



HHS Public Access

Author manuscript

J Biomed Inform. Author manuscript; available in PMC 2020 January 14.

Published in final edited form as:

J Biomed Inform. 2018 March ; 79: 98–104. doi:10.1016/j.jbi.2018.02.012.

Electronic health record case studies to advance environmental public health tracking

Gonza Namulanda^{a,*}, Judith Qualters^{b,1}, Ambarish Vaidyanathan^a, Eric Roberts^c, Max Richardson^c, Alicia Fraser^d, Katharine H. McVeigh^e, Scott Patterson^f

^aEnvironmental Health Tracking Branch, Division of Environmental Hazards and Health Effects, National Center for Environmental Health, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, MS F-60, Atlanta, GA 30341, United States

^bDivision of Environmental Hazards and Health Effects, National Center for Environmental Health, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, MS F-60, Atlanta, GA 30341, United States

^cCalifornia Environmental Health Tracking Program, Public Health Institute, c/o Environmental Health Investigations Branch, 850 Marina Bay Pkwy, P-3, Richmond, CA 94804, United States

^dMassachusetts Department of Public Health, Bureau of Environmental Health, 250 Washington Street, 7th Floor, Boston, MA 02108, United States

^eDivision of Family and Child Health, New York City Department of Health and Mental Hygiene, 42-09 28th Street, Queens, NY 11101, United States

^fMissouri Department of Health and Senior Services, PO Box 570, Jefferson City, MO 65102, United States

Abstract

Data from traditional public health surveillance systems can have some limitations, e.g., timeliness, geographic level, and amount of data accessible. Electronic health records (EHRs) could present an opportunity to supplement current sources of routinely collected surveillance data. The National Environmental Public Health Tracking Program (Tracking Program) sought to explore the use of EHRs for advancing environmental public health surveillance practices. The Tracking Program funded four state/local health departments to obtain and pilot the use of EHR data to address several issues including the challenges and technical requirements for accessing EHR data, and the core data elements required to integrate EHR data within their departments' Tracking Programs. The results of these pilot projects highlighted the potential of EHR data for public health surveillance of rare diseases that may lack comprehensive registries, and surveillance of prevalent health conditions or risk factors for health outcomes at a finer geographic level. EHRs therefore, may have potential to supplement traditional sources of public health surveillance data.

*Corresponding author. fos0@cdc.gov (G. Namulanda).

¹Current Address: Division of Analysis, Research and Practice Integration, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Highway, NE MS F-62, Atlanta, GA 30341, United States.

Conflict of interest

The authors do not report any conflicts of interest.

Keywords

Electronic health records; EHR; Environmental public health tracking; Environmental public health surveillance

1. Introduction

Public health surveillance provides a foundation for public health programs and policies. For environmental public health, the ability collect and use data is an essential service in addressing community environmental public health problems [1]. However, the effort to systematically collect, manage, analyze, interpret, and disseminate data to those who can take action to improve the population's health has traditionally relied on a patchwork of data sources. For many non-communicable and environmentally-related diseases and their risk factors, these may include disease and exposure registries, vital statistics data, national and local health surveys and administrative data systems such as hospital discharge data [2]. Despite recent advances in public health surveillance science and practice, challenges remain in availability and accessibility of data for specific diseases and risk factors; population coverage to better characterize burden, especially at geographic resolution needed for local action; and the timeliness, quality, and efficiency with which data is collected and made available [3,4]. Electronic health records (EHRs) provide an opportunity to address these challenges [5].

1.1. Electronic health records

An electronic health record (EHR) can be defined as a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting, across more than one health care organization [6,7]. An EHR can contain information on a patient's demographics, diagnoses, medications, medical history, immunizations, laboratory test results, and radiology reports and images [6,7]. Although some information may be incomplete or lacking (e.g., behavioral information) [8], ideally, an EHR can bring together all information about a patient's health in one place [7]. Focus on the implementation of EHRs has increased since the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, which incentivizes meaningful use of EHR technology in eligible health care providers, hospitals or organizations [9], and more recently, the Medicare Access and CHIP Reauthorization Act (MACRA) of 2015 and its Merit-based Incentive Payment System (MIPS), which requires participating health care providers to use certified EHR technology [10].

The assessment and use of EHRs for public health surveillance have also increased. Much of this use has focused on the surveillance of infectious diseases with fewer examples focusing on EHR data for chronic disease surveillance [11]. An example is New York City's (NYC) Department of Health and Mental Hygiene (DOHMH) which is using EHR data to produce annual estimates of prevalence of diabetes, body mass index (BMI), smoking, and other population health indicators with the goal of informing municipal health policy [12]. Massachusetts Department of Public Health is using EHR data for surveillance of influenza-like illness, asthma, and diabetes [13]. Colorado Health Observation Regional Data Service

(CHORDS) regional pilot project uses EHR data to monitor public health trends on obesity, tobacco use, mental health conditions and cardiovascular disease [14]. Other uses have focused on EHR data for public health surveillance of eye health [15], diabetes [16], and asthma [17]. These applications demonstrate the feasibility, varying levels of validity, and generalizability of using EHR data to quantify the burden of specific health outcomes and behavioral factors such as smoking, among populations. They show that EHRs can provide more timely data at a finer geographic resolution (e.g., at a city/town or neighborhood level) [12,13,17] compared to traditional population-based surveillance methods such as the National Health and Nutrition Examination Survey (NHANES) (national level) or the Behavioral Risk Factor Surveillance System (BRFSS) (state level) which have limited sample sizes and cannot provide measurements in small geographic areas [17]. Moreover, systems querying aggregate data from EHRs can help streamline the disease reporting process through automation, making it more simple and cost effective [5].

The objectives of this report are to describe the utility of EHRs for advancing environmental public health surveillance practices (such as increasing the amount and timeliness of data available) and to present case studies summarizing innovative uses of EHRs by state and local health departments funded by the National Environmental Public Health Tracking Program (Tracking Program).

1.2. Environmental public health tracking

The Tracking Program was initiated in 2002 to establish a network for the ongoing collection, integration, analysis, interpretation and dissemination of data on environmental hazards, human exposure to environmental hazards, and potentially related health outcomes [18]. The National Environmental Public Health Tracking Network (Tracking Network) was subsequently launched in 2009 to provide data to help understand the associations between environmental exposures and chronic diseases, which can be used to drive public health actions that improve the health of communities [18]. The Tracking Network currently receives data from 26 funded state and local health departments (grantees), and presents these data in a standardized manner on the Tracking Network portal (<https://ephtracking.cdc.gov/showHome.action>) along with data from national partners such as the Environmental Protection Agency (EPA) [18]. These data cover environmental hazards, such as outdoor air quality; health outcomes, including asthma, cancer, and birth defects; biomonitoring results from childhood blood lead testing; chronic disease risk factors such as obesity; and population characteristics including socioeconomic measures [18].

Since 2002, the Tracking Program has greatly increased the capacity of grantee state and local health departments to perform environmental public health tracking, including expanding data collection and developing tools and standards for data analysis and dissemination [4]. Numerous public health actions resulting from state Tracking Programs have been reported [19]. Despite these successes, the effort continues to expand the utility of data on the Tracking Network portal by addressing certain challenges in obtaining and disseminating data that can be used to drive public health action. Some of these challenges are not having data from some states for some topics such as birth defects when no state registry exists, not having more timely data to facilitate rapid response, not having data at

finer geographic resolution (sub-county levels) to inform community-level public health actions, and not having data on specific health outcomes such as neurologic disorders, or on exposures to environmental hazards such as persistent organic pollutants [4]. EHRs could present an opportunity to supplement current sources of routinely collected surveillance data and mitigate some of these challenges; therefore, the Tracking Program sought to explore the use of EHRs to advance environmental public health tracking.

2. Case studies

In 2014, the Tracking Program funded state/local health departments in Massachusetts, New York City, California and Missouri to pilot the use of EHRs in their Tracking Programs. The health departments were tasked with obtaining EHR data, and addressing any of the following questions: (1) How can EHRs be used in Tracking? (2) What are the technical requirements for integrating EHR data into the state/local networks and the National Tracking Network? (3) What are core data elements needed to apply EHRs to Tracking? (4) What are the challenges and barriers to acquiring and processing EHRs? and (5) What are the innovative and emerging approaches to utilize EHRs within Tracking?

New York City, Massachusetts, and California addressed several of these questions while Missouri focused solely on the technical requirements and establishment of a secure data portal and data warehouse to enhance the ability to receive, validate, and process electronic health records data. The approach taken by each state varied based on existing infrastructure, partnerships with health care data providers, and data needs. New York City and Massachusetts's projects built off ongoing work to develop and evaluate an EHR-based surveillance system, New York City MacroScope and MDPHnet respectively. California partnered with Kaiser Permanente Northern California and focused on a subset of EHRs, electronic laboratory records. Details of the New York City and California projects have been published elsewhere [20-22]. We summarize below the purposes, methods, and results of the California, New York City, and Massachusetts projects as Case Studies. Missouri was not able to finish the project within the funded period; however, lessons learned from this project are included in the Discussion section.

2.1. California Tracking Program: electronic health records as a resource for public health surveillance—case study of glycohemoglobin testing and diabetes surveillance

2.1.1. Objective and relevance to environmental public health tracking—

Diabetes affects an estimated 30.3 million people in the U.S. population; about 23 million diagnosed and 7.2 million undiagnosed [23]. Type 2 diabetes accounts for 90–95% of all diabetes cases [23]. Research is increasingly pointing to the role of environmental hazards in increased diabetes risks [24-26]. After diagnosis, diabetes status and control are monitored primarily through repeat testing of patient HbA1c. HbA1c, measured as a percentage of HbA1c in the sample of hemoglobin, provides measures of healthy (< 5.7%), pre-diabetes (5.7–6.4%), controlled or partially controlled diabetes (6.5–8.4%), and uncontrolled diabetes (8.5% and above), which are strong indicators of a patient's ability to control their diabetes and subsequently avoid adverse health outcomes. The objective of this project was to pilot the use of electronic laboratory reports of blood glucose, hemoglobin A1c, as a proxy of

diabetes prevalence, for surveillance of type 2 diabetes that could inform community level public health intervention.

2.1.2. Methods and benchmark surveillance data—Data used as a benchmark/gold standard for type 2 diabetes surveillance is a clinical diagnosis as recorded in the Kaiser Permanente Northern California (KPNC) Diabetes Registry. A clinical diagnosis of diabetes may not be consistently available within varying providers' EHR systems, presenting a barrier to diabetes surveillance. A laboratory measure of glycohemoglobin, also known as hemoglobin A1c (HbA1c), may be more readily accessible from both healthcare providers and laboratories, and may function as a proxy measure for type 2 diabetes surveillance.

The California Tracking Program partnered with KPNC, the largest healthcare provider in California, to evaluate the utility of HbA1c for type 2 diabetes surveillance. California Tracking Program and KPNC analyzed HbA1c results from electronic laboratory records of the 412,400 KPNC members who lived in Contra Costa or Solano counties any time during 2010–2014, and were 18 years old or older as of January 1, 2010 (which represents about 27% of the population 18 years and older in both counties) [20]. The research team calculated the lowest, average, and maximum HbA1c result for each patient and aggregated the data using a priori cut points of 6.5%, 7%, 7.5%, 8%, 8.5% and 9% [20]. Each HbA1c metric was then tested for sensitivity, specificity and positive predictive value when compared to a gold standard of type 2 diabetes (clinic diagnosis of type 2 diabetes). The measure with the highest correlation to actual diagnosis was used to examine type 2 diabetes by income and race at the census tract level.

2.1.3. Results—The 5-year highest glycohemoglobin 6.5% measure most closely correlated with the gold standard clinical diagnosis of type 2 diabetes [20]. Similar to known trends using population level-data for type 2 diabetes, this measure decreased with increase in median family income by census tract, and increased with proportion of residents who are either non-Hispanic Black or Hispanic [20].

2.1.4. Implications and limitations—The project demonstrated that glycohemoglobin laboratory test results are a practical and valid candidate for more geographically refined, ongoing public health surveillance efforts of type 2 diabetes and diabetes control in this geographic region, despite a market penetration of KPNC of 27% among those 18 years and older. However, more work needs to be done to address issues such as data quality (e.g., de-duplication), the impact of membership tenures, and those with well-controlled type 2 diabetes [20]. Questions remain about the generalizability of this result to geographic regions with lower market penetration, or where there are other health care providers with different standards of care than KPNC [20].

2.2. Massachusetts Tracking Program: electronic health records for public health surveillance of pediatric asthma

2.2.1. Objective and relevance to environmental public health tracking—

Asthma is a leading chronic health condition in children and adolescents in the United States [27]. Studies have reported an association between asthma (hospitalizations and emergency

department visits) and air pollution exposures, such as particulate matter and ozone [28]. The objective of this project was to explore the utility of EHRs to conduct surveillance of pediatric asthma.

2.2.2. Methods and benchmark surveillance data—Pediatric asthma in Massachusetts is tracked through a school-based asthma surveillance program. Current environmental public health tracking methods track asthma attacks using hospitalization discharge data, but methods to track overall asthma disease prevalence among children or at sub-county geographic levels may not be available for some states. In Massachusetts, the school-based asthma surveillance system is considered a benchmark for pediatric asthma and represents an opportunity for comparison and evaluation of EHR data.

The Electronic medical record Support for Public Health (ESPnet) provides the architecture, software, and governance rules that enable clinical practices to securely share population-level aggregate EHR data with public health agencies [12]. The Massachusetts ESPnet is called MDPHnet; it connects ESPnet implementations in three large, independent, multi-specialty group practices that collectively provide care for 1.3 million people or about 15% of the population in Massachusetts [12]. Using data queried from MDPHnet, the Massachusetts Tracking Program evaluated the prevalence of pediatric asthma in 12 towns across the state using two different candidate case algorithms. One algorithm was based on the International Classification of Diseases, Ninth Revision (ICD-9) coding for asthma and the other was based on the Healthcare Effectiveness Data and Information Set (HEDIS) definition for asthma, which incorporates ICD-9 coding and medication prescriptions. Asthma prevalence estimates were adjusted using U.S. Census data to account for differences in the race/ethnicity and sex distributions in each town compared to those in the MDPHnet population.

2.2.3. Results—Comparing the MDPHnet asthma prevalence estimates to the benchmark estimates, the MDPHnet definition that was based only on ICD-9 codes fared better than the one that included both ICD-9 code and drug prescription information (average absolute difference of 4% vs. 7% when compared to benchmark prevalence). MDPH also found that the concordance of MDPHnet based pediatric asthma prevalence estimates with the benchmark estimates was positively correlated with the percent of population coverage by MDPHnet in the 12 covered communities.

2.2.4. Implications and limitations—The project highlighted the ability of EHR-based data to provide comparable surveillance results to those obtained by more established benchmark methods, though possibly dependent on sufficient population coverage of the EHR system. The project also brings to awareness the challenges of determining a standardized definition for the outcome of interest before EHR-based surveillance can be implemented, along with the need for a strong benchmark surveillance system for comparison.

2.3. Massachusetts Tracking Program: electronic health records for public health surveillance of amyotrophic lateral sclerosis

2.3.1. Objective and relevance to environmental public health tracking—

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder that leads to gradual upper and lower motor neuron degeneration and eventual death, usually within five years [29]. Although the causes of ALS remain unknown, it has been associated with several genetic and environmental factors [29]. Studies have shown conflicting or inconclusive results of the association between ALS and chemicals such as pesticides and heavy metals [29]. The purpose of this project was to explore the utility of EHR data for the surveillance of ALS.

2.3.2. Methods and benchmark surveillance data—

ALS is a rare disease that is currently tracked in Massachusetts by a specialized state-wide registry with mandatory physician reporting. The Argeo Paul Cellucci ALS Registry of Massachusetts is a resource intensive system and represents the gold-standard for disease surveillance with comprehensive annual case reporting, secondary catchment through vital records review, full medical record abstraction, and physician verification of each case [30]. It offers an ideal comparison for evaluating EHR-based ALS prevalence estimates.

Using data from MDPHnet, the Massachusetts Tracking Program evaluated five ALS algorithms to compare the accuracy of statewide prevalence estimates based on ICD-9 diagnosis codes alone, prescription of the drug riluzole used to treat ALS, or a combination of the two.

2.3.3. Results—When comparing MDPHnet prevalence estimates to the benchmark estimate for crude ALS prevalence in Massachusetts, the algorithm that most closely matched the benchmark of 6.9 cases per 100,000 was that which required two or more medical visits with an ALS ICD-9 code in the past two years with or without a riluzole prescription. Giving a modest undercount of 5.9 cases per 100,000, this algorithm fared much better than the others whose estimates varied widely from 1.7 to 10.5 per 100,000. The algorithm identified 91 cases from the MDPHnet population for the years 2007–2011, years for which benchmark data were available. Of those, 78% had confirmed ALS diagnoses in the Massachusetts statewide registry, 10% were documented in the registry but had unconfirmed diagnoses, 5% were reported to the registry but did not have ALS ($n = 4$) or were not residents of Massachusetts ($n = 1$), and 7% were never reported to the statewide registry, leaving their diagnoses in question. Relaxing the definition to 1 encounter in 2 years would have captured 11 more confirmed ALS cases, but would also have included 9 more non-ALS patients and 17 more patients with no information in the statewide registry.

2.3.4. Implications and limitations—The project demonstrated that EHR-based surveillance may be possible for rare diseases but systems for verification of diagnoses may be needed, especially for conditions with complicated diagnostic criteria such as ALS. As with the Massachusetts asthma analysis, the project highlights the need for a strong benchmark surveillance system for comparison, and to inform selection of an appropriate case algorithm.

2.4. New York City Tracking Program: evaluating the diagnostic validity of electronic health record-based surveillance indicators—the NYC Macroscopic chart review study

2.4.1. Objective and relevance to environmental public health tracking—

Obesity and smoking are risk factors for several environmentally-related health outcomes including diabetes and hypertension [25,31]. The objective of this project was to assess the diagnostic validity of the following five NYC Macroscopic indicators relative to population-level estimates from the New York City Health and Nutrition Examination Survey (NYC HANES): smoking, obesity, hypertension, diabetes, and elevated cholesterol.

2.4.2. Methods and benchmark surveillance data—

NYC HANES is a local version of the National HANES (NHANES) that assesses the health and nutrition status of NYC residents [32]. NYC HANES provides data on some conditions, e.g., diabetes, high blood pressure, high cholesterol, and depression, and environmental exposures such as second-hand tobacco smoke, lead, and mercury [32]. NYC HANES serves as a benchmark to compare EHR-based measurements of these conditions.

The New York City Macroscopic electronic health record surveillance system obtains EHR data from the NYC Department of Health and Mental Hygiene Primary Care Information Project's Hub Population Health Network (Hub). The Hub is a distributed query network that allows the health department to obtain aggregate data on patients from over 700 ambulatory practices in NYC [33].

NYC Tracking Program identified patients represented in both NYC HANES and NYC Macroscopic, who had signed medical record release forms addressed to their primary care provider, and had visited the provider within a year prior to the NYC HANES interview data. The NYC Tracking Program built a chart abstraction tool for secure data entry and data management of the EHR charts of these patients, reviewed the charts, classified them according to the five outcomes of interest, and calculated the prevalence for these indicators. The Program compared measures from NYC HANES and NYC Macroscopic by assessing the percent agreement, sensitivity, specificity, and positive and negative predictive values.

2.4.3. Results—

NYC Macroscopic indicator definitions applied to data abstracted in the chart review for smoking, obesity, hypertension, and diabetes attained a high level of validity (sensitivity and specificity > 0.90) relative to the NYC HANES self-reported survey measures [21,22]. Diagnostic validity, based on a smaller age-restricted sample (men age 40 and women age 45), was lower for cholesterol (sensitivity = 0.69 and specificity = 0.75) [21].

2.4.4. Implications and limitations—

The results suggest that NYC Macroscopic definitions and measures for smoking, obesity, hypertension and diabetes can be useful for tracking the health of New Yorkers. These data are usually only available through self-reported survey data. By accessing medical chart data directly, this pilot project provided a degree of rigor to the validation of EHR data for public health surveillance. However a limitation of this project was having a small sample size for analysis.

3. Discussion

The results of these pilot projects, also summarized in Table 1, highlight the potential to use EHRs for environmental public health tracking and broader non-communicable disease and risk factor surveillance. The results also demonstrate the potential of using EHRs for surveillance of rare diseases that may lack comprehensive registries, and surveillance of prevalent health conditions or risk factors for health outcomes at a finer geographic level. A synthesis of results of all the funded projects with an emphasis on challenges and lessons learned, as well as future tracking efforts related to EHRs is described in the following sections.

3.1. Electronic health records for environmental public health tracking

The pilot projects demonstrated ways that EHR-based estimates can be used to fill current gaps in surveillance data. The California Tracking Program was able to demonstrate the use of EHR-based estimates to map the distribution of type 2 diabetes at census tract level, a more geographically refined level than is available using existing survey data. The California Tracking Program used these measures to map the spatial distribution of age-adjusted rates of glycohemoglobin results, which revealed disparities in type 2 diabetes prevalence in populations based on median family income and race/ethnicity. Massachusetts demonstrated the possibility of EHR-based surveillance of prevalent conditions such as pediatric asthma as well as rare diseases such as ALS. In states that do not have a specialized registry for rare health outcomes such as ALS or birth defects, EHRs could provide data to develop prevalence estimates for these outcomes. However, these pilots have demonstrated the need for validating EHR-based estimates, which would be a challenge without a population-based benchmark. New York City demonstrated the use of EHRs for surveillance of some common risk factors of adverse health outcomes, which are usually only available through self-reported survey data. EHRs could be explored to provide less biased estimates using actual measurements of risk factors e.g., elevated cholesterol, though some risk factor data e.g., smoking, could also be self-reported in EHRs and therefore, may suffer the same bias as survey data.

3.2. Data elements needed and challenges in applying electronic health records data to environmental public health tracking

The several data elements available to develop EHR-based population estimates include demographics, diagnosis and co-diagnosis, prescription information, laboratory test results, social history, or information on risk factors for adverse health outcomes. Some challenges were identified in using these data to develop measures for public health surveillance. One challenge is identifying the correct definition of measures that can be used for public health surveillance. Massachusetts tested several algorithms with different combinations of diagnosis and medication prescription information to define prevalence measures of asthma and ALS. As a result, prevalence measures that closely matched the population-based estimates were identified based on different algorithms for each outcome using ICD-9 codes only for both asthma and ALS. An interesting observation was that medication information did not increase the validity of either the asthma or ALS measures. It is possible that algorithms suitable to develop measures for public health surveillance using EHR data may

differ for different health outcomes, therefore, more work will need to be done to identify and possibly standardize algorithms for health outcomes that are of interest to the Tracking Program. Interest has increased in exploring the potential of using EHR-based estimates to enable public health surveillance where no registry or population-based data exist. However, these pilot projects have demonstrated the need for population-based estimates to inform the validity of EHR-based estimates. More work will be needed to identify other methods of validating EHRs-based estimates, such as medical chart reviews.

Another challenge is the percent of the population covered in the electronic health records. The project results in Massachusetts showed that the accuracy of the EHR-based estimates were positively correlated to the percentage of the population covered by EHRs. For towns in Massachusetts with lower population EHR coverage, EHR-based asthma prevalence tended to underestimate asthma prevalence, while towns with higher EHRs coverage had asthma prevalence estimates closer to the population-based estimates.

3.3. Technical requirements and barriers to accessing electronic health records data

These pilot projects presented different models of storing, accessing, and protecting the confidentiality of EHR data. New York City's MacroScope and Massachusetts's MDPHnet distributed systems allow individual healthcare providers to house EHR data while allowing the state/local departments of public health access to only aggregate data through queries [12,13]. The distributed model protects the confidentiality of the patient and also limits the amount of data accessible by the public health programs [12,13]. The Missouri Tracking Program on the other hand is working on developing a common platform structure to store and access data. Missouri was also funded to pilot the use of EHR data within their Tracking Program, though they were not able to complete the pilot within the project period. The objective of the project was to develop an infrastructure to receive, validate, and process EHR data from Missouri's Electronic System for the Early Notification of Community-based Epidemics (ESSENCE) and from its Health Strategic Architectures and Information Cooperative (MOHSAIC) system. When complete, de-identified or aggregate EHRs and electronic laboratory data from ESSENCE and MOHSAIC respectively would be accessed from a central location. The California Tracking Program was able to access de-identified electronic laboratory data through partnership with the health care provider. These examples show that EHR data will likely be accessed via varied methods across Tracking states: a combination of distributed systems, direct collaboration with healthcare providers, and other implementations of health information exchanges (HIE).

HIEs facilitate the accessing, aggregating, sharing, and exchanging of EHR data. In 2014, 36 out of 50 states had laws relating to a HIE or Health Information Organizations (HIO) (e.g., their implementation or regulation) [34]. Additionally, in 2015, 84% of non-federal acute care hospitals, an increase from 9% in 2008, had adopted basic EHRs systems, which include electronic clinical information with clinician notes, computerized medications order entry, laboratory and radiology reports/images, and diagnostic test results [35]. Tracking Program states with no HIEs may find it more challenging to access EHR data.

There are also non-technical challenges of accessing EHR data. With the increase of meaningful use of EHR technology, sharing of reportable health outcome data from EHRs to

public health agencies may soon become simpler and routine. Because laws regulating the reporting of health outcomes and conditions may differ among states [36], accessing EHR data for health outcomes and conditions not yet reportable may be more of a policy issue than a technical issue. Access to non-reportable chronic disease data, e.g., diabetes may be simplified by mandating reporting. Doing so would make it easier for public health agencies to access these data without, for example, lengthy institutional review board (IRB) approval processes.

3.4. Emerging approaches and opportunities to utilize electronic health records in environmental public health tracking

As demonstrated by the case studies presented above, successful implementation of EHRs for public health surveillance and environmental public health tracking is the result of a collaborative effort between state health departments and groups that develop systems which aggregate EHR data for public health access and use. Tracking Programs that are interested in using EHRs data could begin by identifying groups within their state, similar to NYC Department of Health and Mental Hygiene's Primary Care Information Project or Harvard Medical School's Department of Population Medicine in Massachusetts (and their collaborators), that aggregate clinical data for public health practice, so as to initiate collaborations with them. In states where these groups are yet to be established, Tracking Programs could take advantage of their existing collaborations with health care providers and work toward expanding on them to include sharing of EHR data.

In addition to the different ways the pilot projects accessed EHR data, Tracking Programs could also consider other new and ongoing efforts focusing on interoperable solutions to facilitate access and sharing of EHR data, e.g., Fast Healthcare Interoperability Resources (FHIR). FHIR is a standard and architecture for exchanging health information electronically between healthcare applications [37,38]. FHIR interoperability resources provide a common way to define the information contents and structure for the core information that is shared by systems [37]. This enables sharing of data between diverse EHRs system technologies, and also aggregating of data from several EHRs [39]. FHIR is still a draft standard.

As discussed earlier, the recommended elements required to develop EHR-based estimates may depend on the selected health outcome and algorithm. More pilot projects are needed to define the best algorithm to develop EHR-based public health estimates for different health outcomes of interest to the Tracking Program, including defining the best combination of different data elements accessible through EHRs. The pilots could also help determine if these algorithms can be standardized across states to facilitate national comparisons of public health surveillance measures. Tracking Programs could also consider how they would validate the EHR-based estimates, which includes selecting appropriate benchmark data and methods for validating EHR-based estimates against population-based estimates, and additional methods such as medical chart reviews.

More work is also needed to identify other uses of EHRs for environmental public health tracking. For example, EHR data could be used to obtain more timely health outcome data to support responses to public health incidents such as wildfires, extreme heat or other extreme

weather events. EHR data may also be useful in developing estimates of exposure to environmental hazards (e.g., pesticides and lead) among population groups. Exposure data for several hazards are not available for most states, underscoring a need for exploring how EHR data can supplement the surveillance of exposures to hazards.

Finally, the Public Health Informatics Institute EHR toolkit for planning an EHR-based surveillance program is a resource that could be useful to state public health programs in their efforts to use EHR data for public health surveillance. The toolkit includes a set of field-tested tools to help programs plan public health surveillance systems using EHRs [40].

4. Conclusions

These pilot projects have demonstrated the use of data from EHRs for public health surveillance of health outcomes and risk factors of adverse health outcomes that are of interest to environmental public health tracking. More work needs to be done to define the best algorithms for public health surveillance measures using EHR data. More work is also needed to determine if these algorithms and their implementation could be standardized across states to provide comparable measures nationally. Other potential uses of EHR data, e.g., for the surveillance of exposure to environmental hazards need to be explored.

Acknowledgments

We would like to acknowledge and thank all the Environmental Public Health Tracking Program cooperative agreement partners and their collaborators who participated on the EHRs pilot projects: Katharine H. McVeigh, Elizabeth Lurie, Remle Newton-Dame and Wendy Mckelvey from, or collaborating with, the New York City Tracking Program; Robert Knorr and Alicia Fraser from the Massachusetts Tracking Program; Max Richardson, Eric Roberts and Paul English from the California Tracking Program; and Roger Gibson and Scott Patterson from the Missouri Tracking Program.

Funding

This work was supported by the Centers for Disease Control and Prevention's Environmental Public Health Tracking Program.

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

- [1]. Stanbury M, Anderson H, Blackmore C, Fagliano J, Heumann M, Kass D, McGeehin M, Functions of environmental epidemiology and surveillance in state health departments, *J. Public Health Manage. Pract* 18 (5) (2012).
- [2]. Smith PF, Hadler JL, Stanbury M, Rolfs RT, Hopkins RSCSTE Surveillance Strategy Group, "Blueprint version 2.0": updating public health surveillance for the 21st century, *J. Public Health Manage. Pract* 19 (3) (2013).
- [3]. Richards CL, Iademarco MF, Atkinson D, Pinner RW, Yoon P, MacKenzie WR, et al., Advances in public health surveillance and information dissemination at the centers for disease control and prevention, *Public Health Rep.* 132 (4) (2017).
- [4]. Kearney GD, Namulanda G, Quakers JR, Talbott EO, A decade of environmental public health tracking (2002–2012): progress and challenges, *J. Public Health Manage. Pract* 21 (2) (2015) S23–S35.
- [5]. Friedman D, Parrish G, Ross DA, Electronic health records and US public health: current realities and future promise, *Am. J. Public Health* 103 (9) (2013).

- [6]. Healthcare Information and Management Systems Society (HIMSS), Electronic Health Records, 4 2016 Available from: <<http://www.himss.org/library/ehr>>.
- [7]. Office of the National Coordinator for Health Information Technology, Learn electronic health records (EHR) basics 4 2016, Available from: <<https://www.healthit.gov/providers-professionals/learn-ehr-basics>>.
- [8]. Estabrooks PA, Boyle M, Emmons KM, Glasgow RE, Hesse BW, Kaplan RM, et al., Harmonized patient-reported data elements in the electronic health record: supporting meaningful use by primary care action on health behaviors and key psychosocial factors, *J. Am. Med. Inform. Assoc* 19 (2012).
- [9]. Healthcare Information and Management Systems Society (HIMSS), What is the Meaningful Use Incentive Program?, 11 2016 Available from: <<http://www.himss.org/resource/library/TopicList.aspx?MetaDataID=924>>.
- [10]. Office of the National Coordinator for Health Information Technology, FACT SHEET: Quality Payment Program and Health Information Technology, 10 2016 Available from: <https://www.healthit.gov/sites/default/files/macra_health_it_fact_sheet_final.pdf>.
- [11]. Birkhead GS, Klompas M, Shah N, Uses of electronic health records for public health surveillance to advance public health, *Annu. Rev. Public Health* 36 (2015).
- [12]. McVeigh KH, Newton-Dame R, Perlman S, Chernov C, Thorpe L, Singer J, Greene C, Developing an Electronic Health Record-Based Population Health Surveillance System, New York City Department of Health and Mental Hygiene, New York, 2013.
- [13]. Vogel J, Brown J, 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* 104 (12) (2014).
- [14]. Colorado Health Observation Regional Data Service (CHORDS), 1 2018 Accessible from: <<http://denverpublichealth.org/home/health-information-and-reports/data-and-informatics-tools/colorado-health-observation-regional-data-service-chords>>.
- [15]. Elliott A, Davidson A, Lum F, Chiang M, Saaddine JB, Zhang X, et al., Uses of electronic health records and administrative data for public health surveillance of eye health and vision-related conditions, *Am. J. Ophthalmol* 154 (60) (2012).
- [16]. Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R, Automated detection and classification of Type 1 Vs Type 2 diabetes using electronic health record data, *Diabetes Care* 36 (4) (2013).
- [17]. Tomasallo CD, Hanrahan LP, Tandias A, Chang TS, Cowan KJ, Guilbert TW, Estimating Wisconsin asthma prevalence using clinical electronic health records and public health data, *Am. J. Public Health* 104 (1) (2014).
- [18]. Centers for Disease Control and Prevention, Environmental Public Health Tracking Network, 11 2016 Available from: <www.cdc.gov/ephtracking>.
- [19]. Quakers JR, Strosnider H, Bell R, Data to action: using environmental public health tracking to inform decision making, *J. Public Health Manage. Pract* 21 (2) (2015) S12–S22.
- [20]. Richardson MJ, Van Den Eden SK, Roberts E, Ferrara A, Paulukonis S, English P, Evaluating the use of electronic health records for type 2 diabetes surveillance in 2 California counties, 2010-2014, *Public Health Rep.* 132 (4) (2017).
- [21]. McVeigh KH, Newton-Dame R, Chan PY, Thorpe LE, Schreiberstein L, et al., Can Electronic Health Records Be Used for Population Health Surveillance? Validating Population Health Metrics Against Established Survey Data. *eGEMS (Generating Evidence & Methods to improve patient outcomes)*, vol. 4: Iss.1, Article 27, 2016 Available from: <<http://repository.edm-forum.org/egems/vol4/iss1/27>>.
- [22]. Thorpe LE, McVeigh KH, Perlman S, Chan PY, Bartley K, Schreiberstein L, et al., Monitoring Prevalence, Treatment and Control of Metabolic Conditions in NYC Adults Using 2013 Primary Care Electronic Health Records: A Surveillance Validation Study. *eGEMS (Generating Evidence & Methods to improve patient outcomes)*, vol 4: Iss. 1, Article 28, 2016 Available from: <<http://repository.edm-forum.org/egems/vol4/iss1/28>>.

- [23]. Centers for Disease Control and Prevention. National Diabetes Statistics Report, 2017; Estimates of Diabetes and Its Burden in the United States, 2017 Available from: <<https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf>>.
- [24]. Grun F, Blumberg B, Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling, *Endocrinology* 147 (6 Suppl) (2006) S50–S55. [PubMed: 16690801]
- [25]. Thayer KA, Heindel JJ, Bucher JR, Gallo MA, Role of environmental chemicals in diabetes and obesity: a National Toxicology Program workshop review, *Environ. Health Perspect.* 120 (6) (2012) 779–789. [PubMed: 22296744]
- [26]. Eze IC, Hemkens LG, Bucher HC, et al., Association between ambient air pollution and diabetes mellitus in Europe and North America: systematic review and meta-analysis, *Environ. Health Perspect.* 123 (5) (2015) 381–389. [PubMed: 25625876]
- [27]. Centers for Disease Control and Prevention, Healthy Schools. 12 2016 Available from: <<https://www.cdc.gov/healthyschools/asthma/index.htm>>.
- [28]. Peel JL, Tolbert PE, Klein M, et al., Ambient air pollution and respiratory emergency department visits, *Epidemiology* 16 (2005) 164–217. [PubMed: 15703530]
- [29]. Bozzoni V, Pansarasa O, Diamanti L, Nosari G, Cereda C, Ceroni M, Amyotrophic lateral sclerosis and environmental factors, *Funct. Neurol* 31 (1) (2016).
- [30]. The Argeo Paul Cellucci Amyotrophic Lateral Sclerosis Registry of Massachusetts. 1 2018 Accessible from <<http://www.mass.gov/eohhs/gov/departments/dph/programs/environmental-health/als-registry/>>.
- [31]. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH, The incidence of comorbidities related to obesity and overweight: a systematic review and meta-analysis, *BMC Public Health* 9 (2009) 88. [PubMed: 19320986]
- [32]. New York City Health and Nutrition Examination Survey. 12 2016 Available from: <<http://nychanes.org/>>.
- [33]. Newton-Dame R, McVeigh KH, Schreiberstein L, Perlman S, Lurie-Moroni E, Jacobson L, et al., Design of the New York City Macroscopic: Innovations in Population Health Surveillance Using Electronic Health Records. eGEMS (Generating Evidence & Methods to improve patient outcomes), vol. 4: Iss.1, Article 26, 2016 Available from: <<http://repository.edm-forum.org/egems/vol4/iss1/26>>.
- [34]. Schmit C, Sunshine G, Pepin D, Ramanathan T, Menon A, Penn M, Transitioning from paper to digital: state statutory and regulatory frameworks for health information technology, *Public Health Rep.* (2017).
- [35]. Henry J, Pylypchuk Y, Searcy T, Patel V, Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2015, Office of the National Coordinator for Health Information Technology Data Brief, 2016.
- [36]. Council of State and Territorial Epidemiologists. State Reportable Conditions Assessment Query Results Tool. 2 2017; Available from: <<http://www.cste2.org/izenda/entrypage.aspx>>.
- [37]. Health Level Seven, FHIR overview. 11 2016 Available from: <<http://www.hl7.org/fhir/overview.html>>.
- [38]. Department of Health and Human Services. A Primer on FHIR: Lightweight, Reusable Web Technologies Can Help Solve Substantial Real-World Health Challenges. 11 2016; Available from: <<https://www.hhs.gov/idealab/2016/01/05/a-primer-on-fhir-lightweight-reusable-web-technologies-can-help-solve-substantial-real-world-health-challenges/>>.
- [39]. Georgia Institute of Technology, Health Informatics on FHIR. 11 2016 Available from: <<http://www.vip.gatech.edu/teams/health-informatics-fhir>>.
- [40]. Public Health Informatics Institute, EHRs Toolkit, 11 2016 Available from: <<http://phii.org/ehrtoolkit/introduction>>.

Table 1

Summary of the objectives, methods, results and limitations of state Tracking Program pilots of the use of EHR data for environmental public health tracking.

State Tracking Program	Objective	EHR data access	Measures	Results	Implications and limitations
California	Pilot surveillance of diabetes using hemoglobin A1c (HbA1c) electronic laboratory test results as a proxy of diabetes	Partnership with health care provider to access de-identified EHR data	Percentage of HbA1c in the sample of hemoglobin (6.5%, 7%, 7.5%, 8%, 8.5%, and 9%)	The 5-year highest HbA1c measure most closely correlated with the gold standard clinical diagnosis of diabetes 6.5%	HbA1c electronic laboratory test results are practical and valid for public health surveillance of diabetes. However, provider data represented only 27% of population 18 years
Massachusetts	Pilot the utility of EHR data for asthma surveillance	Distributed system providing aggregate data from independent clinical practices	ICD-9 code only or ICD-9 code and medication prescriptions for asthma	EHR-based prevalence definition based on ICD-9 codes fared better than the one with ICD-9 code and drug prescription when compared to benchmark	EHR-based surveillance of prevalent conditions is possible, and may be dependent on sufficient population coverage of EHR system
Massachusetts	Pilot the utility of EHR data for ALS surveillance	Distributed system providing aggregate data from independent clinical practices	ICD-9 code, medication prescriptions for ALS, or a combination of both	EHR-based prevalence definition based on two or more medical visits with an ALS ICD-9 code most closely matched the benchmark prevalence	EHR-based surveillance for rare diseases, e.g., ALS, may be possible but systems for verification of diagnosis may be needed
New York City	Assess the diagnostic validity of 5 EHR-based indicators relative to population-level estimates	Distributed system providing aggregate data from independent clinical practices	Diagnosis or self-reported measures of smoking, obesity, hypertension, diabetes and elevated cholesterol	EHR-based definitions achieved a high level of validity relative to NYC HANES survey data	EHR-based measures are useful for tracking population health. The project results were based on a small sample size