Title
Real-time Epidemic Forecasting: Challenges and Opportunities.

Journal
Health security, 17(4)

ISSN
2326-5094

Authors
Desai, Angel N
Kraemer, Moritz UG
Bhatia, Sangeeta
et al.

Publication Date
2019-07-01

DOI
10.1089/hs.2019.0022

Peer reviewed

Health Security
Volume 17, Number 4, 2019 © Mary Ann Liebert, Inc.
DOI: 10.1089/hs.2019.0022

SPATIAL/TEMPORAL ANALYSIS IN INFECTIOUS DISEASE OUTBREAKS

Real-time Epidemic Forecasting: Challenges and Opportunities

Angel N. Desai, Moritz U. G. Kraemer, Sangeeta Bhatia, Anne Cori, Pierre Nouvellet, Mark Herringer, Emily L. Cohn, Malwina Carrion, John S. Brownstein, Lawrence C. Madoff, and Britta Lassmann

Infectious disease outbreaks play an important role in global morbidity and mortality. Real-time epidemic forecasting provides an opportunity to predict geographic disease spread as well as case counts to better inform public health interventions when outbreaks occur. Challenges and recent advances in predictive modeling are discussed here. We identified data needs in the areas of epidemic surveillance, mobility, host and environmental susceptibility, pathogen transmissibility, population density, and healthcare capacity. Constraints in standardized case definitions and timely data sharing can limit the precision of predictive models. Resource-limited settings present particular challenges for accurate epidemic forecasting due to the lack of granular data available. Incorporating novel data streams into modeling efforts is an important consideration for the future as technology penetration continues to improve on a global level. Recent advances in machine-learning, increased collaboration between modelers, the use of stochastic semi-mechanistic models, real-time digital disease surveillance data, and open data sharing provide opportunities for refining forecasts for future epidemics. Epidemic forecasting using predictive modeling is an important tool for outbreak preparedness and response efforts. Despite the presence of some data gaps at present, opportunities and advancements in innovative data streams provide additional support for modeling future epidemics.

Keywords: Infectious diseases, Epidemic management/response, Surveillance, Disease modeling

Angel N. Desai, MD, is a Clinical Infectious Disease Research Fellow, Division of Infectious Diseases, Brigham & Women's Hospital, Boston, MA. Moritz U. G. Kraemer, DPhil, is in the Department of Zoology, University of Oxford, UK; Computational Epidemiology group, Boston Children's Hospital, Boston, MA; and Harvard Medical School, Harvard University, Boston, MA. Sangeeta Bhatia, PhD, is a Research Associate, and Anne Cori, PhD, is a Lecturer; both in the MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London. Pierre Nouvellet, PhD, is a Reader, School of Life Sciences, University of Sussex, Brighton, UK. Mark Herringer, Business Analyst, is the Founder of the Global Healthsites Mapping Project, London. Emily L. Cohn, MPH, is Program Manager, Computational Epidemiology Lab, Boston Children's Hospital, Boston, MA. Malwina Carrion, MPH, is a Lecturer, Department of Health Science, Sargent College, Boston University, Boston, MA. John S. Brownstein, PhD, is Chief Innovation Officer, Innovation and Digital Health Accelerator, Boston Children's Hospital, Boston, MA, and Professor, Department of Pediatrics, Harvard Medical School, Boston, MA. Lawrence C. Madoff, MD, is Professor of Medicine, University of Massachusetts Medical School, Worcester, MA, and Editor, ProMED, International Society for Infectious Diseases, Brookline, MA. Britta Lassmann, MD, is Program Director, International Society for Infectious Diseases, Boston, MA.

THE RECENT OUTBREAKS OF Ebola virus in the Democratic Republic of the Congo, yellow fever in Brazil, and Nipah virus in India demonstrate the continued threat of emerging and reemerging infectious diseases.[1-3] Rapid globalization has led to increases in human mobility and the trade of animals, plants, and food throughout the world. Infectious disease outbreaks that begin in the most remote parts of the globe can now spread swiftly to urban centers and across regions, putting large segments of human populations at risk for infection.[4] In tandem, changes in climate and land use contribute to an increased risk of spillover between animals and humans and to the spread of disease vectors to regions where they were not previously endemic.[5] All of these factors have contributed to large epidemics in the past 2 decades. In order to support outbreak preparedness, the ability to forecast the potential spread of a disease is paramount for implementing public health interventions and effective resource allocation. This is particularly critical for low- and middle-income countries, as they often disproportionately bear the burden of infectious diseases and are hindered by limitations in resources available to tackle them.

Epidemic forecasting that models global risks posed by outbreak events present an opportunity to address the growing need for rapid, open, and accurate data sources. While traditional surveillance mechanisms remain the cornerstone of outbreak investigations, they often require time, resources, and infrastructure rarely available in the setting of an epidemic. Novel data streams, such as epidemic case incidence data provided by digital disease detection tools, demographic data estimates aided by geospatial mapping tools, and advances in mathematical modeling, can support efforts to contain emerging outbreaks. Predictive models can leverage these novel data streams to offer timely case count projections and potential geographic spread of an emerging epidemic in real-time.[6]

Despite the potential that epidemic forecasting has in outbreak scenarios, some key challenges remain. For example, unreliable data on basic epidemiologic parameters and disease dynamics in the setting of an emerging outbreak can limit predictive models. While rapid assessments are paramount to disease prevention and control, no standardized or validated forecasting tools exist, and they must therefore be developed in the course of each new outbreak. In addition, effectively communicating model results to key stakeholders can be disrupted due to limitations in available data needed to produce a reliable forecast. This review aims to provide an overview of data needs, challenges, and recent advances we have identified in our own experiences with real-time surveillance and epidemic forecasting.

Challenges of Global, Real-time Epidemic Forecasting

Accurate data streams are critical to enhancing current forecasting capabilities. The ability to account for population movements, potential changes in pathogen transmissibility over time, and drug and vaccine availability require data sources that are updated in real-time. In addition, global, comprehensive datasets regarding outbreak preparedness that are committed to open data sharing are necessary to make risk assessments and predictions.

Ensuring Updated Data

A key challenge during an outbreak is preserving up-to-date, reliable data. This ensures that estimates of the risk of disease spread reflect the current epidemiologic, demographic, and environmental situation. The frequency at which data need to be updated is dependent on context. The natural history of a disease is an important factor in determining the ideal frequency of case count updates; diseases with a fast progression (eg, influenza) will require daily updates, while in those with a slower progression (eg, HIV), monthly updates may be sufficient. More frequent case counts may also be required at times when transmissibility is expected to change (eg, when interventions are initiated, enhanced, or stopped).

The frequency at which demographic and mobility data should be updated will also depend on the time horizon during which forecasts are made and whether demographic and mobility patterns are affected by the disease. For environmental susceptibility, the ideal frequency will depend on the timescale at which environmental factors affect disease transmission change. Fortunately, such data are typically readily available at a fine spatiotemporal scale. Host susceptibility will be affected by previous and ongoing epidemics as well as public health endeavors such as vaccination. If interventions are ongoing, more frequent updates are necessary. Similarly, healthcare capacity may change over the course of an outbreak, and the ideal temporal and spatial resolution for such data also depend on the desired resolution of the forecast.

Model Uncertainties

All parameters used in models forecasting the spread of infectious diseases are subject to uncertainties. For example, emerging outbreaks may lack vital information regarding pathogen transmissibility. Additional structural issues can stem from the choice of model used—for example, those that describe connectivity versus environmental susceptibility. Limitations involving model outputs need to be communicated clearly and transparently to users and stakeholders to avoid undue concern as well as to present the range of plausible scenarios. Bayesian methodologies have become a tool of choice as they incorporate uncertainties as well as expert knowledge through the choice of prior probabilities. Parametric uncertainties can be explored by numerically sampling all plausible combinations of available values and reporting on the full distribution of predicted risks.[7]

To account for structural uncertainties, a range of model structures should be considered. The natural history of the pathogen and data availability should inform the model as well. Methodological advances in model comparison and averaging are promising avenues to improve the characterization of forecasting uncertainty. In particular, the performance of ensemble predictions where various model structures and their individual uncertainty is aggregated is increasingly recognized.[8] During the validation phase, historical data are used to train the forecasting model by comparing observed and predicted trends of disease spread. Among all model structures and plausible sets of parameters, only those consistent with historical observations are retained.

Data Needs

Several data gaps exist that limit the precision with which epidemic forecasting can occur in real-time. Table 1 demonstrates some of these data needs. In particular, epidemic case incidence, mobility, host and environmental susceptibility, healthcare capacity, and geospatial data needed for real-time case projections remain incomplete and warrant further discussion here.

Epidemic Surveillance

Although digital disease surveillance tools offer the rapid dissemination of epidemic data that can be incorporated into forecasting models, several gaps remain. In particular, the absence of standardized case definitions in the setting of an outbreak may affect the precision of model estimates. In addition, achieving the timely sharing of incidence data during an outbreak to incorporate into epidemic forecasts is paramount, but is often hindered for a variety of reasons. Innovative digital disease surveillance tools, such as ProMED, the Program for Monitoring Emerging Diseases, and HealthMap, generally provide more sensitive disease signals over traditional reporting mechanisms and allow for the rapid sharing of epidemic case incidence data in real-time.[24] ProMED is an internet-based infectious disease surveillance tool dedicated to the rapid dissemination of information on global human, animal, and plant infectious disease and toxin outbreaks. Fifty human subject matter experts located in 35 countries verify and contextualize reports.[9] HealthMap aggregates more than 200,000 data sources and uses natural language processing and algorithms to tag, filter, analyze, validate, and map real-time surveillance of emerging threats.[10] Despite the increasing global reach of both tools, specific challenges were noted in regards to outbreak detection and digital incidence data curated from these systems during the 2013-2016 West Africa Ebola outbreak. Delays in accurate and timely case detection, a lack of granular data during the epidemic as a result of the relative paucity of local and online news media in the region, unreliable internet connectivity, few supporting resources, and generally poor infrastructure were some of the challenges encountered. In addition, digital

Table 1. Summary of Data Needs for Real-time Global Epidemic Forecasting

Aim	*Data Needs*	*Examples of Open-Access Data Sources*
Case counts	Case counts including confirmed, probable, and suspected cases Open sharing of case data	ProMED[9] HealthMap[10] World Health Organization[11]
Mobility	Movement of individuals and populations Flight and travel networks	Flowminder[12] Flirt[13]
Host susceptibility	Immunization coverage data: pediatric and adult	GHSA[14] World Health Organization[11] UNICEF[15]
Environmental susceptibility	Climate data such as temperature and precipitation Environmental characteristics, eg, flooding Vector mapping Ecological niche mapping	NOAA[16] NASA Earthdata[17] Natural Earth[18]
Healthcare capacity	GPS latitude and longitude coordinates of hospitals, clinics, and health posts Number of beds Number of physicians Number of nurses Number of critical care beds	Healthsites.io[19] World Bank[20]
Population density and spatial demographic data	Census data Shapefiles for all countries Corresponding estimates of population sizes	LandScan[21] WorldPop[22] Facebook Population Maps[23]

disease incidence data could be resolved only up to the country level and were not available at subnational spatial resolution (eg, district level).

Other digital disease tools used for epidemic surveillance include the Medical Information System (MedISys), a fully automatic event-based system that monitors reporting on infectious diseases threats, and the Global Public Health Intelligence Network (GPHIN), developed by Health Canada in collaboration with the World Health Organization.[25,26] GPHIN is a restricted-access, internet-based multilingual tool that continuously searches global media sources to identify information about disease outbreaks.

Despite the challenges surrounding epidemic surveillance as discussed above, it is expected that digital disease surveillance data availability will continue to improve through higher internet penetration in low-income settings.[27] Further refinements in disease dictionaries, scraping targets, and algorithm tuning may be warranted to ensure accurate incidence case reporting on subnational, national, and regional levels in order to project real-time case numbers and disease spread within countries and across regions.

Pathogen Transmissibility

As an epidemic evolves, current levels of transmissibility may be inferred directly from case incidence data by estimating the reproduction number Rt, the average number of secondary cases generated by a typical infected individual. This information can then be used to forecast future incidence.[28,29] Accurate case incidence data are crucial for informing transmissability. Although under-reporting can be accounted for, difficulties quickly arise if the relative level of reporting is temporally variable and not appropriately quantified. When transmissibility cannot be directly inferred from the data, previously published estimates must be used. If these are not available, or an epidemic is caused by a pathogen not currently known to cause human disease, critical information regarding disease transmissibility may be missing.

Population Density and Spatial Demographic Data

Integrating population density and spatial demographic data into forecasting models is critical in order to predict the occurrence of new cases and to inform spatial spread. Availability of demographic data, however, is constrained in low-income settings where traditional census data are often of poor quality, reported irregularly, and at low granularity, with limited breakdown of population classifications into subcategories. Without accurate baseline demographic data, developing tools for outbreak preparedness that are specific to a particular region is challenging.

Novel global geospatial datasets and recent methodological advances provide new opportunities and are available through different sources. For example, Landscan, an online repository of population density data, is available at approximately 1 km resolution and is updated annually.[21] WorldPop, a digital spatial database focused on low- and middle-income countries, provides open-access population estimates by integrating census, survey, satellite, and GIS in a flexible machine-learning framework to produce high-resolution maps of population counts and densities.[22] WorldPop also produces estimates of population demographics such as age, births, and pregnancies. Other examples include high-resolution population maps that are produced jointly by the Facebook Connectivity Lab and the Center for International Earth Science Information Network (CIESIN). These programs provide data on the distribution of human populations at a 30-meter spatial resolution.

Mobility Data

Human mobility is essential to understanding the geographic spread of infectious diseases. Unfortunately, data regarding mobility dynamics on spatial and temporal scales are sparse, limiting predictive model accuracy. Estimating risks of disease importation within and between countries is challenging due to difficulties in independently ascertaining the relevance and reliability of available data. Similarly, open access information about population flows is often scarce or unavailable at the height of any given epidemic.

Examples of currently available real-time mobility data sources include transportation information such as flight and mobile phone data, although high-resolution information from mobile phones is often not available during outbreaks.[13,30] During the 2013-2016 West Africa Ebola outbreak, analyses of Orange Telecom mobile phone data produced initial maps of human movement.[31] Facebook currently generates and shares anonymous, aggregated population movement data as part of their Disaster Maps project, but its accuracy and utility have not yet been formally assessed.[23] Finally, reliable census data are critical. In the absence of detailed information on human mobility, relative risk levels of disease spread can only be estimated, thereby limiting model accuracy.

Host Susceptibility Data and Environmental Drivers

Host and environmental susceptibility may help elucidate the drivers of infectious disease transmission and can enhance the predictive accuracy of epidemic models.[32,33] For example, human host factors such as prior immunity acquired either naturally or through vaccination is available for many pathogens and collated through the Global Health Security Agenda (GHSA).[14] Baseline human population vaccination rates are often updated regularly by countries, but immunity through vaccination campaigns during outbreaks might be more difficult to capture and integrate in disease transmission models.[34]

Developing models for livestock epidemics comes with its own challenges and, in the absence of livestock population and immunity data in many under-resourced countries, relies on built-in assumptions. Agricultural practices and breeding techniques can further influence disease spread in livestock.[35] There are numerous other host factors that can influence pathogen susceptibility in both humans and animals, such as sex, age, genetics, nutritional status, and co-morbidities, providing additional opportunities for model refinements.

Remote sensing and satellite data allow for real-time monitoring of environmental drivers of epidemics such as temperature, precipitation, humidity, flooding, and other characteristics that influence the spread of certain pathogens. From these data it is possible to derive risk maps of transmission for diseases that have a strong environmental component such as vector-borne diseases like Zika, chikungunya, yellow fever, and dengue.[36-38] Other examples of environmental effects on infectious disease transmission and spread include political instability, conflict, overcrowding, poor sanitation, and urbanization. While these drivers have been shown to influence disease spread, using them in the context of real-time epidemic forecasting remains distant.[39]

Healthcare Capacity Data

Healthcare capacity is an important factor that can influence disease control and therefore disease dynamics and spread. Including this kind of data may help refine case projections and geographic disease spread. Health facility types and geographic coordinates provide parameters that offer some specificity to predictive models. Additional indicators of capacity could include the number of physicians and nurses at each facility, the number of hospital beds, and the presence of specialized equipment in the form of ventilators and isolation rooms. Taken together, these parameters would help quantify the capacity of a healthcare system to prevent, predict, and respond to an outbreak. The quality of the health system in a country will also influence case reporting rates, and models can adjust the number of reported cases accordingly. Unfortunately, such health facility attributes are often not available in a timely manner and at a regional level. To address this, the Global Healthsites Mapping Project is building a "global commons" of health facility data using OpenStreetMap, a collaborative online project that allows users to create free, editable maps.[19] This collaborative approach promotes the sharing of health facility data with the goal of establishing an accessible baseline of global data using an open data approach. The open data approach is focused on 3 key components:

1. Enable national health agencies and organizations to share and contribute data to OpenStreetMap;
2. Enable collaboration between national health agencies and volunteer communities; and
3. Connect multiple data streams to build higher quality data.

Additional Data Sources

Other promising data sources to improve the accuracy of forecasting include biological drivers of disease transmission such as strain-specific transmission dynamics, vaccine effectiveness, sequencing data, and electronic health record data. Realistically, the prospect of using such data in the context of epidemic real-time forecasting remains distant.

Considerations

Novel data streams can provide valuable information to public health officials that complement traditional data sources and reporting mechanisms. While traditional disease surveillance remains the backbone of outbreak investigation and routine data collection, informal disease surveillance tools allow for more rapid dissemination and detection of case incidence data. In addition, novel data streams can provide important input for outbreak modeling in regions that do not have functioning public health systems because of ongoing conflict or poor infrastructure. It also has advantages in data aggregation at a global level, as data are collected using a unified methodology. Rapid epidemic incidence data reporting across national boundaries can also assist modeling of epidemic spread and epidemic potential in geographic areas that do not have surveillance programs.

Recent Advances and Opportunities

Advances

While existing data needs present several challenges for epidemic forecasting, recent advances in high-level methodologies as well as novel open data sources provide opportunities to refine predictive disease models for future outbreak scenarios. As digital disease surveillance tools have continued to evolve over time, real-time case information extracted from these systems has provided significant gains for forecasting efforts.[24]

Demographic and geographic data are becoming more readily available through programs such as LandScan™ and WorldPop. Free software for statistical computing and graphics such as "R" are used to create the next generation of analytics tools for informing the response to disease outbreaks and health emergencies. For example, the R Epidemics Consortium (RECON) brings together experts in data science, modeling methodology, public health, and software development with the goal of developing tools for handling, visualizing, and analyzing outbreak data.[40]

Other examples of infectious disease forecasting tools include FluSight, a consortium from Reich Lab at the University of Massachusetts Amherst, and MRIIDS, the Mapping the Risk of International Disease Spread project. FluSight displays live influenza predictions for the United States based on weekly data provided by the US Centers for Disease Control and Prevention. It visualizes interactive prediction models while also displaying historical data for comparison.[41] The MRIIDS prototype was developed for the 2013-2016 West Africa Ebola outbreak. The tool communicates risks posed by outbreak events in real-time by combining digital disease case incidence data with multiple open access data streams into a single probabilistic framework.[42] It was validated against WHO case incidence data and performed well in a retrospective analysis, but its use and impact as a real-time forecasting tool remains to be determined in future outbreaks.

Model Methodologies

Many methodologies exist to forecast case numbers during an outbreak. The development of new methods and the testing of established models need to take place between outbreak events in preparation for future epidemics. Opportunities such as the RAPIDD Ebola Forecasting Challenge coordinate the modeling community ahead of the next epidemic.[8] These projects provide a deeper understanding of model accuracy and data requirements in a controlled environment and can be extended to other known and unknown pathogens. They are also an opportunity to improve coordination and collaboration between modeling groups. The RAPIDD Ebola Forecasting Challenge demonstrated that for short-term, 1- to 4-week incidence predictions, model performance did not improve with increased complexity, underscoring that even simple models provide valuable information for future case counts. Stochastic semi-mechanistic models may also be an important tool in combating future outbreaks, as they combine the power of mechanistic models with the flexibility of including uncertainty about precise outbreak dynamics.[7,43] More recently, machine learning and climate data were used to develop a model for dengue case count forecasting.[44]

To facilitate comparisons across multiple models, common data standards need to be defined. We expect to see continued forecast improvement through methodological innovation, technological advances, and collaboration. Development of modular models that are built on a core of simple models and are able to incorporate novel data streams as they become available could also improve performance.

Open Data Sharing

Recent advances in digitization, connectivity, and big data provide valuable sources of information. By combining these data streams with conventional methods for monitoring infectious diseases, tools can be developed to assist the public health community to respond to outbreaks swiftly. Data should be shared openly, transparently, and free of political constraints. Definitions of incentives to share data under an open data license need to be considered. To foster scale-up and ensure sustainability over time, forecasting tools should be developed in a modular, flexible manner and use open-source statistical software to make relevant code publicly available. Improved interoperability among data sources, health ministries, and electronic health management systems will also be vital. Open data sharing will allow multiple teams to develop, test, validate, and compare models.

Integration

Many public health officials already use innovative digital disease surveillance tools in their daily work to complement traditional disease surveillance efforts.[45] Refining forecasting targets to more closely align with public health priorities will improve integration of model outputs with decision making. Clear communication of model uncertainties will increase confidence in model outputs. Close collaborations, such as those between the Global Outbreak Alert and Response Network (GOARN), a partnership of institutions, and networks including traditional and innovative digital disease surveillance systems that pool human and technical resources for rapid identification, confirmation, and response to outbreaks of international importance are paramount.[46]

Conclusions

Outbreaks of emerging and reemerging infectious diseases will continue to present challenges in the future. Providing actionable insights such as accurate forecasting of case counts and the potential geographic reach of infectious disease spread is critical for resource allocation and preparedness planning. While several gaps in current data streams provide certain constraints on epidemic forecasting at present, recent advances in the field provide opportunities for continued refinement of future predictive models. Clearly communicating the limitations of outbreak prediction will ensure the adoption of predictive tools by public health officials, operations managers, and healthcare practitioners. In addition, innovation using novel data streams and methodologies need to take place between global outbreak events in order to enhance epidemic preparedness. A modular approach that is rapidly scalable for new pathogens and allows for the easy integration of additional data streams to refine output will be essential for future iterations of epidemic forecasting models.

Acknowledgments

This article was made possible in part by USAID's Combating Zika and Future Threats: A Grand Challenge for Development program and Grant Number T32 AI007433 from the National Institute of Allergy and Infectious Diseases. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or USAID. None of these funding sources assisted in the conception or writing of the manuscript.

References

1. Nuzzo JB, Inglesby TV. Ramping up the response to Ebola. *N Engl J Med* 2018;379(26):2490-2491.
2. Friedrich MJ. Mapping yellow fever virus in Brazil. *JAMA* 2018;320(19):1969.
3. Arunkumar G, Chandni R, Mourya DT, et al. Outbreak investigation of Nipah virus disease in Kerala, India, 2018. *J Infect Dis* 2019;219(12):1867-1878.
4. Jones BA, Betson M, Pfeiffer DU. Eco-social processes influencing infectious disease emergence and spread. *Parasitology* 2017;144(1):26-36.
5. Medlock JM, Leach SA. Effect of climate change on vector-borne disease risk in the UK. *Lancet Infect Dis* 2015;15(6): 721-730.
6. Eisenstein M. Infection forecasts powered by big data. *Nature* 2018;555(7695):S2-S4.
7. Asher J. Forecasting Ebola with a regression transmission model. *Epidemics* 2018;22:50-55.
8. Viboud C, Sun K, Gaffey R, et al. The RAPIDD Ebola forecasting challenge: synthesis and lessons learnt. *Epidemics* 2018;22:13-21.
9. International Society for Infectious Diseases. About ProMED-mail. https://www.promedmail.org/aboutus/. Accessed July 9, 2019.
10. HealthMap. About HealthMap. http://www.diseasedaily.org/about. Accessed July 9, 2019.
11. World Health Organization. https://www.who.int/. Accessed July 9, 2019.
12. Flowminder. http://flowminder.org. Accessed July 9, 2019.
13. Flirt. https://flirt.eha.io. Accessed July 9, 2019.
14. Global Health Security Agenda. https://www.ghsagenda.org/. Accessed July 9, 2019.
15. UNICEF. https://www.unicef.org/. Accessed July 9, 2019.
16. National Oceanic and Atmospheric Administration. https://www.noaa.gov/. Accessed July 9, 2019.
17. NASA Earthdata. https://earthdata.nasa.gov/. Accessed July 9, 2019.
18. Natural Earth. http://www.naturalearthdata.com. Accessed July 9, 2019.
19. Saameli R, Kalubi D, Herringer M, Sutton T, de Roodenbeke E. Healthsites.io: the Global Healthsites Mapping Project. In: Hostettler S, Najih Besson S, Bolay JC, eds. *Technologies for Development: From Innovation to Social Impact.* Cham, Switzerland: UNESCO, Springer; 2018.
20. Flowminder. http://flowminder.org. Accessed July 9, 2019.
21. Oak Ridge National Laboratory. LandScan™: Geographic Information Science & Technology. https://landscan.ornl.gov/. Accessed July 9, 2019.
22. WorldPop. Mapping population distributions. http://www.worldpop.org.uk/. Accessed July 9, 2019.
23. Facebook Data for Good. Disaster maps. https://dataforgood.fb.com/tools/disaster-maps/. Accessed July 9, 2019.
24. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the web for public health surveillance. *N Engl J Med* 2009;360(21):2153-2155, 2157.
25. Medisys. Recent disease incidents. http://medisys.newsbrief.eu/medisys/helsinkiedition/en/home.html. Accessed July 9, 2019.
26. Government of Canada. About GPHIN. https://gphin.canada.ca/cepr/aboutgphin-rmispenbref.jsp?language=en_CA. Accessed July 9, 2019.
27. Murphy J, Roser M. Internet. Our world in data. https://ourworldindata.org/internet. Accessed July 9, 2019.
28. Nouvellet P, Cori A, Garske T, et al. A simple approach to measure transmissibility and forecast incidence. *Epidemics* 2018;22:29-35.
29. Cori A, Donnelly CA, Dorigatti I, et al. Key data for outbreak evaluation: building on the Ebola experience. *Philos Trans R Soc Lond B Biol Sci* 2017;372(1721).
30. Kraemer MUG, Golding N, Bisanzio D, et al. Utilizing general human movement models to predict the spread of emerging infectious diseases in resource poor settings. *Sci Rep* 2019;9(1):5151.
31. Wesolowski A, Buckee CO, Bengtsson L, Wetter E, Lu X, Tatem AJ. Commentary: containing the Ebola outbreak—the potential and challenge of mobile network data. *PLoS Curr* 2014;6.
32. Kraemer MUG et al. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. *eLife* 2015;4: e08347.
33. Kraemer MUG, Hay SI, Pigott DM, Smith DL, Wint GRW, Golding N. Progress and challenges in infectious disease cartography. *Trends Parasitol* 2016;32(1):19-29.
34. Shearer FM, Moyes CL, Pigott DM, et al. Global yellow fever vaccination coverage from 1970 to 2016: an adjusted retrospective analysis. *Lancet Infect Dis* 2017;17(11):1209-1217.
35. Brooks-Pollock E, de Jong MC, Keeling MJ, Klinkenberg D, Wood JL. Eight challenges in modelling infectious livestock diseases. *Epidemics* 2015;10:1-5.
36. Nsoesie EO, Kraemer MU, Golding N, et al. Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Euro Surveill* 2016;21(20).
37. Messina JP, Kraemer MU, Brady OJ, et al. Mapping global environmental suitability for Zika virus. *Elife* 2016 Apr 19;5.
38. Shearer FM, Longbottom J, Browne AJ, et al. Existing and potential risk infection zones of yellow fever worldwide: a modelling analysis. *Lancet Glob Health* 2018;6(3):e270-e278.
39. Kilpatrick AM, Randolph SE. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet* 2012; 380(9857):1946-1955.
40. R Epidemics Consortium. RECON. https://www.repidemicsconsortium.org/. Accessed July 9, 2019.

41. Tushar A, Reich NG, Ray EL. flusight: Influenza Forecasts Visualizer (v2.0). 2017.
42. Mapping the Risk of International Disease Spread. www.mriids.org. Accessed 1.29.19.
43. Funk S, Camacho A, Kucharski AJ, Eggo RM, Edmunds WJ. Real-time forecasting of infectious disease dynamics with a stochastic semi-mechanistic model. *Epidemics* 2018; 22:56-61.
44. Guo P, Liu T, Zhang Q, et al. Developing a dengue forecast model using machine learning: a case study in China. *PLoS Negl Trop Dis* 2017;11(10):e0005973.
45. Bijkerk P, Monnier AA, Fanoy EB, Kardamanidis K, Friesema IH, Knol MJ. ECDC Round Table Report and ProMed-mail most useful international information sources for the Netherlands Early Warning Committee. *Euro Surveill* 2017;22(14):30502.
46. World Health Organization. Global Outbreak Alert and Response Network. https://www.who.int/ihr/alert_and_response/outbreak-network/en/. Accessed July 9, 2019.

Manuscript received January 30, 2019;
revision returned April 18, 2019;
accepted for publication April 23, 2019.

Address correspondence to:
Angel Desai, MD
Infectious Disease Clinical Research Fellow
Massachusetts General Hospital
Boston, MA

Email: angel.n.desai@gmail.com