found incidence rates 4–7 times higher than those from traditional surveillance. EHR data were also used to evaluate temporal trends in sexually transmitted disease testing, positivity, and re-testing in several primary care settings. Multiple studies were also able to control for additional confounders in multivariable models, such as number of sexual partners and concurrent infections, because of the breadth of data available in EHR systems.

Summary Studies highlighted in this review demonstrate that EHR data enhance provider-based and laboratory-based disease reports and may facilitate more complete case reporting. EHR data also provides corollary patient information that enables longitudinal disease reporting and analysis of important health outcomes. As public health infrastructure and investment allow health departments to establish closer relationships with healthcare providers, EHR data use in public health surveillance activities should continue to increase.

Keywords Electronic health records · Public health surveillance · Infectious diseases · Automated public health surveillance

This article is part of the Topical Collection on *Technology and Infectious Disease*

✉ Sarah J. Willis
sarah_willis@harvardpilgrim.org

Noelle M. Cocoros
noelle_cocoros@harvardpilgrim.org

Liisa M. Randall
liisa.randall@state.ma.us

Aileen M. Ochoa
aileen_ochoa@harvardpilgrim.org

Gillian Haney
gillian.haney@state.ma.us

Katherine K. Hsu
katherine.hsu@state.ma.us

Alfred DeMaria, Jr
alfred.demaria@state.ma.us

Michael Klompas
mklompas@bwh.harvard.edu

[1] Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA 02215, USA

[2] Bureau of Infectious Disease and Laboratory Sciences, Massachusetts Department of Public Health, Boston, MA, USA

[3] Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA

Published online: 26 August 2019

Introduction

Public health surveillance includes collection, analysis, and interpretation of health data [1], intended to inform public health actions that reduce morbidity and mortality. Currently, there are approximately 120 nationally notifiable conditions under public health surveillance in the United States (US) [2], the majority of which are infectious diseases. Responsibility for collecting, analyzing, and interpreting health data related to reportable diseases resides with local, state, and territorial health departments. However, healthcare providers, laboratories, and hospitals have the responsibility of reporting these diseases to health departments. The Council of State and Territorial Epidemiologists (CSTE), in conjunction with the Centers for Disease Control and Prevention (CDC) and health departments, develops case definitions and recommends which data elements should be collected for each disease to facilitate consistent disease reporting across jurisdictions [2]. CSTE case definitions include laboratory test results, supporting clinical information such as symptoms and treatment, and exposures or risk factors for disease transmission [2, 3].

Since the late 1990s, local, state, and territorial health departments in the US have integrated electronic laboratory reporting (ELR) into their surveillance programs [4, 5]. ELR is the automated transmission of reportable disease laboratory results to health departments from commercial, public health, and hospital laboratories. ELR has greatly improved the timeliness and completeness of disease reporting when compared with traditional, or paper-based, healthcare provider reporting [4, 5]. For example, in a study comparing ELR reports with traditional provider reports in Marion County, Indiana, ELR reports were received, on average, 8 days earlier than traditional provider reports, and the health department received approximately four times as many ELR reports than traditional provider reports for the same diseases [5]. The CDC also estimates that in 2018, 80% of laboratory reports received by state health departments were via ELR [6].

Despite these improvements, gaps in the required elements for case reports, including exposures or risk factors for disease acquisition and transmission, and supporting clinical information like symptoms, pregnancy status, and treatment, are not typically available through laboratory reporting. Indeed, ELR can paradoxically lead to a decrease in the completeness of reporting for variables not typically found in laboratory information systems [5]. Thus, health departments must still rely on non-automated methods (e.g., telephone calls and faxes to healthcare providers) to gather data for case reports, and work burden may increase due to efforts to improve case data completeness for an increasing number of reported cases.

Electronic health records (EHRs) may be an excellent alternative source of information on exposures, past medical history, symptoms, and treatment. Healthcare providers record data such as vital signs and patient-reported information, order medications, and laboratory tests, and refer patients to specialty providers through their EHR systems. Thus, EHR data may facilitate more complete case reporting and provide corollary patient information that enables longitudinal disease reporting and analysis of important health outcomes. Recent legislation has also encouraged providers to adopt or upgrade their EHR systems. In 2009, the American Recovery and Reinvestment Act (ARRA) was signed into law and included nearly $30 billion in incentives for hospitals and physician practices to adopt EHR systems [7]. By 2018, the CDC estimated that approximately 87% of healthcare providers were using EHR systems [6]. In addition, ARRA also provided monetary incentives for healthcare providers to demonstrate "Meaningful Use" of EHR data. This helped catalyze substantive improvements in EHR's capabilities, including their capacity to support data interchange and public health reporting. Indeed, improving population and public health was one of the five pillars of health outcomes promoted as "Meaningful Use" in this legislation [7, 8].

The purpose of this article is to review the current state of EHR use for public health infectious disease surveillance in the US. To identify studies, we conducted a PubMed search of manuscripts published between January 1, 2018, and April 30, 2019, using the key terms 'electronic medical records' or 'electronic health records' and 'public health surveillance.' We reviewed the manuscript abstracts to identify studies that focused on infectious disease surveillance conducted with US-based local, state, or territorial health departments (Table 1). We also included several conference abstracts and studies related to the surveillance platform 'Electronic medical record Support for Public health' (ESP) that the authors were familiar with (Table 1). The manuscripts and presentations highlighted throughout this review represent curated examples of EHR data use in public health surveillance. These studies are meant to demonstrate the variety of uses for EHR data but were not meant to be comprehensive. The studies described below demonstrate that EHR data, whether automated or not, can improve the completeness of disease reporting and provide corollary patient information that enables longitudinal disease reporting and the analysis of important health outcomes.

Electronic medical record Support for Public Health

Many of the studies we reviewed leveraged ESP, which is an open-source surveillance platform developed in partnership by public health officials, researchers, and clinicians in Massachusetts. ESP uses daily extracts of data from EHR systems to identify and report conditions of public health interest to local, state, or territorial health departments (esphealth.org).

Table 1 Electronic health record use in public health surveillance in the US, 2018–2019

Study, year	Jurisdiction/setting	Automated feed to public health surveillance	Objectives
Dee, 2019 [9••]	Three multisite ambulatory practice groups in Massachusetts	Automated feed to public health surveillance (ESP)	Assess repeat *Chlamydia trachomatis* testing and positivity patterns between 2010 and 2015
Dee, 2018 [10••]	Three multisite ambulatory practice groups in Massachusetts	Automated feed to public health surveillance (ESP)	Investigate whether observed increases in gonorrhea among men was due to more testing, increased incidence, or a combination of these two factors
Willis, 2019 [11••]	Single multisite ambulatory practice group in eastern Massachusetts	Automated feed to public health surveillance (ESP)	Assess the impact of an electronic best practice alert on EPT provision and chlamydia reinfection
Krakower, in press [12••]	Two ambulatory practice groups in Massachusetts	Automated feed to public health surveillance (ESP)	Develop an HIV risk prediction algorithm for providers to assess patient candidacy for PrEP
Lewis, in press [13••]	Six community health centers in Massachusetts	Automated feed to public health surveillance (ESP)	Describe efforts to improve identification of HIV infection, linkage to care, and retention in HIV care services through enhanced partnerships between state public health and community health centers. The paper also describes ESP implementation for HIV reporting.
Reilley, 2018 [14••]	Indian Health Service	Not an automated feed. Authors used EHR data as a new data source for surveillance purposes	Examine the rates, demographic characteristics, and annual trends in new HIV diagnoses among the American Indian/Alaskan Native population between 2005 and 2014
Moon, 2019 [15••]	Single multisite ambulatory practice group in Pennsylvania	Not an automated feed. Authors used EHR data as a new data source for surveillance purposes	(1) Compare EHR-based incidence rates for Lyme disease to traditional surveillance incidence rates (2) Evaluate individual and community risk factors for incident Lyme disease (3) Characterize the proportion of Lyme disease cases with post-treatment Lyme disease syndrome from EHR data
Kourtis, 2019 [16••]	US hospital-based database with administrative, health care utilization, and financial information (Premier database) and a multi-state inpatient EHR database (Cerner)	Not an automated feed. Authors used these EHR data as a new data source for surveillance purposes	Describe trends in the incidence of hospital-onset and community-onset MRSA and MSSA bloodstream infections
Tymejczyk, 2018 [17••]	Sexual health clinics in New York City	Not an automated feed. Authors merged EHR data with public health surveillance data	Investigate HIV care receipt and viral suppression among HIV-positive patients who also sought services from a sexual health clinic in 2012
Wagaw, 2018 [18••]	Primary-care safety-net clinics in Chicago, IL	Not an automated feed. Authors merged EHR data with health survey data	Explore the feasibility of electronically collecting self-administered patient survey data in primary care-clinics and linking and comparing the survey data with the patients' EHR data
Barnes, 2018 [19••]	Two hospitals in Pima County, Arizona	Not an automated feed. Authors merged EHR data with electronic death certificate data	(1) Quantify mortality among adult patients hospitalized for severe acute respiratory infections (2) Identify potential risk factors for severe acute respiratory infection mortality
Kunzweiler, in press [20••]	Six community health centers in Massachusetts	Not an automated feed. Authors merged EHR data with public health surveillance data.	Describe procedures used to identify and engage individuals who failed to fully link or engage in HIV care, and the sociodemographic characteristics and HIV health outcomes among these individuals

ESP Electronic medical record Support for Public Health, *EPT* expedited partner therapy, *EHR* electronic health record, *MRSA* methicillin-resistant *Staphylococcus aureus*, *MSSA* methicillin-susceptible *Staphylococcus aureus*, *HIV* human immunodeficiency virus

Through ESP, EHR data fields are mapped to common terms; these terms are analyzed for notifiable diseases or updates to existing cases of a notifiable disease, and then automatically submitted via electronic case report to health departments' surveillance systems [21–25]. Importantly, developing and maintaining ESP infrastructure requires substantial investment of time and resources on the part of health departments and healthcare providers to fully realize its promise.

As of 2019, ESP is used for notifiable disease case reporting in Massachusetts, Tarrant County, Texas, and Ohio. In Massachusetts, the following notifiable disease algorithms are in use: chlamydia, gonorrhea, acute hepatitis A,

acute hepatitis B, acute and chronic hepatitis C, HIV infection, syphilis, and tuberculosis. Notifiable disease algorithms are also in development for Lyme disease, anaplasmosis, babesiosis, and chronic hepatitis B. The algorithms vary in complexity; chlamydia is defined solely by positive laboratory tests while HIV infection is identified by specific combinations of laboratory tests, diagnosis codes, and medication orders. Recent studies using ESP demonstrate how additional patient information from EHRs improve completeness of disease reporting and enable longitudinal disease reporting for enhanced analyses of health outcomes [9••, 10••, 11••, 12••, 13••]. Several of these studies are described below.

Temporal Patterns in Chlamydia Repeat Testing in Massachusetts [9••]

The CDC's Sexually Transmitted Disease Treatment and Prevention Guidelines recommend test-of-cure for pregnant women and test for reinfection for all patients with chlamydia infection [26]. However, little is known about retesting rates and whether these rates are changing over time. Therefore, researchers and public health officials in Massachusetts used EHR data to evaluate temporal trends in chlamydia test-of-cure for pregnant females and test for reinfection for all patients between 2010 and 2015 [9••]. EHR data were collected via the ESP platform from three independent clinical practice groups [9••]. EHR data were necessary for these analyses because test-of-cure and test for reinfection rates require data on all chlamydia laboratory tests performed, regardless of the result. Currently, most health departments receive reports of *positive* chlamydia results from ELR but do not receive reports on negative or indeterminate laboratory tests.

The study identified an index positive chlamydia test for 972 pregnant females, 10,309 non-pregnant females, and 4973 males. Test-of-cure at the recommended interval of 3 to 5 weeks occurred in 37% of pregnant females. Test for reinfection at the recommended interval of 2 to 4 months occurred in 39% of pregnant females, 18% of non-pregnant females, and 9% of males. There were no significant increases in test-of-cure or test for reinfection rates between 2010 and 2015 [9••]. These findings identify a considerable gap between CDC's chlamydia retesting recommendations and what is occurring in practice. Thus, this study highlights the need for interventions that improve translation of clinical care guidelines into clinical practice.

Gonorrhea Testing Patterns Among Males in Three Large Clinical Practices in Massachusetts, 2010–2017 [10••]

Gonorrhea case rates in the US increased by 75% between 2009 and 2017, predominantly among men [27]. The Massachusetts Department of Public Health observed concomitant increases in gonorrhea rates among men since 2009 [28]. It is unclear whether the increase in gonorrhea among men is being driven by more testing for gonorrhea, an increase in gonorrhea incidence, or a combination of these two factors.

Men ≥ 15 years of age, who received care between 2010 and 2017 in one of three clinical practice groups in Massachusetts, were eligible for this study. All patient data were captured from the clinical sites' EHRs using the ESP platform [10••]. By using EHR data, the authors captured all gonorrhea laboratory results among patients with medical encounters. This allowed them to accurately calculate the percentage of men tested for gonorrhea and the percentage of men tested for gonorrhea with a positive result. In addition, symptoms related to gonorrhea were available in the EHR and allowed them to examine temporal trends in symptomatic testing.

Significant increases were observed in the percentage of men tested for gonorrhea and the percentage of men tested for gonorrhea with a positive result. In addition, the percentage of men who had symptoms at the time of their gonorrhea test decreased slightly, indicating that there may have been a small increase in asymptomatic screening of men during this time-period. The findings of this study support the idea that both more testing and more disease are driving the increase in gonorrhea among men in Massachusetts.

Impact of an Electronic Medical Record Best Practice Alert on Expedited Partner Therapy for Chlamydia Infection and Reinfection [11••]

Patients with *C. trachomatis* infections are at high risk for reinfection if their sexual partners are not treated [29–31]. Expedited partner therapy (EPT) allows providers to prescribe treatment to sex partners of patients diagnosed with chlamydia without examining the partner [32]. Atrius Health, a primary health care system located in eastern Massachusetts, implemented an electronic best practice alert (BPA) in their EHR in October 2014 to encourage providers to prescribe EPT. The objective of this analysis was to assess the impact of the BPA on EPT provision and chlamydial reinfection [11••]. Patients aged ≥ 15 years, with at least one positive chlamydia laboratory test between January 2013 and August 2018, were included. The ESP platform allowed researchers to collect patient-level data and physicians' response to the BPA from Atrius's EHR.

There were 6639 laboratory-confirmed chlamydia diagnoses identified, and EPT was provided to 21%. EPT provision frequency increased each month from 15% of chlamydia cases in January 2013 to 24% in October 2014 ($p = 0.03$). Paradoxically, the frequency of EPT provision decreased after introducing the BPA, to 16% of chlamydia cases in August 2018 ($p < 0.01$). On average, 65% of chlamydia cases received a test for reinfection and 11% were re-infected; there

were no changes in these percentages after the BPA was implemented. Thus, the authors concluded that EPT provision declined after introducing a BPA to facilitate EPT prescribing. However, this analysis demonstrates the value of EHR data to evaluate the impact of quality improvement initiatives.

Development and Validation of an Automated HIV Prediction Algorithm to Identify Candidates for Pre-exposure Prophylaxis [12••]

Pre-exposure prophylaxis (PrEP) is an effective method to prevent HIV infection [33–36]. However, PrEP is underutilized [37] in part because clinicians lack tools to identify patients who might benefit from it [38]. Therefore, the authors used EHR data from a health care system in eastern Massachusetts to develop an algorithm to identify individuals at increased risk for HIV acquisition. The authors also validated the algorithm with EHR data from a separate health care system in Massachusetts. EHR data from both health care systems were collected via ESP, as ESP enables gathering the same data from multiple EHR systems with different EHR vendors.

The best-performing algorithm discriminated between patients with and without incident HIV infection in the algorithm development health care system. When validated in the external health care system, the algorithm was also able to discriminate between patients with and without incident HIV infection and between patients with and without PrEP prescriptions. The authors concluded that prediction algorithms can identify persons at increased risk for HIV acquisition, and these algorithms could be integrated into EHRs to alert providers about patients who may benefit from PrEP.

Digital Bridge

The CDC released a report summarizing their efforts to improve public health surveillance in 2018 [6]. One of the goals for the CDC is to enhance public health surveillance through innovation, and a priority initiative is Digital Bridge [6]. Digital Bridge is a collaboration among federal agencies, state and local health departments, healthcare providers, and EHR vendors, and, when fully implemented, will result in the automated flow of initial case reports of notifiable diseases from EHRs to public health agencies [39, 40]. Digital Bridge is working with eight health departments across the US to pilot the program, and the first disease algorithms to be implemented include pertussis, gonorrhea, chlamydia, salmonellosis, and Zika virus infection [39, 40]. As of May 2019, two of the pilot sites, in Texas and Utah, are receiving electronic case reports from the healthcare providers they partnered with.

Additional Applications of EHR Data for Public Health Infectious Disease Surveillance

Several of the reports we identified in our PubMed search did not use automated data feeds from EHR systems to public health surveillance databases. A number of these used EHR data extracts as a new source of surveillance data [14••, 15••, 16••], while several others linked EHR data with existing surveillance data to enhance the quality and breadth of information available [17••, 18••, 19••, 20••]. The examples provided below again demonstrate that EHR data can improve the completeness of disease reporting and provide additional patient information to enable analysis of important health outcomes.

The Indian Health Service EHR system, which is used in 45 hospitals and more than 300 health centers in the US, was used to analyze HIV diagnoses among American Indian and Alaskan Natives between 2005 and 2014. Reilley et al. used International Classification of Diseases (ICD) diagnosis codes to identify newly diagnosed HIV infections and ICD diagnosis codes associated with opportunistic infections to identify newly diagnosed HIV infections that were considered late-stage [14••]. The authors emphasize that due to missing or inaccurate racial and ethnic data in public health surveillance and vital statistics databases, it is often hard to accurately assess the prevalence of HIV infection and other diseases among American Indians and Alaskan Natives. In addition, published analyses in the US often categorize these racial and ethnic groups as 'Other' and combine them with persons who self-report Asian race or multiple races. Their analyses identified 2273 newly diagnosed HIV infections among American Indian and Alaskan Natives in care in the Indian Health Service between 2005 and 2014, and approximately 20% of these infections were identified as late-stage, meeting criteria for AIDS.

In Pennsylvania, EHR data from the Geisinger Health system were used to assess the burden of Lyme disease between 2006 and 2014 [15••]. During the study period, Geisinger served 479,344 primary care patients across 38 counties in Pennsylvania. Using EHR data, Moon et al. calculated the incidence of Lyme disease and post-treatment Lyme disease syndrome. Lyme disease case definitions included a combination of ICD diagnosis codes, medication orders, and laboratory test orders and results. The authors were also able to compare their incidence estimates to incidence estimates from routine surveillance for Lyme disease in Pennsylvania. They found that the Lyme disease incidence estimates from EHR data were four to seven times higher than traditional surveillance estimates due to underreporting. These results highlight the utility of EHR data for case-finding, risk factor evaluation, and characterization of post-treatment Lyme disease syndrome.

The Premier Health Care Database and the Cerner Health Facts EHRs were used by Kourtis et al. to estimate trends in the incidence of hospital-onset and community-onset

methicillin-resistant *Staphylococcus aureus* (MRSA) and methicillin-susceptible *Staphylococcus aureus* (MSSA) infections between 2012 and 2017 in the US [16••]. The epidemiology of MSSA is not well understood, and existing data sources may not be nationally representative. The databases included data from more than 400 acute care hospitals during the study period. The authors estimated a significant decline in hospital-onset MRSA infection rates each year; however, there was no change in community-onset MRSA infection. In addition, hospital-onset MSSA infection rates did not change between 2012 and 2017, and the rates of community-onset MSSA infection significantly increased.

Lastly, we provide an example of an analysis that linked EHR data with existing surveillance data. In New York City, HIV surveillance data were linked to the city's sexual health clinics EHR system [17••]. The New York City Department of Health and Mental Hygiene was then able to determine whether persons diagnosed with HIV infection and "out-of-care" (i.e., no HIV viral load or CD4 laboratory result within the past 12 months in their surveillance system) were receiving services at the city's sexual health clinics. This would enable patient contact by the health department to reengage these patients in regular HIV care. Linkage of HIV surveillance data to EHR data also allowed the authors to examine risk factors for out-of-care status, such as concurrent STIs, number of sexual partners, and neighborhood poverty level (based upon current residence at time of sexual health clinic visit). These risk factors are often not regularly available in public health surveillance databases.

Conclusions

In this review, we describe several platforms that enable EHR data to be used for public health surveillance and provide examples of how EHR data enhance infectious disease surveillance in the US. The ESP platform is currently used in Massachusetts, Texas, and Ohio, and there are planned expansions in Maryland, Pennsylvania, and Washington. The CDC is working with eight public health jurisdictions, healthcare providers, and EHR vendors to implement Digital Bridge. While both platforms require substantial time and resources to implement and maintain, these platforms provide automated data feeds to public health surveillance systems. Automated feeds provide more timely and complete data and can minimize the burden on medical and surveillance staff to generate and receive case reports. In addition to these automated platforms, several public health agencies used EHR data to supplement their routine surveillance data through record linkages or by using an EHR data extract as a new source for surveillance information.

We demonstrate that EHR data enhance provider-based and laboratory-based disease reports and facilitate more complete reporting and case reports. For example, EHR data enable public health agencies to calculate incidence and prevalence estimates for reportable diseases that do not require laboratory confirmation. For some reportable diseases, like Lyme disease and influenza, providers observe patients' vital signs, current signs and symptoms, and empirically treat observations that indicate the presence of the disease. As these diseases are also common, it is unlikely that healthcare providers report all cases they encounter using traditional public health reporting methods. Therefore, EHR disease detection algorithms that are based on a combination of diagnoses, symptoms, medication orders, and laboratory results may identify cases of common diseases more efficiently when compared with surveillance methods that rely upon a positive laboratory result only. In fact, the Lyme disease study in Pennsylvania identified incidence estimates four to seven times higher in EHR data than in traditional surveillance [15••].

EHR data can also shed light on the incidence and prevalence of disease in marginalized or under-identified groups. As the example from the Indian Health Service shows, patients of American Indian and Alaskan Native heritage are often misclassified as another race or lumped into the category of 'other' race in surveillance systems and published analyses [14••]. Therefore, this leads to faulty estimation of the burden of disease in this population. By specifically studying this population using EHR data, the authors were able to accurately calculate trends in diagnoses over time. Race and ethnicity data are routinely captured in EHR systems and not always in ELR. In addition, data on the number of persons tested, and the number of persons with a positive result, are available for each race and ethnicity in EHR systems. Other populations in which EHR data could provide accurate estimates of incidence and prevalence include LGBT individuals and persons with substance use disorders, to the extent that sexual behavior or drug use is routinely captured in health systems' EHRs.

We also highlighted several studies that show how EHR data enable longitudinal disease surveillance and the analysis of important health outcomes. Health departments typically receive reports of *positive* laboratory tests for disease surveillance purposes, and these data allow them to calculate disease rates in their jurisdictions. Health departments do not routinely receive reports of negative or indeterminate laboratory test results (except in circumstances where jurisdictional regulations name certain negative tests, such as those related to hepatitis C), nor do they receive data on medical encounters in a given time-period. Thus, EHR data, as demonstrated in the studies by Dee et al., provide health departments additional information to calculate the prevalence of disease testing, retesting, and test positivity [9••, 10••]. Secondly, many health departments across the US promote EPT for chlamydia treatment but may have limited means of measuring the frequency it is prescribed and its impact on recurrent infection [41, 42].

Longitudinal data from EHRs can fill these data gaps and provide a better understanding of EPT, or other notifiable disease treatments, prescribing practices and effectiveness.

Several of the studies we reviewed were able to account for a wide variety of potential confounders in multivariable models because of the breadth of data available in EHR systems. In the ESP analysis of gonorrhea testing and gonorrhea test positivity, the authors controlled for STI testing history, use of HIV PrEP, and clinical signs and symptoms associated with gonorrhea [10••]. In the NYC linkage study of EHR and HIV surveillance data, the authors were able to assess concurrent STIs, neighborhood level poverty, and number of sexual partners as possible predictors for out-of-care status [17••]. These data fields are often missing or marked as "unknown" on case report forms and are not available through ELR reports at all.

There are limitations in using EHR data for public health surveillance. Key information, such as symptoms, are often captured in free-text clinical notes rather than structured fields in EHRs. Therefore, this information cannot be easily identified and automatically transmitted to health departments. However, natural language processing (NLP) programs are becoming more common. As researchers overcome challenges in processing free-text notes in clinically meaningful ways, NLP could be tailored to public health surveillance. A second limitation is that information about disease exposure, particularly for diseases related to travel and sexual or drug use behavior, is frequently missing in EHR systems. These data could be missing because the EHR does not prompt providers to collect the information; data are collected in free-text notes, or providers and patients may be reluctant to discuss sexual partners or drug use. Third, EHR systems typically apply to only one health care system. If patients seek care outside of their primary health care system, patient data would not be available to the health department if the additional care site(s) do not also submit EHR data. However, this limitation can be overcome if health departments are able to work with all or multiple health care systems in their jurisdiction. Lastly, substantial gaps exist with respect to health departments' abilities to receive, manage, and analyze data from EHRs. ARRA provided monetary incentives for healthcare providers who demonstrate "Meaningful Use" of EHR data and helped catalyze EHRs' capacity to support public health reporting. However, ARRA did not provide financial support for health departments to upgrade their surveillance systems to receive and manage EHR data. Therefore, significant investment in public health infrastructure is needed to support EHR data management activities and to develop better partnerships between public health and clinical entities.

The purpose of this review was to identify what types of activities are currently underway to enhance public health infectious disease surveillance data systems in the US with EHR data. There is a wide variety of EHR data use in public health surveillance, from automated data feeds to manual linkages with existing surveillance data. As investment in public health infrastructure allows health departments to establish closer relationships with healthcare providers, the use of EHR data in surveillance activities should continue to rise.

Compliance with Ethical Standards

Conflict of Interest Sarah J. Willis, Noelle M. Cocoros, Liisa M. Randall, Aileen M. Ochoa, Gillian Haney, Katherine K. Hsu, Alfred DeMaria, Jr., and Michael Klompas declare that they have no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
•• Of importance

1. Thacker S, Berkelman RL. Public health surveillance in the United States. Epidemiol Rev. 1998;10:164–90.
2. Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System. 2018. https://wwwn.cdc.gov/nndss/document/NNDSS-Fact-Sheet-508.pdf. Accessed on April 17, 2019.
3. Centers for Disease Control and Prevention. Surveillance case definitions for current and historical conditions. 2017. https://wwwn.cdc.gov/nndss/conditions/. Accessed on April 25, 2019.
4. Effler P, Ching-Lee M, Bogard A, Ieong M, Nekomoto T, Jernigan DB. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. JAMA. 1999;282:1845–50.
5. Overhage JM, Grannis S, McDonald CJ. A comparison of the completeness and timelines of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. Am J Public Health. 2008;98:344–50.
6. Centers for Disease Control and Prevention. CDC Public Health Surveillance Strategy Report: 2014–2018. 2018. https://www.cdc.gov/surveillance/pdfs/Surveillance-Series-Bookleth.pdf. Accessed on April 24, 2019.
7. Blumenthal D. Stimulating the adoption of health information technology. N Engl J Med. 2009;360:1477–9.
8. Marcotte L, Seidman J, Trudel K, Berwick DM, Blumenthal D, Mostashari F, et al. Achieving meaningful use of health information technology. Arch Intern Med. 2012;172:731–6.
9.•• Dee EC, Hsu KK, Kruskal BA, Menchaca JT, Zambarano B, Cocoros N, et al. Temporal patterns in chlamydia repeat testing in Massachusetts. Am J Prev Med. 2019;56:458–63. **The EHR data in this study were from an automated feed to a public health surveillance database.**
10.•• Dee EC, Willis SJ, Hsu KK, Elder H, Cocoros NM, Young J, et al. Gonorrhea testing patterns among males in three large clinical practices in Massachusetts, 2010–2017. Washington: Paper presented at the National STD Prevention Conference. 26–30 August 2018. Available at: https://www.esphealth.org/resources/publications-and-presentations. **The EHR data in this study were from an**

automated feed to a public health surveillance database.. Accessed on April 05, 2019.

11.•• Willis SJ, Elder H, Cocoros N, Callahan M, Flagg EW, Hsu KK, *et al*. Impact of an electronic medical record best practice alert on expedited partner therapy for chlamydia infection and reinfection. Paper presented at the Annual CSTE Conference. Raleigh, NC. 02–05 June 2019. Abstract available at: https://cste.confex.com/cste/2019/meetingapp.cgi/Paper/11435. Accessed on June 10, 2019. **The EHR data in this study were from an automated feed to a public health surveillance database.**

12.•• Krakower DS, Gruber S, Hsu K, Menchaca JT, Maro JC, Kruskal BA, *et al*. Development and validation of an automated HIV prediction algorithm to identify candidates for preexposure prophylaxis. Lancet. In press. **The EHR data in this study were from an automated feed to a public health surveillance database.**

13.•• Lewis S, Morrison M, Randall LM, Roosevelt K. The Partnerships for Care Project in Massachusetts: developing partnerships and data systems to increase linkage and engagement in care for individuals living with HIV. J Acquir Immune Defic Syndr. In press. **Study describes an automated feed of EHR data that is submitted for public health surveillance, along with electronic laboratory and provider reporting.**

14.•• Reilley B, Haberling DL, Person M, Leston J, Iralu J, Haverkate R, et al. Assessing new diagnoses of HIV among American Indian/Alaska Natives served by the Indian Health Service, 2005–2014. Public Health Rep. 2018;133:163–8. **The EHR data in this study were a new source of information for public health surveillance purposes.**

15.•• Moon KA, Pollak J, Hirsch AG, Aucott JN, Nordberg C, Heaney CD, et al. Epidemiology of Lyme disease in Pennsylvania 2006–2014 using electronic health records. Ticks Tick Borne Dis. 2019;10:241–50. **The EHR data in this study were a new source of information for public health surveillance purposes.**

16.•• Kourtis AP, Hatfield K, Baggs J, Mu Y, See I, Epson E, et al. Vital signs: epidemiology and recent trends in methicillin-resistant and in methicillin-susceptible Staphylococcus aureus bloodstream infections – United States. MMWR. 2019;68:214–9. **The EHR data in this study were a new source of information for public health surveillance purposes**.

17.•• Tymejczyk O, Jamison K, Pathela P, Braunstein S, Schillinger JA, Nash D. HIV care and viral load suppression after sexual health clinic visits by out-of-care HIV-positive persons. AIDS Patient Care STDs. 2018;32:390–8. **EHR data were merged with public health surveillance data to achieve the study objectives.**

18.•• Wagaw F, Okoro CA, Kim S, Park J, Rachman F. Linking data from health surveys and electronic health records: a demonstration project in two Chicago health center clinics. Prev Chronic Dis. 2018;15:170085. https://doi.org/10.5888/pcd15.170085 **EHR data were merged with public health surveillance data to achieve the study objectives.**

19.•• Barnes SR, Wansaula Z, Herrick K, Oren E, Ernst K, Olsen SJ, et al. Mortality estimates among adult patients with severe acute respiratory infections from two sentinel hospitals in southern Arizona, United States, 2010–2014. BMC Infect Dis. 2018;18:78. https://doi.org/10.1186/s12879-018-2984-1 **EHR data were merged with public health surveillance data to achieve the study objectives.**

20.•• Kunzweiler C, Kishore N, John B, Roosevelt K, Lewis S, Klevens RM, Morrison M, Randall LM, DeMaria A. Using HIV surveillance and clinic data to optimize data to care efforts in community health centers in Massachusetts: The Massachusetts Partnerships for Care Project. J Acquir Immune Defic Syndr. In press. **EHR data were used to create line lists of patients with HIV from six community health centers who were defined as out-of-care or diagnosed with HIV, but not linked to care.**

21. Centers for Disease Control and Prevention. Automated detection and reporting of notifiable diseases using electronic medical records versus passive surveillance — Massachusetts, June 2006–July 2007. MMWR. 2008;57:373–6.

22. Klompas M, McVetta J, Lazarus R, Eggleston E, Haney G, Kruskal BA, et al. Integrating clinical practice and public health surveillance using electronic medical record systems. Am J Public Health. 2012;102(Suppl 3):S325–32.

23. Klompas M, Cocoros NM, Menchaca JT, Erani D, Hafer E, Herrick B, et al. State and local chronic disease surveillance using electronic health record systems. Am J Public Health. 2017;107:1406–12.

24. Lazarus R, Klompas M, Campion FX, McNabb SJ, Hou X, Daniel J, et al. Electronic support for public health: validated case finding and reporting for notifiable diseases using electronic medical data. J Am Med Inform Assoc. 2009;16:18–24.

25. Vogel J, Brown JS, Land T, Platt R, Klompas M. MDPHnet: secure, distributed sharing of electronic health record data for public health surveillance, evaluation, and planning. Am J Public Health. 2014;104:2265–70.

26. Centers for Disease Control and Prevention. Sexually transmitted diseases treatment guidelines. 2015 MMWR Recomm Rep. 2015;64:1–137.

27. Centers for Disease Control and Prevention. 2017 Sexually Transmitted Disease Surveillance Report. 2018. https://www.cdc.gov/std/stats17/default.htm Accessed on April 25, 2019.

28. Massachusetts Department of Public Health, Bureau of Infectious Disease and Laboratory Sciences. 2016 Integrated HIV/AIDS, STD, and Viral Hepatitis Surveillance Report. 2017. http://www.mass.gov/eohhs/docs/dph/cdc/aids/std-surveillance-2016.pdf Accessed September 14, 2018.

29. Gaydos C, Wright C, Wood BJ, Waterfield G, Hobson S, Quinn TC. Chlamydia trachomatis reinfection rates among female adolescents seeking rescreening in school-based health centers. Sex Transm Dis. 2008;35:223–37.

30. Fung M, Scott KC, Kent CK, Klausner JD. Chlamydial and gonococcal reinfection among men: a systematic review of data to evaluate the need for retesting. Sex Transm Infect. 2007;83:304–9.

31. Hosenfeld CB, Workowski KA, Berman S, Zaidi A, Dyson J, Mosure D, et al. Repeat infection with chlamydia and gonorrhea among females: a systematic review of the literature. Sex Transm Dis. 2009;36:478–89.

32. Centers for Disease Control and Prevention. Expedited partner therapy in the management of sexually transmitted diseases. 2006. https://www.cdc.gov/std/treatment/EPTFinalReport2006.pdf. Accessed on September 18, 2018.

33. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med. 2010;363:2587–99.

34. Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N Engl J Med. 2012;367:399–410.

35. Thigpen MC, Kebaabetswe PM, Paxton LA, Smith DK, Rose CE, Segolodi TM, et al. Antiretroviral preexposure prophylaxis for heterosexual HIV transmission in Botswana. N Engl J Med. 2012;367:423–34.

36. Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock PA, Leethochawalit M, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2013;381:2083–90.

37. Smith DK, Van Handel M, Grey J. Estimates of adults with indications for HIV pre-exposure prophylaxis by jurisdiction, transmission risk group, and race/ethnicity, United States, 2015. Ann Epidemiol. 2018;28:850–7.

38. Chou R, Evans C, Hoverman A, Sun C, Dana T, Bougatsos C, et al. Pre-exposure prophylaxis for the prevention of HIV infection: a systematic review for the U.S. Preventive Services Task Force. Evidence synthesis no. 178. AHRQ publication no. 18–05247-EF-1. Rockville: Agency for Healthcare Research and Quality; 2018.
39. Cooney MA, Iademarco MF, Huang M, MacKenzie WR, Davidson AJ. The public health community platform, electronic case reporting and the digital bridge. J Public Health Manag Pract. 2018;24:185–9.
40. Digital Bridge Initiative: PHII Website. https://digitalbridge.us/infoex/. Accessed on April 24, 2019.
41. Okah E, Rogers AV, Kim M, Schillinger JA. Sentinel surveillance for expedited partner therapy prescriptions using pharmacy data, in 2 New York City neighborhoods, 2015. Sex Transm Dis. 2017;44:104–8.
42. Schillinger JA, Gorwitz R, Rietmeijer C, Golden MR. The expedited partner therapy continuum: a conceptual framework to guide programmatic efforts to increase partner treatment. Sex Transm Dis. 2016;43(Suppl 1):S63–75.