

MDPHnet: Secure, Distributed Sharing of Electronic Health Record Data for Public Health Surveillance, Evaluation, and Planning

Electronic health record systems contain clinically detailed data from large populations of patients that could significantly enrich public health surveillance. Clinical practices' security, privacy, and proprietary concerns, however, have limited their willingness to share these data with public health agencies.

We describe a novel distributed network for public health surveillance called MDPHnet. The system allows the Massachusetts Department of Public Health (MDPH) to initiate custom queries against participating practices' electronic health records while the data remain behind each practice's firewall.

Practices can review proposed queries before execution and approve query results before releasing them to the health department. MDPH is using the system for routine surveillance for priority conditions and to evaluate the impact of public health interventions. (*Am J Public Health*. Published online ahead of print October 16, 2014: e1–e6. doi:10.2105/AJPH.2014.302103)

Joshua Vogel, MPH, Jeffrey S. Brown, PhD, Thomas Land, PhD, Richard Platt, MD, MS, and Michael Klompas, MD, MPH

ELECTRONIC HEALTH RECORD

(EHR) systems present substantial opportunities to improve public health surveillance and evaluation. EHRs are rich in detailed clinical data that could significantly enhance public health agencies' capacity to monitor the spread of communicable diseases, measure burden and care patterns for chronic diseases, and evaluate the impact of public health interventions.¹ EHR data tend to be more accurate than self-reports, more granular than hospital discharges or claims data, more detailed than death certificates, more inclusive than survey samples, more complete than clinician-initiated reports, and more timely than all current surveillance sources with the possible exception of telephone notifications.^{2–5} Notwithstanding EHRs' potential to revolutionize public health surveillance and evaluation, very few public health agencies have been able to take advantage of EHR data to date.^{6–8} Barriers include clinical practices' reluctance to "give" their data to government agencies, heterogeneity between EHR systems, EHRs' limited capacity for interoperability, and health departments' lack of capacity to receive and analyze these types of data.^{6–10} Clinical practices are often concerned about the possibility of security breaches, privacy violations, the exposure of commercial data that might benefit competitors, and use of their data for involuntary

comparisons to other practices and practitioners.^{11,12}

We describe in this article a new public health surveillance and evaluation tool called MDPHnet that provides the Massachusetts Department of Public Health (MDPH) with secure, controlled access to EHR data from multiple independent clinical practices via a distributed network. Distributed networks effectively manage the tensions between privacy, security, and public health by allowing institutions to retain complete control of their health data, while simultaneously enabling authorized users to submit queries for authorized purposes.^{12–15} We provide an overview of the design, architecture, governance, and implementation of MDPHnet. We describe 2 proof-of-concept queries that demonstrate MDPHnet's capacity to enhance routine public health surveillance and evaluation. We also discuss additional potential applications for the system and future options to improve MDPHnet by adding new features, data sources, and links to other distributed networks.

DESCRIPTION OF MDPHnet

MDPHnet consists of a distributed network that provides the MDPH with the capacity to query clinical practices' EHR data. The distributed network obviates the need for practices to send all of

their EHR data to the health department; rather, electronic queries are distributed to participating practices for local execution behind their firewalls. Results are then securely returned to the department for aggregation and review.

Three large practices in Massachusetts currently participate in MDPHnet: Atrius Health, Cambridge Health Alliance, and Massachusetts League of Community Health Centers. Participation is voluntary. Together these practices serve approximately 1.3 million people (15% of the state population). This makes MDPHnet an ideal starting point for evaluating many population-based public health interventions. MDPHnet was built jointly by MDPH, Harvard Medical School, Massachusetts eHealth Institute (MeHI), and the clinical practices mentioned. The project was funded by a Challenge Grant from the Office of the National Coordinator for Health Information Technology.

MDPHnet was created by coupling together 2 open-source software applications: Electronic medical record Support for Public Health (ESP) and PopMedNet (popmednet.org). The combined system is called ESPnet. MDPHnet is the Massachusetts implementation of ESPnet.

ESP (esphealth.org) is an open source software application that extracts data from EHRs on a nightly basis, organizes the data into a standard format stored across multiple data tables, and

applies sophisticated algorithms to find conditions of public health interest.^{1,2,9} ESP's case-finding algorithms integrate diagnosis codes, laboratory tests, prescriptions, and vital signs to maximize accuracy.^{2-4,16-19} Cases are recorded in the local ESP database and transmitted to the state health department as either HL7 electronic case reports or population-level aggregate summaries as appropriate. ESP currently has modules for notifiable diseases, chronic diseases, influenza-like illness (ILI), and vaccine adverse event surveillance. ESP is installed locally and operates behind the host practice's firewall. It is EHR agnostic; data from any EHR system can be extracted and stored in the ESP data model.

PopMedNet is a software application that facilitates the creation and operation of distributed health data querying networks. It has sophisticated security and access control features, query creation, distribution, and response workflows, and a modular design that allows implementation of new features without disrupting the underlying data networking architecture. PopMedNet includes a secure web portal that handles query creation, distribution, response, and aggregation. Clinical practices install a client application behind their firewall, or connect via a secure virtual private network. The client application downloads queries from the portal, manages local execution of the query, allows the clinical practice the option to review query results before release, and returns approved query results to the portal. Data partners can choose to review all queries before execution and all results before release, or they can choose to allow some queries to run automatically. Automatic execution is customizable

based on specific requesters, specific query types, or specific projects. PopMedNet currently supports several large-scale distributed networks.²⁰⁻²³

In the combined implementation of ESP and PopMedNet (ESPnet), ESP generates standardized tables of EHR data that can then be queried using PopMedNet's query interfaces. The advantage of using ESP to generate tables for PopMedNet is that it provides a generalizable solution compatible with different kinds of EHRs and it makes ESP's case finding algorithms for communicable diseases, diabetes, obesity, asthma, smoking, and other conditions available for queries.

MDPHnet FEATURES

MDPHnet has several important features that together create a platform that meets the security, privacy, and confidentiality needs of clinical data partners while providing MDPH with a powerful new surveillance tool. Most importantly, MDPHnet allows MDPH epidemiologists to easily query up-to-date (next day) clinical data for a variety of purposes.

The MDPHnet secure portal lets MDPH epidemiologists create queries in 2 different ways. The Query Composer tool is a menu-driven interface that allows users to construct queries using check boxes and dropdown menus. The Query Composer lets users identify populations using *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)*²⁴ codes or a dropdown list of high priority conditions flagged by ESPnet's disease algorithms.^{20,21} The dropdown list of predefined conditions currently includes ILI, asthma, type 1 diabetes, type 2 diabetes, prediabetes, and gestational diabetes (GDM). Users

can create multiple patient "cohorts" that can be combined using inclusion or exclusion criteria. For example, a user could identify all patients with diabetes in the past ten years who have had ILI in the past 30 days. Queries can be stratified by gender, race, age group, smoking status, zip code, health center, and time period (week, month, year). Queries can also be stratified and aggregated by 3, 4, or 5 digit *ICD-9-CM* code.

When the Query Composer is not sufficient, users can develop custom SQL queries that can execute across the full range of structured EHR data captured by ESP. Users can schedule both Query Composer and custom SQL queries for distribution and execution at defined intervals. The system also includes tools that allow users to extrapolate rates from the current MDPHnet population to the state at large and to selected towns by adjusting for differences in age, gender, and race/ethnicity distributions in MDPHnet versus the state or target town as appropriate.

PARTICIPANTS AND GOVERNANCE

An Advisory Panel responsible for overseeing all aspects of the network governs MDPHnet. This Advisory Panel includes representatives from each participating organization: MDPH, Atrius Health, Cambridge Health Alliance, the Massachusetts League of Community Health Centers, the Department of Population Medicine of Harvard Medical School and Harvard Pilgrim Health Care Institute, the Massachusetts eHealth Institute (MeHI), and the informatics vendors that support MDPHnet. The Advisory Panel defines the purpose and scope of the network, sets the rules for

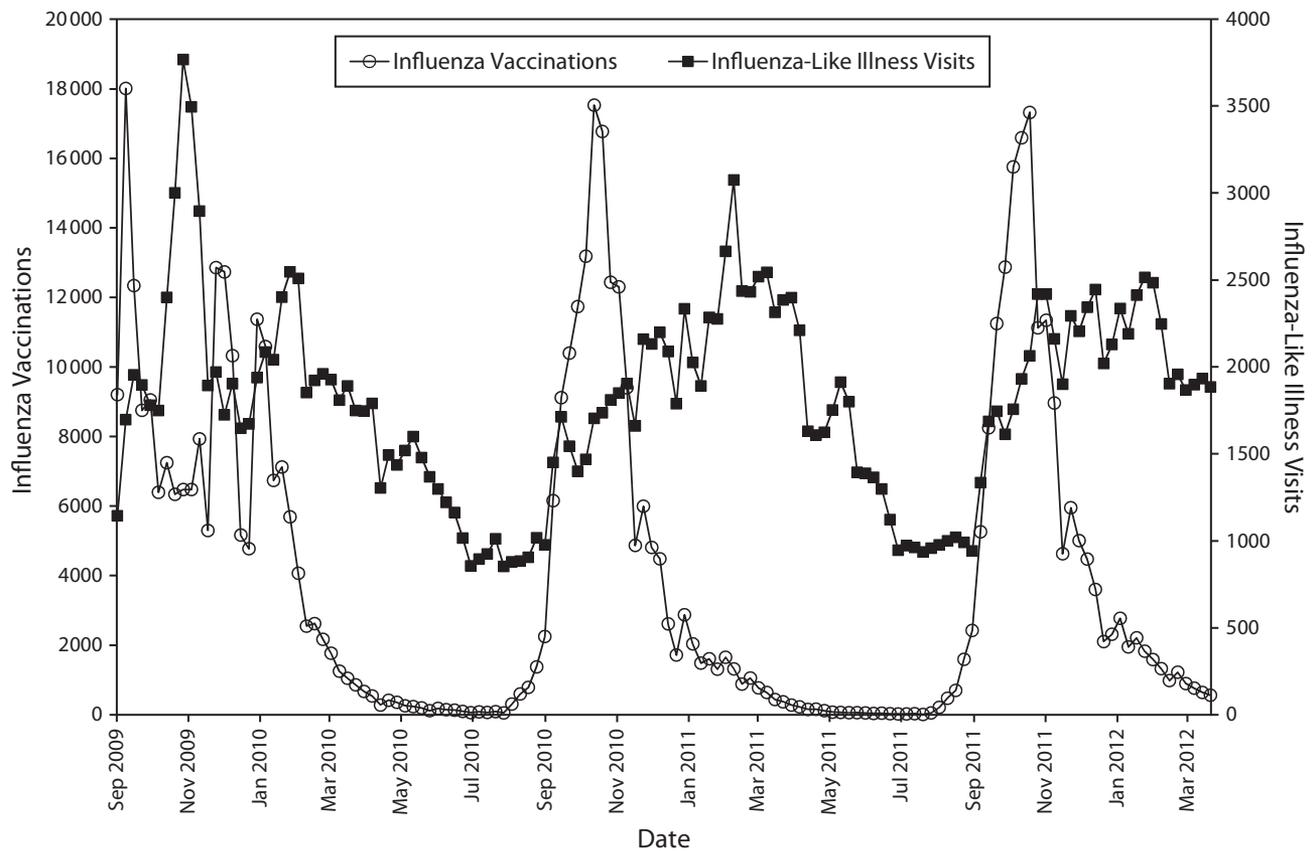
engagement between participants, works to secure funding, and facilitates communication between partners. The Panel is responsible for reviewing any requests to change the scope, access controls, or allowable activities for the network. The Advisory Panel meets in person regularly with additional communication by e-mail. Their deliberations are informed by written rules of governance.

PROOF-OF-CONCEPT ASSESSMENTS

To demonstrate ways in which MDPH epidemiologists are making use of MDPHnet, we describe 2 proof-of-concept projects. The first focuses on ILI surveillance and the second on diabetes testing rates.

For the first proof-of-concept project, we assessed whether influenza vaccination and ILI case counts mirror expected seasonal trends. These data are of great interest to the health department. Tracking vaccine distribution patterns helps inform vaccine supply management. Tracking ILI can help identify disease clusters, estimate disease burden, flag high-risk populations, inform prevention and management campaigns, and track the impact of interventions. We assessed weekly influenza vaccine and ILI counts over 3 influenza seasons, spanning September 2009 to March 2012 (Figure 1). ILI cases were defined on the basis of suggestive combinations of *ICD-9-CM* codes and fever.¹⁶ As expected, vaccine counts peaked during the fall and ILI counts peaked during the winter each year.

For diabetes, we investigated the impact of a public service announcement (PSA) that MDPH sponsored for the month of June 2011. The PSA encouraged young Hispanic women to get tested for diabetes. We used



Note. Data were gathered and aggregated from the electronic health records of Atrius Health and the Massachusetts League of Community Health Centers using MDPHnet.

FIGURE 1—Influenza vaccine and influenza-like-illness visit counts: Massachusetts, September 2009–March 2012.

MDPHnet to look for an increase in the count of Hemoglobin A1C tests among Hispanic women of childbearing age in the 8 months following the announcement compared with the 6 months before the announcement. After conducting a time-series analysis, we found no difference in diabetes testing counts among Hispanic women before versus after release of the PSA (Figure 2). Paradoxically, the lack of evidence of the effectiveness of the PSA was a victory for MDPHnet because it was proof of the Department's new-found capacity to use EHR data to evaluate public health initiatives. Historically, evaluations have required chart reviews or surveys by

telephone or mail, both of which can be extremely costly, slow, and, in the case of survey work, less accurate compared with medical record review. With MDPHnet, the health department can now assess amenable interventions rapidly and electronically. We anticipate that MDPH's new capacity to query EHR data will help the department better evaluate the relative effectiveness of its different programs and use this information to allocate limited resources more effectively.

BENEFITS TO MDPH

MDPHnet has several major advantages for MDPH as compared

with currently available capabilities. It can help detect and follow both predictable (e.g., ILI) and unpredictable (e.g., cancer clusters) events, facilitate rapid epidemiological investigations, and help evaluate the impact of public health interventions. MDPHnet allows the department access to the breadth of data contained in EHRs, including many data streams not previously available for analysis such as body mass index, blood pressure, liver function tests, urine microalbumin tests, smoking rates, and hemoglobin A1C levels. ESPnet's custom disease detection algorithms further enrich the system by providing sophisticated case detection

and classifications that are not routinely available to public health. For example, ESPnet's diabetes detection algorithm can accurately distinguish between type 1 and type 2 diabetes.¹⁷

The modular design and flexible nature of ESPnet make it possible for public health agencies to easily add features and capabilities to the system. A public health department using ESPnet could, for example, develop and implement sophisticated custom reports for monitoring the incidence, prevalence, management, and outcomes associated with priority conditions such as asthma, diabetes, obesity, hypertension, and smoking. Health departments

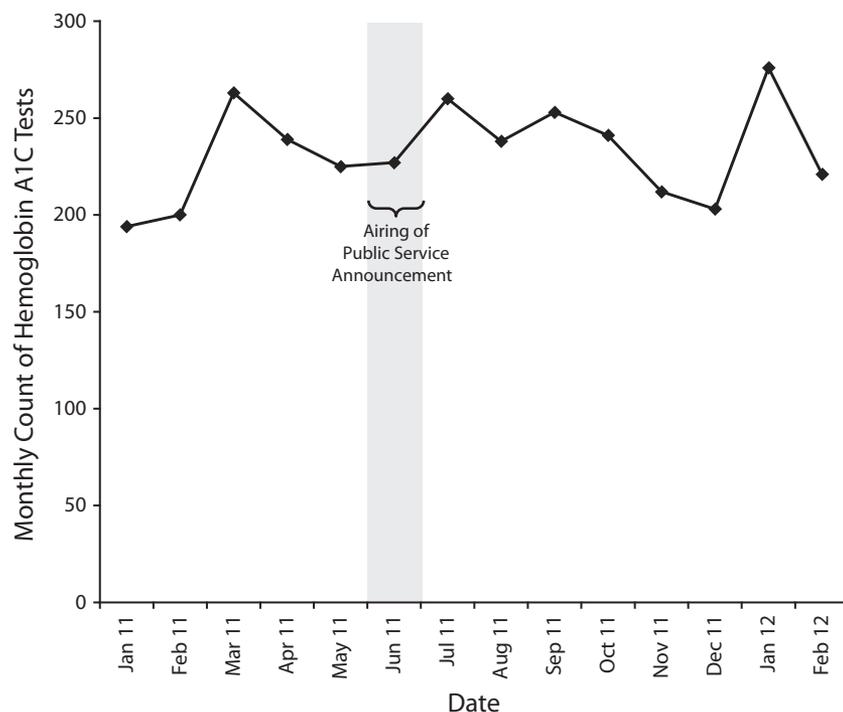


FIGURE 2—Hemoglobin A1C test counts among Hispanic women aged 18–45 years prior to and following the airing of a public service announcement encouraging diabetes testing: Massachusetts, September 2009–March 2012.

could also create reports tracking the prevalence of selected conditions along with the proportions of patients prescribed key medications, their incidence of urgent care visits, and their rates of hospitalization. An example would be prescribing rates for corticosteroid inhalers in children with asthma. These reports could be stratified by age group, gender, race/ethnicity, or geographic location and structured to provide monthly or quarterly rates so as to facilitate time series analyses. The reports could be run in parallel for different regions to facilitate comparative analyses. Other potential surveillance targets could include body mass index, blood pressures, diabetes testing rates, average hemoglobin A1C levels, microalbuminuria screening rates, use of

nephroprotective medications in diabetic patients, depression screening rates, diagnoses of attention deficit hyperactivity disorder, smoking rates, prescriptions for smoking cessation aids, influenza vaccination rates in pregnant women, antibiotic prescriptions for otitis media or bronchitis, and the year-on-year changes in each of these parameters, again stratified by age group, gender, race/ethnicity, or neighborhood.

The timeliness of EHR data are a major advantage for monitoring new health programs because it facilitates evaluation within rather than following the life cycles of new programs. This is a particular strength during the deployment of new interventions where real-time monitoring could be used to inform midcourse corrections and

identify further opportunities for more targeted implementation. Telephone surveys and claims data analyses, by contrast, can take months to years before results are available, making it difficult or impossible to use those data to modify interventions within a program's life cycle. MDPHnet data is available within a day following the clinical encounter.

BENEFITS TO PARTNER PRACTICES

Participating in MDPHnet should be beneficial to clinical partners as well as to the health department. Clinical practices currently devote considerable time and effort to generating and submitting reports for various government agencies. MDPHnet

can simplify this process. ESPnet already has the capacity to detect and report a number of notifiable disease including chlamydia, gonorrhea, syphilis, acute hepatitis A, acute hepatitis B, acute hepatitis C, lyme, pertussis, giardia, and active tuberculosis. The health department has the potential to simplify future reporting requests by developing and distributing custom queries to run automatically against practices' EHR data, reducing the time and cost for practices to respond to new reporting requests. This centralized querying approach should also achieve greater consistency in reports because clinical partners would execute the same algorithms against similarly structured data.

MDPHnet can also help practices fulfill some meaningful use criteria along with their associated financial incentives and future penalties. MDPHnet participants in Massachusetts, for example, are already eligible for meaningful use credit for specialized registries. Other meaningful use options that MDPHnet may be able to help practices fulfill include electronic laboratory reporting, syndromic surveillance, vaccine registries, and monitoring care for patients with high priority conditions such as diabetes.

Finally, practices can take advantage of MDPHnet's Query Composer tool themselves to quickly and easily explore their own data. Ultimately, participation in MDPHnet is good public citizenship that demonstrates practices' commitment to their communities and willingness to engage with public health agencies to try to improve the health of society.

LIMITATIONS

Surveillance using EHR data does have important limitations.

Data on health behaviors (e.g. exercise, nutrition, seat belt use, sun exposure, high-risk sexual encounters) are either not recorded in EHRs or are recorded as unstructured text that cannot easily be queried. EHRs only include information on the subset of the population that seeks clinical care, and they can only provide information that is documented by providers and that can be standardized for querying. Data are typically missing or incomplete for patients that seek care outside of the system. Data may also be incomplete for events that tend to be poorly documented in EHRs, such as falls. Surveillance using data from a single EHR or a small network of practices may provide a biased picture of statewide health because of non-random coverage of a relatively small fraction of the state population. Finally, patients that seek care from multiple practices may be double-counted. These limitations may diminish over time as more practices adopt EHRs and as EHRs become more functional and interoperable in response to the federal government's meaningful use incentives.²⁵

There are also important limitations inherent in using distributed networks for public health surveillance. Clinical practices' participation in the network, and their response to any specific query, is entirely voluntary. Practices have the right to withdraw from the network at any time, which in turn can threaten the breadth, stability, and reproducibility of surveillance. There is no centralized standing repository of clinical data hence the departure of a practice from the network excludes their data from future queries and precludes the possibility of repeating or refining prior

queries. This can complicate longitudinal comparisons.

SUSTAINABILITY

There are several options for funding a network like MDPHnet. These include integrating the network into the state's health information exchange (which clinical partners pay to use), adding support for additional meaningful use options (which might increase the value proposition for clinical partners), using current departmental funds earmarked for surveillance and evaluation, or negotiating new appropriations from the state government. A parallel option is to amortize the fixed costs of maintaining the network infrastructure over several separate public health networks. This might decrease each health department's software development and maintenance costs for hosting, maintenance, and support of the network, and help spread the costs of developing new reports and other capabilities.

GENERALIZABILITY

Both ESP and PopMedNet were designed to facilitate generalizability. ESP is compatible with any EHR capable of exporting structured data either as carat delimited text or as HL7 messages. Likewise, PopMedNet is agnostic to the type and source of data that it queries. The key requirement is simply that the tables exposed to PopMedNet must conform to a common format that can be targeted by distributed queries.

The greater challenge in organizing a network like MDPHnet is eliciting practices' agreements to participate, ensuring a stable funding stream, securing agreements between practices and informatics vendors, programming

and validating data extracts, standardizing the contents of different data fields, validating disease identification algorithms, and establishing rules of governance agreeable to all involved parties. National programs such as meaningful use and the Query Health Initiative may help overcome some of the technical hurdles to building networks like MPDHnet but significant logistical, financial, and governance rules still need to be tackled.¹⁵

OPPORTUNITIES

Notwithstanding MDPHnet's considerable current capacity to improve operational surveillance and programmatic evaluation, there are many potential options to further develop MDPHnet. These include adding additional clinical partners to enhance coverage of the state population, creating a library of specialized queries tailored specific public health priorities, developing visualization tools to help map disease prevalence and highlight differences between communities and demographic groups, and incorporating statistical tools to identify outlier populations and assess the impact of public health interventions more rigorously.

In addition, there are significant potential benefits to be had by linking MDPHnet to other distributed networks.¹⁵ Several large distributed health data networks have been implemented over the past few years using the PopMedNet platform. These include the Food and Drug Administration Mini-Sentinel, the National Institutes of Health/Health Care System Research Collaboratory Distributed Research Network, and the PCORnet Distributed Research Network.²⁰⁻²³ These networks collectively contain health

data from claims and EHRs on many millions of patients. The potential exists to link MDPHnet to all these networks because they all share similar data structures and all use PopMedNet as their distributed querying platform. This would enable, for example, MPDH epidemiologists to submit queries to other networks, or for other networks to submit queries to MDPHnet partners. This type of cross-network interoperability is technically simple through PopMedNet, but requires clear governance rules and agreements. Linking networks could expand coverage within jurisdictions and allow agencies to do comparative analyses between jurisdictions. It would further allow agencies to share complex queries that have been developed locally but may be of interest more broadly, in essence, crowd-sourcing analytics.

Finally, given MDPHnet's current breadth of coverage, the Department plans to use MDPHnet to help evaluate the Massachusetts Prevention and Wellness Trust. The Massachusetts Prevention and Wellness Trust is a first-of-its-kind initiative that provides substantial resources to community-based organizations and clinical practices to foster partnerships to advance the health of their communities. MDPHnet is ideally suited to assess the impact of interventions supported by the Fund on clinical parameters and outcomes.

CONCLUSIONS

MDPHnet is a groundbreaking tool for MPDH in particular and for public health agencies in general. Distributed networking technology has succeeded in overcoming clinical practices' traditional concerns regarding privacy and data control that have

prevented them from sharing their data in the past. The capacity to provide health departments with routine access to large amounts of EHR data from diverse practices for both routine and custom queries is novel and unprecedented. These data will allow more timely and granular measurements of health care processes and outcomes for large populations. Rapid and detailed analyses of public health interventions as they're still unfolding will allow health departments to refine interventions before they're complete to better ensure their ultimate success. ■

About the Authors

Joshua Vogel and Thomas Land are with the Office of Health Information Policy and Informatics, Massachusetts Department of Public Health, Boston. Jeffrey S. Brown, Richard Platt, and Michael Klompas are with the Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA.

Correspondence should be sent to Michael Klompas MD, MPH, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, 133 Brookline Avenue, 6th Floor, Boston, MA 02215 (e-mail: mklompas@partners.org). Reprints can be ordered at <http://www.ajph.org> by clicking the "Reprints" link.

This article was accepted May 12, 2014.

Contributors

All authors contributed to the conception and design of this article. J. Vogel, M. Klompas, and J. Brown wrote the first draft. All authors helped revise the article and approved the final version.

Acknowledgments

Funding for this work was provided by the Office of the National Coordinator for Health Information Technology (contract 90HT0038/01) and the Centers for Disease Control and Prevention (grant 1P01HK00088).

The authors wish to acknowledge the members of the MPDHnet advisory panel including Laurance Stuntz and Keely Benson from the Massachusetts eHealth Institute; Michael Lee, MD, from Atrius Health; Brian Herrick, MD, Michelle Weiss, and Sihui Li from the Cambridge

Health Alliance; Ellen Hafer and Mark Josephson from the Massachusetts League of Community Health Centers; Michael Sullivan and Bruce Swan from Lincoln Peak Partners; and Catherine Rocchio, Richard Schaaf, and David Fram from Commonwealth Informatics. The authors also thank Paul Oppedisano, Saul Franklin, James Buszkiewicz, and Andrea Pinzon for assistance in collecting the data for this article.

Human Participant Protection

This study was reviewed and approved by the institutional review board of Harvard Pilgrim Health Care Institute.

References

- Klompas M, McVetta J, Lazarus R, et al. Integrating clinical practice and public health surveillance using electronic medical record systems. *Am J Public Health*. 2012;102(suppl 3):S325–S332.
- 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 Morb Mortal Wkly Rep*. 2008;57(14):373–376.
- Allen-Dicker J, Klompas M. Comparison of electronic laboratory reports, administrative claims, and electronic health record data for acute viral hepatitis surveillance. *J Public Health Manag Pract*. 2012;18(3):209–214.
- Calderwood MS, Platt R, Hou X, et al. Real-time surveillance for tuberculosis using electronic health record data from an ambulatory practice in eastern Massachusetts. *Public Health Rep*. 2010;125(6):843–850.
- Overhage JM, Grannis S, McDonald CJ. A comparison of the completeness and timeliness of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. *Am J Public Health*. 2008;98(2):344–350.
- Buck MD, Anane S, Taverna J, Amirfar S, Stubbs-Dame R, Singer J. The Hub Population Health System: distributed ad hoc queries and alerts. *J Am Med Inform Assoc*. 2012;19(e1):e46–e50.
- Davis RL, Kolczak M, Lewis E, et al. Active surveillance of vaccine safety: a system to detect early signs of adverse events. *Epidemiology*. 2005;16(3):336–341.
- Centers for Disease Control and Prevention. State electronic disease surveillance systems—United States, 2007 and 2010. *MMWR Morb Mortal Wkly Rep*. 2011;60(41):1421–1423.
- Lazarus R, Klompas M, Campion FX, 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(1):18–24.
- Lenert L, Sundwall DN. Public health surveillance and meaningful use regulations: a crisis of opportunity. *Am J Public Health*. 2012;102(3):e1–e7.
- McMurry AJ, Murphy SN, MacFadden D, et al. SHRINE: enabling nationally scalable multi-site disease studies. *PLoS ONE*. 2013;8(3):e55811.
- Brown JS, Holmes JH, Maro JC, et al. Design Specifications for Network Prototype and Cooperative to Conduct Population-Based Studies and Safety Surveillance. Rockville, MD: Agency for Healthcare Research and Quality; 2009.
- Maro JC, Platt R, Holmes JH, et al. Design of a national distributed health data network. *Ann Intern Med*. 2009;151(5):341–344.
- Brown JS, Holmes JH, Shah K, Hall K, Lazarus R, Platt R. Distributed health data networks: a practical and preferred approach to multi-institutional evaluations of comparative effectiveness, safety, and quality of care. *Med Care*. 2010;48(6, suppl):S45–S51.
- Klann JG, Buck MD, Brown J, et al. Query Health: standards-based, cross-platform population health surveillance. *J Am Med Inform Assoc*. 2014;21(4):650–656.
- Yih WK, Cocoros NM, Crockett M, et al. Automated influenza-like illness reporting—an efficient adjunct to traditional sentinel surveillance. *Public Health Rep*. 2014;129(1):55–63.
- Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R. Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data. *Diabetes Care*. 2013;36(4):914–921.
- Klompas M, Bialek SR, Kulldorff M, Vilks Y, Harpaz R. Herpes zoster and postherpetic neuralgia surveillance using structured electronic data. *Mayo Clin Proc*. 2011;86:1146–1153.
- Klompas M, Haney G, Church D, Lazarus R, Hou X, Platt R. Automated identification of acute hepatitis B using electronic medical record data to facilitate public health surveillance. *PLoS ONE*. 2008;3(7):e2626.
- Curtis LH, Weiner MG, Boudreau DM, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. *Pharmacoepidemiol Drug Saf*. 2012;21(suppl 1):23–31.
- Behrman RE, Benner JS, Brown JS, McClellan M, Woodcock J, Platt R. Developing the Sentinel System—a national resource for evidence development. *N Engl J Med*. 2011;364(6):498–499.
- NIH Distributed Research Network. Available at: <https://http://www.nihcollaboratory.org/Pages/distributed-research-network.aspx>. Accessed February 4, 2014.
- PCORnet | National Patient-Centered Clinical Research Network. Available at: <http://pcornet.org>. Accessed February 4, 2014.
- International Classification of Diseases, Ninth Revision, Clinical Modification*. Hyattsville, Md: National Center for Health Statistics; 1980. DHHS publication PHS 80-1260.
- Blumenthal D, Tavenner M. The “meaningful use” regulation for electronic health records. *N Engl J Med*. 2010;363(6):501–504.