Electronic Health Record–Based Algorithm for Monitoring Respiratory Virus–Like Illness

Noelle M. Cocoros, Karen Eberhardt, Vu-Thuy Nguyen, Catherine M. Brown, Alfred DeMaria, Lawrence C. Madoff, Liisa M. Randall, Michael Klompas

Viral respiratory illness surveillance has traditionally focused on single pathogens (e.g., influenza) and required fever to identify influenza-like illness (ILI). We developed an automated system applying both laboratory test and syndrome criteria to electronic health records from 3 practice groups in Massachusetts, USA, to monitor trends in respiratory viral–like illness (RAVIOLI) across multiple pathogens. We identified RAVIOLI syndrome using diagnosis codes associated with respiratory viral testing or positive respiratory viral assays or with fever. After retrospectively applying RAVIOLI criteria to electronic health records, we observed annual winter peaks during 2015–2019, predominantly caused by influenza, followed by cyclic peaks corresponding to SARS-CoV-2 surges during 2020–2024, spikes in RSV in mid-2021 and late 2022, and recrudescent influenza in late 2022 and 2023. RAVIOLI rates were higher and fluctuations more pronounced compared with traditional ILI surveillance. RAVIOLI broadens the scope, granularity, sensitivity, and specificity of respiratory viral illness surveillance compared with traditional ILI surveillance.

Respiratory viral illnesses place an enormous burden on human health and the healthcare system (*1*–*3*). Although multiple pathogenic respiratory viruses circulate, often simultaneously, public health has traditionally dedicated most of its attention to monitoring trends in laboratory-confirmed influenza and influenza-like illness (ILI). Illness and death associated with seasonal respiratory syncytial virus (RSV) spikes, the SARS-CoV-2 pandemic, and occasional clusters of infection from other respiratory pathogens, however, illustrate the importance of expanding monitoring to include all respiratory viral–like illness activity. Relying on laboratory testing alone will not accomplish this goal because most persons with respiratory viral illnesses do not seek care, many who do seek care are not tested, and not everyone tested is tested for all respiratory viruses.

Public health agencies have traditionally relied on syndromic surveillance to monitor conditions for which testing rates are low and variable (*4*). The Centers for Disease Control and Prevention's outpatient Influenza-like Illness Surveillance Network and emerging systems designed to monitor COVID-19–like illness are prime examples (*5*–*9*). However, syndromic surveillance systems tend to provide little or no information about which particular pathogens are circulating, and most jurisdictions require fever to define ILI, a requirement that increases specificity but lowers sensitivity (fever occurs in fewer than half of persons with laboratory-confirmed influenza) (*10*). Surveillance focusing on single pathogens (e.g., influenza, SARS-CoV-2), viral testing alone, or syndromic definitions alone provides an incomplete picture of respiratory illness activity and can miss critical trends and developments (*11*,*12*). Extending surveillance to include multiple pathogens, using both laboratory testing and syndromes, and decreasing reliance on fever as a gatekeeper symptom are necessary to provide public health agencies and healthcare institutions with the data needed to improve situational awareness for planning, resource use, internal and external communications, and targeted prevention activities.

To regularly monitor overall respiratory viral illness activity associated with multiple pathogens, we developed an integrated surveillance strategy using a combination of laboratory and syndromic indicators, incorporating logic to identify the relative contributions of different individual pathogens. We describe

Author affiliations: Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA (N.M. Cocoros, V.-T. Nguyen, M. Klompas); Harvard Medical School, Boston (N.M. Cocoros, M. Klompas); Commonwealth Informatics, Waltham, Massachusetts, USA (K. Eberhardt); Massachusetts Department of Public Health, Boston (C.M. Brown, A. DeMaria, L.C. Madoff, L.M. Randall); Brigham and Women's Hospital, Boston (M. Klompas)

DOI: https://doi.org/10.3201/eid3006.230473

our data-driven approach to developing a routine, automated respiratory virus-like illness (RAVIOLI) algorithm for syndromic surveillance in Massachusetts using live electronic health record (EHR) data drawn from 3 large practice groups. Our work was performed as public health surveillance and therefore not subject to institutional review board oversight.

Methods

We used the Electronic Medical Record Support for Public Health (ESP, https://www.esphealth.org) public health surveillance platform to develop the RAVIOLI algorithm. ESP is open-source software that uses automated daily extracts of EHR data to identify and report conditions of public health interest to health departments (*13*–*17*). ESP maps raw data to common terms and then applies algorithms to identify conditions using diagnosis codes, prescriptions, laboratory tests, and vital signs. In Massachusetts, ESP is used for automated reporting of infectious disease cases to the Massachusetts Department of Public Health, aggregate reporting of chronic diseases, and continuum-of-care assessments (*18*–*21*).

Three multisite clinical practice groups that use ESP for infectious disease reporting, Atrius Health, Cambridge Health Alliance, and Boston Medical Center, contributed data for our project. Atrius Health (https://www.atriushealth.org) is an ambulatory care group with >30 locations in eastern Massachusetts that provides clinical services for a population of ≈700,000. Cambridge Health Alliance (https://www.challiance.org) is a safety-net system that provides ambulatory and inpatient care to >140,000 patients in communities north of Boston. Boston Medical Center (https://www.bmc.org) is a 514-bed academic medical center and safety-net hospital that provides ambulatory and inpatient care to ≈220,000 persons. We combined data from those 3 sites for this analysis.

We sought to develop an evidence-based set of diagnosis codes to identify respiratory virus–like illnesses and assess whether a subset of those codes might be predictive of specific pathogens. To identify codes associated with respiratory viral illness syndrome, we identified all patients tested for respiratory viruses (Table 1) during October 3, 2015–July 30, 2022. Among patients who tested positive for ≥1 virus, we identified all International Classification of Diseases, 10th Revision (ICD-10), diagnosis codes recorded within 2 days before or after the specimen collection date. For patients without a recorded specimen collection date, we used the test order date; if that was unavailable, we used the result date. We manually removed ICD-10 codes unrelated to respiratory viral illness (e.g., trauma, cancer, chronic disease management). The list of >7,000 excluded codes is available upon request from the authors.

Table 1. Respiratory pathogens and test types included in RAVIOLI algorithm for monitoring respiratory virus–like illness*

Pathogen	Test types
Adenovirus	NAAT
Non–SARS-CoV-2 coronaviruses: OC43,229E, HKU1, NL63	NAAT
Human metapneumovirus	NAAT
Influenza	NAAT, antigen/rapid, culture
Parainfluenza	NAAT
Respiratory syncytial virus	NAAT, antigen
Rhinovirus/enterovirus	NAAT
SARS-CoV-2	NAAT, antigen/rapid

*Respiratory virus–like illness is defined as a clinical encounter with a positive laboratory test result for a respiratory virus, as shown in this table; 1 of the International Classification of Diseases, 10th Revision, diagnosis codes shown in Table 1; or a measured fever >100°F. NAAT, nucleic acid amplification test.

We calculated the positive predictive value (PPV) for each ICD-10 code associated with positive respiratory virus test results. We also calculated the PPV for measured temperature >100°F within 2 days before or after a positive respiratory virus test. We calculated the PPV for each ICD-10 code and fever as the number of encounters with the diagnosis code within 2 days of a positive test divided by the total number of times the diagnosis code occurred across all clinical encounters during the study period. We defined a clinical encounter as a patient receiving a relevant diagnosis code, immunization, vital sign measure, laboratory test, or prescription.

We included in the final algorithm diagnosis codes with a PPV ≥10% for any respiratory virus (all viruses combined) or for a specific individual respiratory virus. We also included encounters with positive respiratory virus tests in the total count of respiratory virus encounters as well as in virus-specific categories of RAVIOLI. We counted each viral encounter only 1 time if the patient had both a positive respiratory virus assay result and ≥1 suggestive diagnosis code. We classified measured fever alone and diagnosis codes with a PPV of ≥10% for any positive respiratory virus test but <10% for any specific respiratory virus in a category referred to as nonspecific for respiratory viral illness syndrome. In summary, we categorized positive cases within RAVIOLI as virus-specific (e.g., influenza, adenovirus), based on a positive test or a diagnosis code with a PPV ≥10% for the specific virus, or nonspecific, based on fever or a diagnosis code with a PPV ≥10% for any positive test of interest.

To better understand the underlying data included in the final RAVIOLI algorithm, we examined the proportion of patients in each virus-specific

category of the algorithm with a positive laboratory test and the proportion of patients in the nonspecific category with a fever. We generated weekly counts during October 3, 2015–January 13, 2024, for clinical encounters with patients meeting the RAVIOLI algorithm, overall and stratified by the probable etiology when possible. For comparison, we also identified the proportion of patients that met the definition of ILI: fever and a diagnosis code for any influenza-like symptom or diagnosis; fever was identified by either a measured fever >100°F or diagnosis code for fever (Appendix Table 1, https://wwwnc.cdc.gov/EID/article/30/6/23-0473-App1.pdf).

Results

Forty-two diagnosis codes (Table 2) and measured fever (>100°F) had a PPV ≥10% for either any positive respiratory virus test (nonspecific) or ≥1 virus-specific positive test; those diagnosis codes and fever are included in the RAVIOLI algorithm. We recorded weekly counts of patients with clinical encounters and calculated the proportion that met the definition

Table 2. ICD-10 diagnosis codes that met the positive predictive value threshold for confirmed respiratory viral illnesses and are included in the RAVIOLI algorithm for monitoring respiratory virus–like illness*

Virus	ICD-10 codes†	Description
Adenovirus	A08.2	Adenoviral enteritis
	B34.0	Adenovirus infection, unspecified
	B97.0	Adenovirus as the cause of diseases classified elsewhere
	J12.0	Adenoviral pneumonia
Non–SARS-CoV-2 coronavirus	B34.2	Coronavirus infection, unspecified
SARS-CoV-2	B34.2	Coronavirus infection, unspecified
	B97.29	Other coronavirus as the cause of diseases classified elsewhere
	J12.82	Pneumonia associated with coronavirus disease 2019
	J12.89	Other viral pneumonia
	J80	Acute respiratory distress syndrome
	R05.1	Acute cough
	R48.1	Agnosia
	U07.1	COVID-19
Human metapneumovirus	B97.81	Human metapneumovirus as the cause of diseases classified elsewhere
	J12.3	Human metapneumovirus pneumonia
	J21.1	Acute bronchiolitis associated with human metapneumovirus
Influenza	J09.X1	Influenza from identified novel influenza A virus with pneumonia
	J09.X2	Influenza associated with identified novel influenza A virus with other respiratory manifestations
	J10.00	Influenza associated with other identified influenza virus with unspecified type of pneumonia
	J10.1	Influenza associated with other identified influenza virus with other respiratory manifestations
	J11.00	Influenza associated with unidentified influenza virus with unspecified type of pneumonia
	J11.1	Influenza associated with unidentified influenza virus with other respiratory manifestations
Parainfluenza	B33.8	Other specified viral diseases
	B34.8	Other viral infections of unspecified site
	J20.4	Acute bronchitis associated with parainfluenza virus
Rhinovirus and enterovirus	B34.0	Adenovirus infection, unspecified
	B34.8	Other viral infections of unspecified site
	B97.10	Unspecified enterovirus as the cause of diseases classified elsewhere
	J20.6	Acute bronchitis associated with rhinovirus
	J45.902	Unspecified asthma with status asthmaticus
Respiratory syncytial virus	B97.4	Respiratory syncytial virus as the cause of diseases classified elsewhere
	J12.1	Respiratory syncytial virus pneumonia
	J20.5	Acute bronchitis associated with respiratory syncytial virus
	J21.0	Acute bronchiolitis associated with respiratory syncytial virus
Any respiratory viral test (nonspecific)	J21.8	Acute bronchiolitis associated with other specified organisms
	R06.03	Acute respiratory distress
	P81.9	Disturbance of temperature regulation of newborn, unspecified
	J12.9	Viral pneumonia, unspecified
	R50.81	Fever manifesting with conditions classified elsewhere
	J96.90	Respiratory failure, unspecified, unspecified whether with hypoxia or hypercapnia
	R05.9	Cough, unspecified
	J96.91	Respiratory failure, unspecified with hypoxia
	J96.92	Respiratory failure, unspecified with hypercapnia
	R57.9	Shock, unspecified

*Respiratory virus-like illness is defined as a clinical encounter with a positive laboratory test result for a respiratory virus listed in Table 1; 1 of the ICD-10 diagnosis codes listed in this table; or a measured fever >100°F. ICD-10, International Classification of Diseases, 10th Revision.
†All of the diagnosis codes in the table had a positive predictive value ≥10% PPV for either any positive respiratory virus laboratory test or 1 of the virus-specific positive tests.

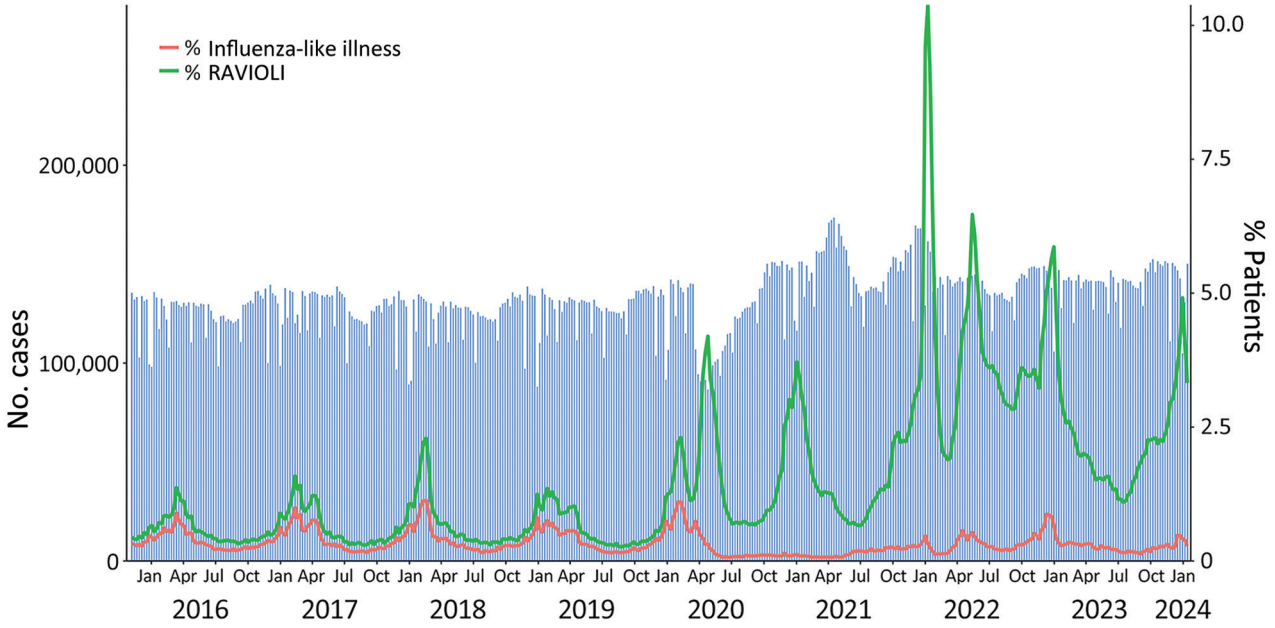


Figure 1. Numbers of patients with a clinical encounter for respiratory virus–like illness and the percentages that met the requirements for influenza-like illness versus those of the RAVIOLI algorithm for monitoring respiratory virus–like illness, by week, Massachusetts, USA, October 2015–January 2024. Patients receiving a diagnosis code, immunization, vital sign measure, laboratory test, or prescription were considered to have a clinical encounter.

for RAVIOLI overall (diagnosis code, fever, or positive respiratory virus test) and, for comparison, the proportion that met the ILI criteria (Figure 1). The percentage of encounters that met the RAVIOLI algorithm showed clear seasonal trends of annual winter spikes during 2015–2019 followed by periodic increases during spring 2020–early 2024, corresponding to emergence or surges of SARS-CoV-2, RSV, and influenza in Massachusetts. RAVIOLI was identified in a much larger proportion of encounters than ILI after March 2020 and, at times (e.g., fall 2021, August–November 2023), ILI did not detect an increase in respiratory virus illness while RAVIOLI did.

We estimated weekly counts of patients with clinical encounters meeting the RAVIOLI algorithm stratified by encounters with virus-specific or nonspecific encounters without a classified virus. We calculated those data for the full study period, October 2015–January 2024 (Figure 2, panel A), and for January 2020–January 2024 (Figure 2, panel B). Before March 2020, most RAVIOLI encounters came from the influenza or nonspecific categories. SARS-CoV-2 subsequently dominated until fall 2021, when the nonspecific category reemerged, along with influenza and RSV. When we examined trends by patient age groups, the highest proportion of encounters that met the RAVIOLI algorithm were among children 0–4 years of age, followed by young persons 5–24 years of age (Figure 3).

Data from January 2023–January 2024 show the proportions of patients in the COVID-19, influenza, and RSV categories with a positive laboratory test versus diagnosis code, as well as the proportion in the nonspecific category with fever (Appendix Table 2). The proportion with a positive test varied by virus and time; patients in the COVID-19 category were least likely and those in the RSV category most likely to have a positive laboratory test. Among patients in the nonspecific category, one third or fewer had evidence of fever, and most were identified by a diagnosis code. We also determined the proportion of RAVIOLI patients identified on the basis of ≥1 positive laboratory test, diagnosis code, or fever during January 2021–January 2024 (Appendix Figure 1); RAVIOLI patients can meet ≥1 criterion (e.g., have both a positive laboratory test and a diagnosis code). Diagnosis codes were the most common element contributing to identification in most weeks, followed by positive laboratory tests and fever.

Discussion

Respiratory viruses continue to impose a high burden on patients, healthcare providers, and society, and multiple pathogens, including SARS-CoV-2, influenza, RSV, and others, contribute to the burden of respiratory illnesses. Both healthcare providers and public health agencies therefore have an interest in having access to timely and granular data on

trends in respiratory viral illnesses and contributing pathogens. We developed an EHR-based algorithm for integrated surveillance of respiratory virus illness syndromes and associated pathogens using historical data to identify diagnosis codes and other characteristics of healthcare visits most predictive of confirmed respiratory viral illnesses. The RAVIOLI algorithm comprises positive laboratory tests, evidence-based diagnosis codes, and measured fever.

We have implemented RAVIOLI surveillance within the ESP automated public health surveillance platform to provide the Massachusetts Department of Public Health and participating practices with weekly reports on RAVIOLI incidence and contributing pathogens. RAVIOLI provides the department

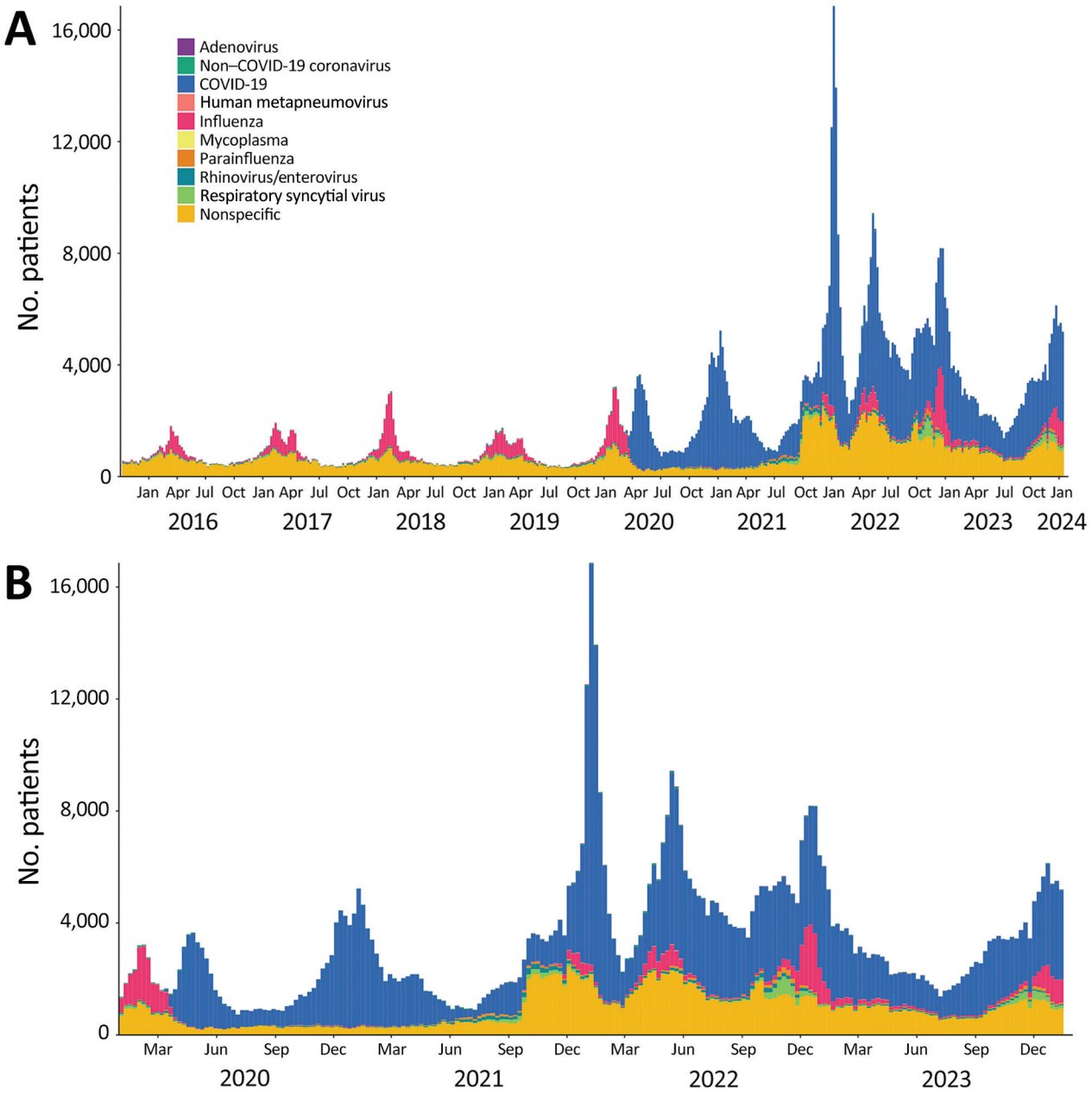


Figure 2. Numbers of patients that met the requirements for the RAVIOLI algorithm for monitoring respiratory virus–like illness, by pathogen category and week, Massachusetts, USA, October 2015–January 2024. A) October 2015–January 2024; B) January 2020–January 2024. Within each virus-specific category are counts of positive test results and diagnosis codes with a positive predictive value (PPV) ≥10% for that specific pathogen. The nonspecific category includes diagnosis codes with a PPV of ≥10% for any positive respiratory viral assay but PPV of <10% for any specific respiratory virus and includes measured fever >100°F.

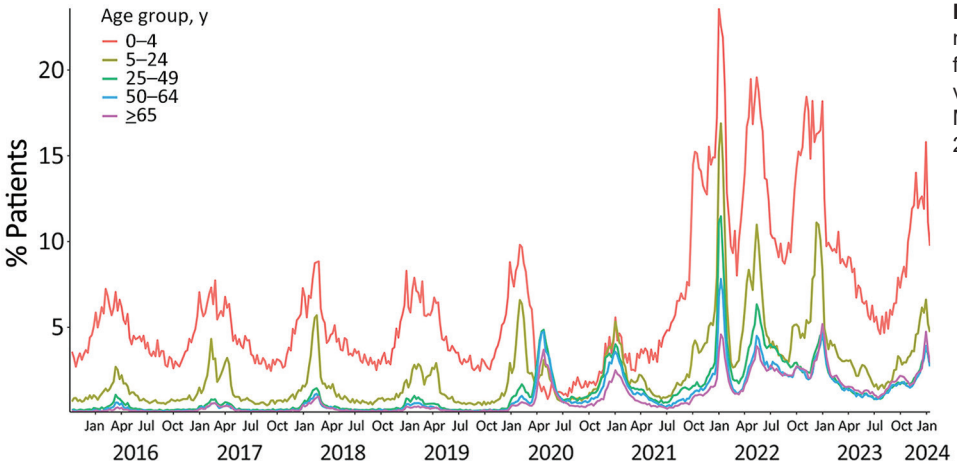


Figure 3. Percentage of patients meeting the RAVIOLI algorithm for monitoring respiratory virus–like illness, by age group, Massachusetts, USA, October 2015–January 2024.

and practices with granular insight into evolving trends in respiratory viral illness rates that both retains the best features of traditional syndromic surveillance (capacity to monitor changes in disease incidence in near real time regardless of whether persons get tested) and simultaneously broadens the scope of surveillance to include multiple pathogens, not just influenza and SARS-CoV-2. The data provide insight into the relative proportions of contributing pathogens across multiple clinical facilities using both test results and diagnosis codes to identify organisms.

When implemented well, syndromic surveillance provides a picture of the frequency, intensity, and trends in indicators of infectious and noninfectious conditions at local and extended scales. Integrating available viral pathogen test results, even if only in a subset of the population under surveillance, as we have done with the RAVIOLI algorithm, can add information about what is or is not contributing to observed increases in respiratory viral activity. Although influenza-like illness and COVID-like illness surveillance have been critical components for monitoring influenza and COVID-19 activity, reliance on fever as a required component of syndromic definitions is problematic because fever occurs only in a minority of laboratory-confirmed influenza and SARS-CoV-2 cases (*22*–*24*). Syndromic surveillance algorithms that require fever can therefore miss critical trends in the incidence of illnesses (*9*). The RAVIOLI algorithm, in contrast, does not require fever as a criterion and uses both laboratory test results and an evidence-based set of diagnosis codes to increase both sensitivity and specificity.

Limitations of RAVIOLI surveillance include its development in a single region of the country using data from just 3 practice groups. Generalizability to other practice groups and regions need to be assessed. Changes in testing practices or coding practices over time and between practices might change the future performance of the RAVIOLI algorithm. The algorithm will require periodic revalidation and possibly modification. Furthermore, the breadth of pathogen capture using the RAVIOLI algorithm depends on the range and frequency of respiratory viral testing by clinicians; greater use of multiplex testing platforms will provide more granular and robust results. RAVIOLI surveillance is limited to patients who seek care, which likely biases the data toward pathogens associated with more severe disease. The PPV of algorithm components may vary by season; whether and how this affects surveillance should be considered. We used a 10% PPV threshold to select diagnosis codes for inclusion. This threshold was arbitrary, but we found using higher thresholds dramatically reduced the number of eligible diagnosis codes. We also found that the terms associated with diagnosis codes with a PPV of ≥10% were specific in their descriptions and not indicative of broad health conditions. However, the PPV threshold for including diagnosis codes should be considered in future revalidation of the algorithm.

The healthcare site data included in developing the algorithm and whose data are part of the weekly reports came from both ambulatory and inpatient care facilities. We observed variation in which RAVIOLI categories (e.g., influenza, RSV) of the algorithm were detected at each site (data not shown). The limited number of sites makes it difficult to know if apparent differences between ambulatory and inpatient sites resulted from differences in catchment populations, illness severity

associated with different viruses, or testing platforms. As the network expands to include a greater number and variety of sites, we plan to examine this question further.

The Massachusetts Department of Public Health has used data from the underlying EHR-based system for infectious disease reporting and surveillance for more than a decade (*18*–*21*,*25*–*28*). This system has been sustained and enhanced over time to meet MDPH needs. As public health agencies consider what they need for the monitoring of current, emerging, and as-yet unidentified pathogens, we have found that a robust EHR data platform is a critical complement to traditional surveillance data.

In conclusion, we developed an integrated, routine, automated EHR-based system for respiratory virus surveillance in Massachusetts. As experience with this approach expands, the hope is that this system will provide early indications of emerging infection trends and prevailing pathogens that render a fuller picture of respiratory viral activity beyond ILI and COVID-like illnesses. A broader view of circulating pathogens will provide public health agencies and healthcare institutions with more precise information useful for informing testing guidance, optimizing health communications; developing more targeted prevention activities, including vaccination; initiating enhanced infection control measures, such as masking and posting of notices in facilities; and generating other policies optimized to minimize the effect on population health of specific circulating pathogens.

Acknowledgments

The authors thank the following colleagues for their contributions: Kelly O'Keefe, Myfanwy Callahan, Dan Flanagan, Jonathan Bress, Brian Herrick, and Michelle Weiss.

This work was supported with funding from the Massachusetts Department of Public Health.

About the Author

Dr. Cocoros is a principal research scientist at the Harvard Pilgrim Health Care Institute and a principal associate in the Department of Population Medicine. Her primary research interests include infectious disease epidemiology, pharmacoepidemiology, and the use of electronic health data for research and surveillance, from pandemic response to pragmatic trials.

References

1. Hansen CL, Chaves SS, Demont C, Viboud C. Mortality associated with influenza and respiratory syncytial virus in the US, 1999–2018. JAMA Netw Open. 2022;5:e220527. https://doi.org/10.1001/jamanetworkopen.2022.0527
2. Ackerson B, Tseng HF, Sy LS, Solano Z, Slezak J, Luo Y, et al. Severe morbidity and mortality associated with respiratory syncytial virus versus influenza infection in hospitalized older adults. Clin Infect Dis. 2019;69:197–203. https://doi.org/10.1093/cid/ciy991
3. Troeger CE, Blacker BF, Khalil IA, Zimsen SRM, Albertson SB, Abate D, et al.; GBD 2017 Influenza Collaborators. Mortality, morbidity, and hospitalisations due to influenza lower respiratory tract infections, 2017: an analysis for the Global Burden of Disease Study 2017. Lancet Respir Med. 2019;7:69–89. https://doi.org/10.1016/S2213-2600(18)30496-X
4. Centers for Disease Control and Prevention. Overview of the National Syndromic Surveillance Program [cited 2022 Aug 9]. https://www.cdc.gov/nssp/documents/NSSP-overview.pdf
5. Boehmer TK, DeVies J, Caruso E, van Santen KL, Tang S, Black CL, et al. Changing age distribution of the COVID-19 pandemic—United States, May–August 2020. MMWR Morb Mortal Wkly Rep. 2020;69:1404–9. https://doi.org/10.15585/mmwr.mm6939e1
6. Elliot AJ, Harcourt SE, Hughes HE, Loveridge P, Morbey RA, Smith S, et al. The COVID-19 pandemic: a new challenge for syndromic surveillance. Epidemiol Infect. 2020;148:e122. https://doi.org/10.1017/S0950268820001314
7. Güemes A, Ray S, Aboumerhi K, Desjardins MR, Kvit A, Corrigan AE, et al. A syndromic surveillance tool to detect anomalous clusters of COVID-19 symptoms in the United States. Sci Rep. 2021;11:4660. https://doi.org/10.1038/s41598-021-84145-5
8. Willis SJ, Eberhardt K, Randall L, DeMaria A, Brown CM, Madoff LC, et al. The evolving nature of syndromic surveillance during the COVID-19 pandemic in Massachusetts. Open Forum Infect Dis. 2021;8(Suppl 1):S695. https://doi.org/10.1093/ofid/ofab466.1401
9. Cocoros NM, Willis SJ, Eberhardt K, Morrison M, Randall LM, DeMaria A, et al. Syndromic surveillance for COVID-19, Massachusetts, February 2020–November 2022: the impact of fever and severity on algorithm performance. Public Health Rep. 2023;138:756–62. https://doi.org/10.1177/00333549231186574
10. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Hellferscee O, et al.; PHIRST group. Asymptomatic transmission and high community burden of seasonal influenza in an urban and a rural community in South Africa, 2017–18 (PHIRST): a population cohort study. Lancet Glob Health. 2021;9:e863–74. https://doi.org/10.1016/S2214-109X(21)00141-8
11. Elson W, Zambon M, de Lusignan S. Integrated respiratory surveillance after the COVID-19 pandemic. Lancet. 2022;400:1924–5. https://doi.org/10.1016/S0140-6736(22)02325-X
12. Ziegler T, Moen A, Zhang W, Cox NJ. Global Influenza Surveillance and Response System: 70 years of responding to the expected and preparing for the unexpected. Lancet. 2022; 400:981–2. https://doi.org/10.1016/S0140-6736(22)01741-X
13. 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;57:373–6.
14. Klompas M, McVetta J, Lazarus R, Eggleston E, Haney G, Kruskal BA, et al. Integrating clinical practice and public

health surveillance using electronic medical record systems. Am J Public Health. 2012;102(Suppl 3):S325–32. https://doi.org/10.2105/AJPH.2012.300811
15. Klompas M, Cocoros NM, Menchaca JT, Erani D, Hafer E, Herrick B, et al. State and local chronic disease surveillance using electronic health record systems. Am J Public Health. 2017;107:1406–12. https://doi.org/10.2105/AJPH.2017.303874
16. Lazarus R, Klompas M, Campion FX, McNabb SJ, Hou X, Daniel J, et al. Electronic Support for Public Health: validated case finding and reporting for notifiable diseases using electronic medical data. J Am Med Inform Assoc. 2009;16:18–24. https://doi.org/10.1197/jamia.M2848
17. Vogel J, Brown JS, Land T, Platt R, Klompas M. MDPHnet: secure, distributed sharing of electronic health record data for public health surveillance, evaluation, and planning. Am J Public Health. 2014;104:2265–70. https://doi.org/10.2105/AJPH.2014.302103
18. Willis SJ, Cocoros NM, Callahan M, Herrick B, Brown CM, Kruskal BA, et al. Assessment of antibiotic prescriptions for Lyme disease after modification of reporting language for positive screening test results. JAMA Netw Open. 2022;5:e2144928. https://doi.org/10.1001/jamanetworkopen.2021.44928
19. Willis SJ, Elder H, Cocoros NM, Callahan M, Hsu KK, Klompas M. Impact of an electronic medical record best practice alert on expedited partner therapy for chlamydia infection and reinfection. Open Forum Infect Dis. 2022;9:ofab574.
20. Willis SJ, Elder H, Cocoros N, Young J, Marcus JL, Eberhardt K, et al. More screening or more disease? Gonorrhea testing and positivity patterns among men in 3 large clinical practices in Massachusetts, 2010–2017. Clin Infect Dis. 2020;71:e399–405. https://doi.org/10.1093/cid/ciaa066
21. Dee EC, Hsu KK, Kruskal BA, Menchaca JT, Zambarano B, Cocoros N, et al. Temporal patterns in chlamydia repeat testing in Massachusetts. Am J Prev Med. 2019;56:458–63. https://doi.org/10.1016/j.amepre.2018.10.006
22. Sutton D, Fuchs K, D'Alton M, Goffman D. Universal screening for SARS-CoV-2 in women admitted for delivery. N Engl J Med. 2020;382:2163–4. https://doi.org/10.1056/NEJMc2009316
23. Ma Q, Liu J, Liu Q, Kang L, Liu R, Jing W, et al. Global percentage of asymptomatic SARS-CoV-2 infections among the tested population and individuals with confirmed COVID-19 diagnosis: a systematic review and meta-analysis. JAMA Netw Open. 2021;4:e2137257. https://doi.org/10.1001/jamanetworkopen.2021.37257
24. Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, et al.; Cochrane COVID-19 Diagnostic Test Accuracy Group. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease. Cochrane Database Syst Rev. 2020;7:CD013665.
25. Elder HR, Gruber S, Willis SJ, Cocoros N, Callahan M, Flagg EW, et al. Can machine learning help identify patients at risk for recurrent sexually transmitted infections? Sex Transm Dis. 2021;48:56–62. https://doi.org/10.1097/OLQ.0000000000001264
26. Gruber S, Krakower D, Menchaca JT, Hsu K, Hawrusik R, Maro JC, et al. Using electronic health records to identify candidates for human immunodeficiency virus pre-exposure prophylaxis: an application of super learning to risk prediction when the outcome is rare. Stat Med. 2020;39:3059–73. https://doi.org/10.1002/sim.8591
27. Yih WK, Cocoros NM, Crockett M, Klompas M, Kruskal BA, Kulldorff M, et al. Automated influenza-like illness reporting—an efficient adjunct to traditional sentinel surveillance. Public Health Rep. 2014;129:55–63. https://doi.org/10.1177/003335491412900109
28. 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:e2626. https://doi.org/10.1371/journal.pone.0002626

Address for correspondence: Noelle M. Cocoros, Harvard Pilgrim Health Care Institute, Department of Population Medicine, 401 Park Dr, Ste 401 East, Boston, MA 02215, USA; email: hphei.harvard.edu

Article DOI: https://doi.org/10.3201/eid3006.230473

Electronic Health Record–Based Algorithm for Respiratory Virus–Like Illness

Appendix

Appendix Table 1. ICD-10-CM diagnosis codes included in algorithm for influenza-like illness (ILI)

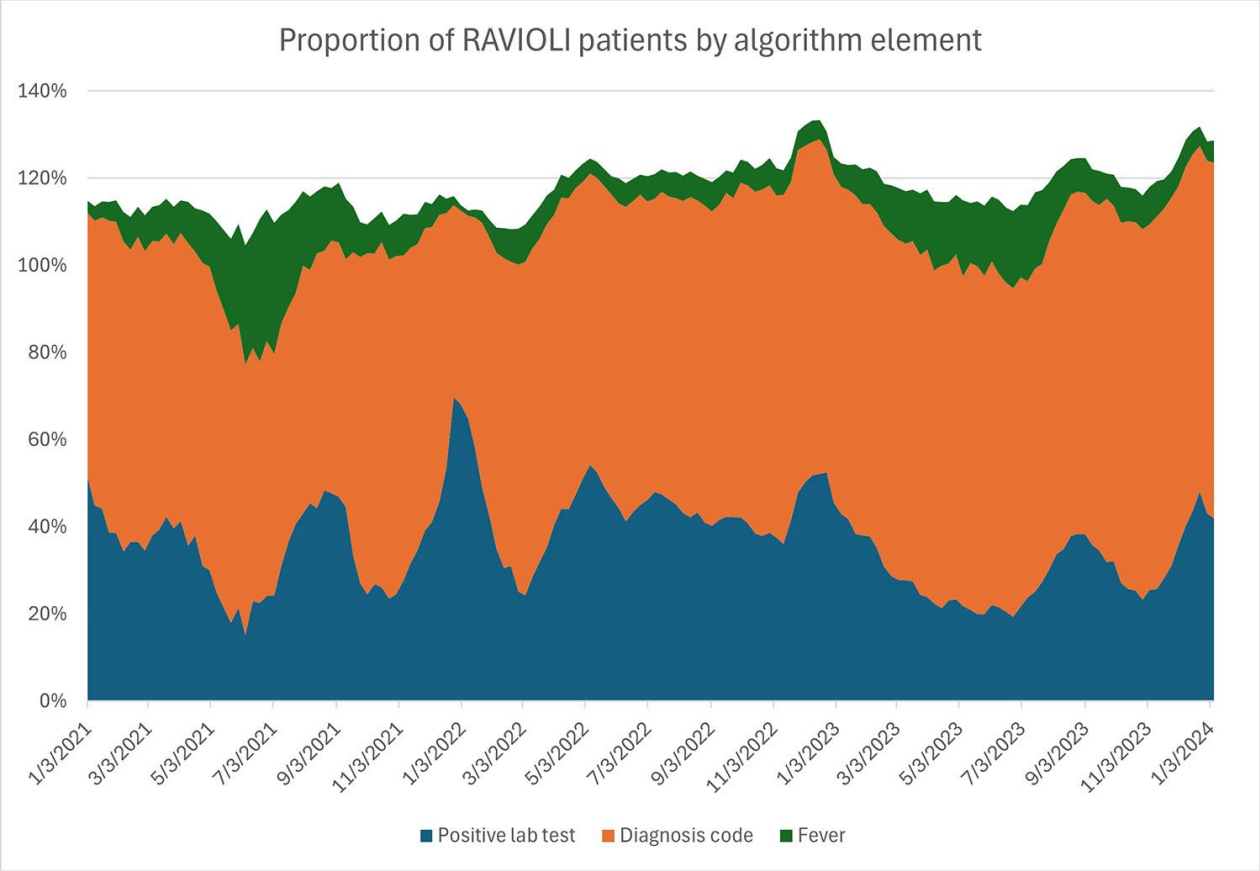
ICD-10 code	Description
Influenza-like illnesses	
B33.8	Other specified viral diseases
B34.1	Enterovirus infection, unspecified
B34.2	Coronavirus infection, unspecified
B34.4	Papovavirus infection, unspecified
B34.8	Other viral infections of unspecified site
B97.19	Other enterovirus as the cause of diseases classified elsewhere
B97.29	Other coronavirus as the cause of diseases classified elsewhere
B97.89	Other viral agents as the cause of diseases classified elsewhere
J00	Acute nasopharyngitis
J02.9	Acute pharyngitis, unspecified
J04.0	Acute laryngitis
J04.10	Acute tracheitis without obstruction
J04.11	Acute tracheitis with obstruction
J04.2	Acute laryngotracheitis
J05.0	Acute obstructive laryngitis
J06.0	Acute laryngopharyngitis
J06.9	Acute upper respiratory infection, unspecified
J20.9	Acute bronchitis, unspecified
J21.8	Acute bronchiolitis associated with other specified organisms
J21.9	Acute bronchiolitis, unspecified
J39.8	Other specified diseases of upper respiratory tract
J39.9	Disease of upper respiratory tract, unspecified
J12.89	Other viral pneumonia
J12.9	Viral pneumonia, unspecified
J13	Pneumonia associated with Streptococcus pneumoniae
J18.1	Lobar pneumonia, unspecified organism
J15.20	Pneumonia associated with staphylococcus, unspecified
J15.211	Pneumonia associated with methicillin susceptible staph
J15.212	Pneumonia associated with Methicillin resistant Staphylococcus aureus
J15.29	Pneumonia associated with other staphylococcus
J17	Pneumonia in diseases classified elsewhere
J18.0	Bronchopneumonia, unspecified organism
J18.8	Other pneumonia, unspecified organism
J18.9	Pneumonia, unspecified organism
J10.00	Influenza associated with other ident influenza virus with unspecified type of pneumonia
J10.08	Influenza associated with other ident influenza virus with other pneumonia
J11.00	Influenza associated with unidentified influenza virus with unspecified type of pneumonia
J11.08	Influenza associated with unidentified influenza virus with specified pneumonia
J12.9	Viral pneumonia, unspecified
J10.01	Influenza associated with other ident influenza virus with same other identified influenza virus pneumonia
J10.1	Influenza associated with other identified influenza virus with other respiratory manifestations
J11.1	Influenza associated with unidentified influenza virus with other respiratory manifestations
J10.2	Influenza associated with other identified influenza virus with GI manifestations
J10.81	Influenza associated with other identified influenza virus with encephalopathy
J10.82	Influenza associated with other identified influenza virus with myocarditis
J10.83	Influenza associated with other identified influenza virus with otitis media
J10.89	Influenza associated with other identified influenza virus with other manifestations
J11.2	Influenza associated with unidentified influenza virus with GI manifestations
J11.81	Influenza associated with unidentified influenza virus with encephalopathy

ICD-10 code	Description
J11.82	Influenza associated with unidentified influenza virus with myocarditis
J11.83	Influenza associated with unidentified influenza virus with otitis media
J11.89	Influenza associated with unidentified influenza virus with other manifestations
R07.0	Cough
R05	Pain in throat
Fever	
R50.2	Drug induced fever
R50.8	Other specified fever
R50.84	Febrile nonhemolytic transfusion reaction
R50.9	Fever, unspecified
R56.00	Simple febrile convulsions

Appendix Table 2. Patient counts and percentages meeting inclusion criteria for select RAVIOLI algorithm categories, January 2023–January 2024

	COVID-19			Influenza			Respiratory syncytial virus			Nonspecific	
Week	N	Positive lab test, %	Diagnosis code, %	N	Positive lab test, %	Diagnosis code, %	N	Positive lab test, %	Diagnosis code, %	N	Fever, %
Jan 1	4,128	50	74	786	71	65	62	77	32	1,011	22
Jan 8	3,550	49	75	532	73	69	42	79	38	1,031	27
Jan 15	2,680	51	75	289	66	75	28	79	39	847	26
Jan 22	2,647	47	78	278	73	78	35	83	37	973	29
Jan 29	2,425	47	76	257	70	81	25	76	32	953	31
Feb 5	2,446	47	78	225	73	81	25	76	44	1,029	31
Feb 12	2,235	45	80	194	71	77	22	73	50	1,025	32
Feb 19	1,841	38	83	129	78	73	14	86	29	850	32
Feb 26	1,863	37	83	145	74	83	23	70	61	1,019	33
Mar 5	1,649	37	83	136	76	82	20	80	35	1,040	33
Mar 12	1,546	35	84	134	84	75	11	91	18	954	34
Mar 19	1,594	35	84	165	80	76	7	86	43	1,016	33
Mar 26	1,531	30	88	149	85	74	7	86	29	1,074	37
Apr 2	1,499	30	89	153	80	77	11	55	73	1,034	36
Apr 9	1,350	28	88	110	83	72	9	56	67	1,038	39
Apr 16	1,122	23	90	106	75	77	12	75	33	802	38
Apr 23	1,124	29	87	104	81	74	7	57	43	854	35
Apr 30	1,189	29	89	94	81	78	2	0	100	833	36
May 7	1,107	26	90	91	85	73	5	60	40	901	42
May 14	1,121	25	90	88	82	75	4	75	50	847	34
May 21	1,168	25	91	100	80	79	1	0	100	873	37
May 28	1,044	27	91	60	80	73	3	100	33	820	39
Jun 4	1,176	28	90	72	83	82	5	60	40	803	39
Jun 11	1,052	29	90	68	72	81	3	67	33	748	44
Jun 18	938	24	91	59	78	78	1	100	100	694	44
Jun 25	814	24	90	65	86	72	6	67	67	691	41
Jul 2	757	30	88	46	74	78	3	33	67	534	43
Jul 9	908	30	87	47	89	77	2	100	0	564	48
Jul 16	919	35	87	43	79	79	2	100	0	591	47
Jul 23	1,035	38	85	45	82	78	10	100	30	638	47
Jul 30	1,163	38	84	58	86	76	9	89	44	560	44
Aug 6	1,396	44	83	38	74	84	5	80	40	579	43
Aug 13	1,474	45	83	49	88	78	5	80	60	591	36
Aug 20	1,730	47	82	62	84	84	17	100	18	585	34
Aug 27	1,921	47	82	61	85	84	12	92	25	575	34
Sep 3	1,837	47	82	51	84	75	17	82	47	609	33
Sep 10	2,117	44	82	81	78	80	25	96	28	717	30
Sep 17	2,303	42	82	97	78	84	30	90	43	837	31
Sep 24	2,372	39	86	87	80	83	39	92	28	859	23
Oct 1	2,326	39	85	100	77	78	55	95	31	955	26
Oct 8	2,123	32	87	101	76	83	72	85	38	1,019	27
Oct 15	2,038	30	88	113	74	81	101	89	35	1,055	24
Oct 22	2,020	27	90	125	78	74	162	92	36	1,080	24
Oct 29	1,863	24	91	131	73	82	196	85	35	1,130	23
Nov 5	1,945	23	91	174	88	72	278	87	38	1,226	25
Nov 12	2,096	24	92	179	80	72	290	87	37	1,268	24
Nov 19	1,927	25	90	191	85	69	258	83	45	958	24
Nov 26	2,717	29	87	371	82	73	318	85	40	1,294	21

Week	COVID-19			Influenza			Respiratory syncytial virus			Nonspecific	
	N	Positive lab test, %	Diagnosis code, %	N	Positive lab test, %	Diagnosis code, %	N	Positive lab test, %	Diagnosis code, %	N	Fever, %
Dec 3	2,875	32	88	573	87	64	313	88	32	1,249	26
Dec 10	3,245	37	86	806	85	68	269	86	35	1,270	26
Dec 17	3,685	40	85	1,043	85	67	203	83	32	1,189	26
Dec 24	3,348	45	82	938	87	66	165	89	24	939	25
Dec 31	3,551	39	83	880	81	70	131	83	32	909	26
Jan 7 (2024)	3,202	40	83	847	80	73	104	87	28	956	27



Appendix Figure. Proportion of RAVIOLI patients by algorithm element, January 2021–January 2024. Data include all categories of RAVIOLI and show the proportion identified on the basis of a positive laboratory test, diagnosis code, and/or fever. Patients could meet ≥1 criterion (e.g., have both a positive laboratory test and a diagnosis code).