Use and Visualization of Electronic Health Record Data to Advance Public Health

Sharon E. Perlman, MPH

ABOUT THE AUTHOR

Sharon E. Perlman is with the Division of Epidemiology, New York City Department of Health and Mental Hygiene, Queens, NY.

See also Cocoros et al., p. 269.

The COVID-19 pandemic has underscored the importance of strong public health infrastructure and timely health surveillance data. Electronic health record (EHR) networks such as MDPHnet can help meet both emergency and routine public health data needs.

In this issue of *AJPH*, Cocoros et al. (p. 269) describe RiskScape, a unique and powerful data visualization platform that makes data easy to access and interpret by providing clear figures, maps, and tables. RiskScape also allows stratification by variables of interest and runs statistical testing in the background to determine whether observed differences are statistically meaningful. RiskScape is based on aggregated data from MDPHnet, an EHR network used to facilitate notifiable disease reporting, examine the cascade of care (e.g., for diabetes), estimate disease prevalence, and monitor the effectiveness of public health interventions. RiskScape is primarily used by the Massachusetts Department of Public Health for surveillance, planning, and advocacy.

COVID-19 highlights the potential of platforms such as RiskScape to help monitor and respond to a rapidly changing situation. The ability to quickly identify disease hotspots may be one of the most valuable uses of EHR networks, providing data that are timely enough to monitor changes essentially in real time. EHR networks can also greatly improve public health surveillance by increasing notifiable disease reporting through automation, decreasing paperwork, and streamlining laboratory testing reporting.

In New York City, two large EHR networks, the Bronx Regional Health Information Organization (Bronx RHIO) and Healthix, have integrated electronic lab reporting data and flag patients with COVID-19, alerting providers in ambulatory and inpatient settings.[1,2] Outpatient providers can also be notified when patients are hospitalized, and hospital providers can easily learn about a patient's underlying risk factors for more severe COVID-19. The Bronx RHIO is using EHR data to learn more about associations with social determinants of health (e.g., housing and neighborhood poverty) and rare outcomes, such as multisystem inflammatory syndrome.[1] Both the Bronx RHIO and Healthix have worked closely with the New York City Health Department in the COVID-19 response.

Just as RiskScape can calculate patient cardiovascular risk scores, which Cocoros et al. note can be used for prevention at the provider level and to assess the overall health of the larger population, one could imagine creating a risk score for COVID-19, making it possible to identify patients and practices with higher risk of developing severe disease. COVID-19 has had a disproportionate impact on those of Black and Hispanic race/ethnicity, older adults, and those with chronic conditions such as diabetes and heart disease.[3] Higher-risk patients and practices could be selected for public health interventions and educational campaigns to decrease transmission, reduce inequities, and encourage patients with symptoms to seek care early. Now that COVID-19 vaccines have been approved and we need to prioritize certain groups for early vaccination, EHR networks could be useful in identifying such patients and tracking the proportion that have been vaccinated. Platforms such as RiskScape could be used to visualize data on high-risk neighborhoods and to monitor vaccination efforts.

EHRs could also potentially be used to examine social determinants of health associated with increased risk for COVID-19, including certain occupations, crowded housing and congregate settings, and intergenerational households.[4] More generally, education, employment, housing stability, food security, and social support are all important predictors of health.[5] However, current data limitations may preclude such analyses. There is a lack of evidence-based, standardized patient-level measures and a need for the

development of such measures, as well as support and incentives for providers to collect this information.[6,7] Another related data limitation in exploring health inequities is missing or incorrect race/ethnicity data in the EHR.[8] This could be improved by incentivizing documentation and focusing on patient self-report of race and ethnicity.

Even in the absence of patient-level data, social determinants of health can be examined by linking EHR data with neighborhood-level data, for example, average household income, housing characteristics, access to parks, crime, and density of fast food and alcohol stores. Neighborhood-level information can provide a helpful context for health risk factors and outcomes.[7,9] Furthermore, linking EHRs to neighborhood-level resources could help providers connect patients in need to nearby food pantries, community centers, and housing assistance.

Other jurisdictions that want to use EHR data for population health estimates, as MDPHnet does, should assess the accuracy and validity of their data and whether those in the EHR network represent the larger population. MDPHnet data are from three large clinical groups that cover about 20% of the population, which is concentrated in the eastern part of Massachusetts. Previous comparisons with Behavioral Risk Factor Surveillance System (BRFSS) data and the Centers for Disease Control and Prevention's 500 Cities estimates (based on BRFSS) showed similar prevalence estimates with EHR-based data.[10] It is noteworthy that Massachusetts has unusually high insurance coverage, with only 2.9% uninsured in 2019,[11] which increases the generalizability of the MDPHnet system to the population as a whole. In states where less of the population is insured and seeking care, it is likely that greater differences will exist between EHR-derived population health estimates and the health of the general population.

Although conditions such as hypertension, diabetes, obesity, and smoking may be good candidates for EHR-based surveillance,[10,12,13] other conditions may be more difficult to monitor with EHR data. As Cocoros et al. note, patients may seek care outside an EHR network. This may particularly be true for cancer and rare diseases, for which patients often seek care in specialty hospitals or from providers specializing in certain conditions. To assess the accuracy and validity of an EHR-based health indicator, whenever possible it is important to compare EHR-based health estimates against established surveys or registries of disease. A second limitation the authors discuss is the inability of MDPHnet and some other EHR networks to de-duplicate records across practices, so a patient with hypertension seeking care from an internist and a cardiologist might be double counted in the system. This protects privacy but can affect data quality.

As Cocoros et al. mention, the future of RiskScape might include linkage to and visualization of hospitalization data, claims records, and mortality data. Clearly there are challenges in linking data from different sources, including how to link and match data, how to protect privacy, and obtaining institutional review board permission.[14] Yet such linkages could lead to a greater understanding of risk factors for disease, hospitalization and death, and the development of interventions for and targeted outreach to special populations. For COVID-19, one could potentially learn more about the progression of disease, time to seeking treatment, and risk factors for hospitalization and death.

The ability to clearly share data and communicate important messages is key. This is especially true in an urgent and rapidly changing public health situation such as the COVID-19 pandemic. Data visualization platforms are often difficult to use and have poor graphic design, reflecting the need for increased collaboration among those knowledgeable about data, data programmers, and graphic designers. RiskScape is a model for good design, easy navigation, and presentation of data in an easily interpretable way. By developing and freely sharing the code and documentation for the platform, MDPHnet researchers have provided a valuable resource for other jurisdictions. Other jurisdictions and health departments can benefit from this valuable tool but need to have a good understanding of the strengths and limitations of their own data to use it appropriately.

Access to RiskScape data is currently limited to participating practices, MDPHnet researchers, and the Massachusetts Department of Public Health. With appropriate explanation of the data and its limitations, RiskScape (or a more limited version) might be made available to researchers and the public, as both could greatly benefit from being able to access these data. In Colorado, CHORDS (the Colorado Health Observation Regional Data Service) has publicly available interactive maps of some health indicators using data from an EHR network covering the Denver metropolitan area.[15] The ability to clearly and transparently share data can help make science more accessible and increase support for public health. AJPH

CORRESPONDENCE

Correspondence should be sent to Sharon Perlman, Director of Special Projects, Division of Epidemiology, CN6, New York City Department of Health and Mental Hygiene, 42-09 28th St., Queens, NY 11101 (e-mail: sperlma1@health.nyc.gov). Reprints can be ordered at http://www.ajph.org by clicking the "Reprints" link.

PUBLICATION INFORMATION

Full Citation: Perlman SE. Use and visualization of electronic health record data to advance public health. *Am J Public Health*. 2021;111(2):180–182.

Acceptance Date: November 16, 2020.

DOI: https://doi.org/10.2105/AJPH.2020.306073

ACKNOWLEDGMENTS

The author thanks R. Charon Gwynn and Hannah Helmy for their review of this editorial.

CONFLICTS OF INTEREST

I have no conflicts of interest to disclose.

REFERENCES

1. Bronx Regional Health Information Organization. Bronx RHIO research and reporting on COVID-19. Available at: https://bronxrhio.org/COVID-19. Accessed November 14, 2020.
2. Healthix. COVID-19 activities. Available at: https://www.healthix.org/covid-19. Accessed November 14, 2020.
3. Centers for Disease Control and Prevention. Coronavirus disease 2019 case surveillance—United States, January 22–May 30, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(24):759–765. https://doi.org/10.15585/mmwr.mm6924e2
4. Figueroa JF, Wadhera RK, Lee D, Yeh RW, Sommers BD. Community-level factors associated with racial and ethnic disparities in COVID-19 rates in Massachusetts. *Health Aff (Millwood)*. 2020;39(11):1984–1992. https://doi.org/10.1377/hlthaff.2020.01040
5. Kushel MB, Gupta R, Gee L, Haas JS. Housing instability and food insecurity as barriers to health care among low-income Americans. *J Gen Intern Med*. 2006;21(1):71–77. https://doi.org/10.1111/j.1525-1497.2005.00278.x
6. Cantor MN, Thorpe L. Integrating data on social determinants of health into electronic health records. *Health Aff (Millwood)*. 2018;37(4):585–590. https://doi.org/10.1377/hlthaff.2017.1252
7. Gottlieb LM, Francis DE, Beck AF. Uses and misuses of patient- and neighborhood-level social determinants of health data. *Perm J*. 2018;22:18–78. https://doi.org/10.7812/TPP/18-078
8. Polubriaginof FCG, Ryan P, Salmasian H, et al. Challenges with quality of race and ethnicity data in observational databases. *J Am Med Inform Assoc*. 2019;26(8–9):730–736. https://doi.org/10.1093/jamia/ocz113
9. Kolak M, Bhatt J, Park YH, Padrón NA, Molefe A. Quantification of neighborhood-level social determinants of health in the continental United States. *JAMA Netw Open*. 2020;3(1):e1919928. https://doi.org/10.1001/jamanetworkopen.2019.19928
10. Klompas M, Cocoros NM, Menchaca JT, et al. State and local chronic disease surveillance using electronic health record systems. *Am J Public Health*. 2017;107(9):1406–1412. https://doi.org/10.2105/AJPH.2017.303874
11. Center for Health Information and Analysis. Findings from the 2019 Massachusetts Health Insurance Survey. Available at: https://www.chiamass.gov/assets/docs/r/survey/mhis-2019/2019-MHIS-Report.pdf. Accessed October 17, 2020.
12. Bacon E, Budney G, Bondy J, et al. Developing a regional distributed data network for surveillance of chronic health conditions: the Colorado Health Observation Regional Data Service. *J Public Health Manag Pract*. 2019;25(5):498–507. https://doi.org/10.1097/PHH.0000000000000810
13. Perlman SE, McVeigh KH, Thorpe LE, Jacobson L, Greene CM, Gwynn RC. Innovations in population health surveillance: using electronic health records for chronic disease surveillance. *Am J Public Health*. 2017;107(6):853–857. https://doi.org/10.2105/AJPH.2017.303813
14. Gliklich RE, Dreyer NA, Leavy MB, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. 3rd ed. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Available at: https://www.ncbi.nlm.nih.gov/books/NBK208618. Accessed November 21, 2020.
15. Colorado Health Institute. Maps of select CHORDS indicators. Available at: https://www.cohealthmaps.dphe.state.co.us/chords_maps/chords_webmap. Accessed November 14, 2020.