Commentary

Emerging Challenges and Opportunities in Infectious Disease Epidemiology

Joseph A. Lewnard* and Arthur L. Reingold

* Correspondence to Dr. Joseph A. Lewnard, Division of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, 2121 Berkeley Way, Berkeley, CA 94720 (e-mail: jlewnard@berkeley.edu).

Initially submitted September 7, 2018; accepted for publication November 29, 2018.

Much of the intellectual tradition of modern epidemiology stems from efforts to understand and combat chronic diseases persisting through the 20th century epidemiologic transition of countries such as the United States and United Kingdom. After decades of relative obscurity, infectious disease epidemiology has undergone an intellectual rebirth in recent years amid increasing recognition of the threat posed by both new and familiar pathogens. Here, we review the emerging coalescence of infectious disease epidemiology around a core set of study designs and statistical methods bearing little resemblance to the chronic disease epidemiology toolkit. We offer our outlook on challenges and opportunities facing the field, including the integration of novel molecular and digital information sources into disease surveillance, the assimilation of such data into models of pathogen spread, and the increasing contribution of models to public health practice. We next consider emerging paradigms in causal inference for infectious diseases, ranging from approaches to evaluating vaccines and antimicrobial therapies to the task of ascribing clinical syndromes to etiologic microorganisms, an age-old problem transformed by our increasing ability to characterize human-associated microbiota. These areas represent an increasingly important component of epidemiology training programs for future generations of researchers and practitioners.

infectious diseases; methods; modeling; surveillance

Abbreviation: EBOV, Ebola virus.

The priority afforded to infectious diseases within epidemiologic research has been fluid over the past 200 years or longer. Despite the lasting prominence of early investigations into measles, cholera, plague, typhoid fever, malaria, and yellow fever (1–6), the intellectual tradition of modern epidemiology stems largely from studies of chronic diseases dating to the post-World War II era, when such conditions came to surpass infectious diseases in morbidity and mortality in high-income countries amid improvements in living conditions and the introduction of numerous antibiotics and vaccines. This epidemiologic transition co-occurred with a shift in focus for epidemiologic research (and training programs) toward the multifactorial etiology of chronic conditions (7). In parallel, early 20th-century work on the “dependent happenings” of communicable diseases (8–11) yielded to the development of today’s core biostatistical methods for chronic diseases, premised on the independence of outcomes among subjects (12–14).

Since the late 20th century, the emergence of human immunodeficiency virus and acquired immunodeficiency syndrome and other infections has renewed interest in infectious diseases and their health, economic, and security implications (15). Outbreaks of severe acute respiratory syndrome, pandemic influenza A H1N1, Ebola virus (EBOV), and Zika virus have prompted international responses, whereas influenza A H7N9 and H5N1, Lassa fever virus, Nipah virus, and Middle East respiratory syndrome coronavirus, among other agents, have been a source of regional concern. The incidence and endemic range of “neglected” infections, including dengue and cholera (16, 17), have expanded, and antimicrobial resistance has threatened to derail the control of tuberculosis, typhoid, malaria, gonorrhea, yaws, and invasive bacterial infections (18–23). Although highly effective vaccines are available, measles and yellow fever have resurfaced on multiple continents, due to gaps in vaccine coverage (24, 25), while short-lived vaccine-induced protection has facilitated unexpected resurgences in diseases once on the path to elimination, such as pertussis and mumps (26, 27).

After decades of relative obscurity in the mid-20th century, infectious disease epidemiology has experienced an intellectual rebirth in response to disease emergence. Repopulation of this field by scientists trained not only in clinical medicine but...
ecology, demography, and quantitative sciences has led to the adoption of methods scarcely addressed in traditional public health training programs. Cluster-randomized trial designs, for example, have become commonplace for evaluating infectious disease interventions, quantifying indirect effects resulting from contagion (28). Classical models of ecological dynamics have been adapted to address the transmission and control of infectious agents, while our expanding ability to integrate such models with epidemiologic data through Bayesian statistics has enhanced their relevance to policymaking (29–31). Most recently, sequencing and phylogenetic analysis have afforded an unprecedented view into the population structure and dynamics of pathogens (32, 33).

Here, we review developments in infectious disease epidemiology together with their implications for research and practice, and for the training of future epidemiologists. We first consider the role of epidemiologic surveillance in the context of pathogen emergence and the integration of surveillance data into quantitative studies of transmission. Next, we discuss challenges in causal inference, including the evaluation of public health interventions against emerging pathogens and the difficulties of attributing clinical syndromes to microbial agents.

SURVEILLANCE OF EMERGING INFECTIONS

Public health practice

Surveillance justifiably has been seen as a core public health function, with its important role articulated by Langmuir in 1963 (34). In the United States and elsewhere, surveillance for diverse infectious (and noninfectious) diseases has typically relied on an essentially passive system of reporting by healthcare providers or laboratories, often mandated by public health laws. Although such passive systems of disease reporting have proven invaluable, their limitations, such as incomplete, often inconsistent detection of cases and delayed detection of outbreaks, are well documented (35). As a result, active surveillance systems that do not depend on providers and laboratories to report have been developed and promoted.

The Centers for Disease Control and Prevention–funded Emerging Infections Program and its components (e.g., Foodnet, ABCs), initiated in 1994, is a notable example of such a system, as are international versions promoted through the Centers for Disease Control and Prevention’s Global Health Security initiative (36). Such systems, while expensive to develop and maintain, are of great value, especially when they collect biological specimens (e.g., isolates of bacterial and viral pathogens) for typing and support analytic epidemiologic studies (e.g., case-control studies of vaccine effectiveness (37)). Nevertheless, these systems, too, may have limitations, particularly in their ability to detect in a timely fashion outbreaks caused by novel microbial agents, prompting interest in alternative methods. In this article, we consider proposed alternatives, highlighting their benefits and challenges.

Surveillance for disease emergence

Beyond efforts to quantify the incidence of infectious diseases of known etiology, disease surveillance has been advocated as a means of mitigating the threat posed by novel pathogens. Large-scale efforts to identify pathogens with the potential to spill over from animals to humans have received notable investment, for instance from the US Agency for International Development Emerging Pandemic Threats program (38). More recently, metagenomic sequencing has enabled the number of known viruses to be multiplied in such studies (39). The aim of those working on the Global Virome Project, launched in 2018, to characterize within 10 years all 1.6 million viruses thought to exist (40).

However, the pathway for translating data sets produced by such activities into actionable threat-reduction programs remains unclear. That only approximately 250 viruses are known to infect humans (and an even smaller handful to cause major epidemics) may constrain the value of large-scale virus discovery for identifying high-risk pathogens, as well as viral determinants of pathogenic potential (41). Spillover events causing recent epidemics—including H1N1 in Mexico, Middle East respiratory syndrome in Saudi Arabia, and EBOV in West Africa—have been poorly predicted by factors long believed to drive disease emergence (42). Accordingly, interest in expanding our catalogue of potential pathogens should be weighed against our persisting need to enhance the detection and control of outbreaks of known pathogens (43). For instance, 2 EBOV epidemics in the Democratic Republic of the Congo in 2018 took weeks to be identified, with dozens of suspected cases having already accumulated (44, 45).

Serological studies have received increasing enthusiasm for monitoring emerging pathogens of significance to humans (41, 46). The prevalence of antibodies indicating previous exposure may provide valuable information about the frequency of animal-human spillover events and the potential for person-to-person spread, overcoming reporting biases that favor detection of large outbreaks under traditional surveillance. Moreover, the low cost of multiplex assays makes integrated surveillance of multiple pathogens plausible. Although serosurveys have bolstered recent efforts to understand the geographic range and clinical spectrum of EBOV and Zika virus infections (47, 48), the enhancement of dengue hemorrhagic fever risk by prior exposure (49), and the role of immunologic history in influenza susceptibility and vaccine response (50), there remain few examples of public health programs undertaking serological studies for routine surveillance, at least in civilian populations (51).

Emerging data and analytics

Outside of laboratory-based surveillance, the increasing availability of passively collected “Big Data” on the health and behaviors of individuals has prompted enthusiasm about enhancing disease surveillance through alternative data streams. Initiatives such as the ProMED-mail network and HealthMap (52, 53) compile and disseminate news about outbreaks from media and other sources, aiming to trigger investigation by public health organizations. Data such as emergency department visits, medication sales, online search queries, and social media postings have also been suggested as real-time indicators of outbreak activity, although their integration into public health responses remains a subject of debate (54–56). The need to overcome reporting biases is a central challenge, because observations may be too nonspecific to distinguish between meaningful and spurious signals in settings with high technological capacity while also

being insensitive to even high-risk events in resource-poor settings (57, 58). Although nontraditional data sources have, in some applications, supported inferences about epidemic dynamics (59), limited information about cases from such sources remains a barrier. For instance, models fitted from news reports of recent measles and mumps outbreaks have yielded considerable underestimates of vaccine coverage (60–62), underscoring the importance of field investigations.

Forecasting the incidence of diseases has been a more successful application of these emerging data streams and data-analytic approaches. Although the acknowledged failure of Google Flu Trends—a prediction approach based on Internet search behavior—yielded important lessons about nonmechanistic forecasts, approaches based on machine learning and crowdsourced human judgment have provided the most accurate within-season predictions of US influenza activity in recent comparisons (63–65). Given expanding interest in forecasting among researchers, funding agencies, and other stakeholders, there is a clear and compelling need to evaluate whether such forecasts can enhance the success and efficiency of public health response efforts.

UNDERSTANDING TRANSMISSION DYNAMICS

Model-data integration for emerging diseases

Mathematical modeling as a means to understanding infectious disease spread dates to studies by Sir Ronald Ross (8). Although the use of models to connect data such as age of infection to transmission dynamics of endemic infections has longstanding precedent (66, 67), assimilation of outbreak data for near-term assessments of control priorities is a comparatively recent phenomenon. Integration of modeling with the public health response to epidemics of bovine spongiform encephalopathy and foot-and-mouth disease in the United Kingdom and the severe acute respiratory syndrome epidemic (68–73) has led to expectations for near real-time modeling studies during major outbreaks. In recent experience, models of the spread of pandemic influenza A H1N1 (74), cholera (75), Middle East respiratory syndrome (76), EBOV (77, 78), Chikungunya virus (79), Zika virus (80, 81), yellow fever (82), and plague (83) have all been published within weeks of the respective outbreak notifications.

Although the circumstances of particular epidemics dictate what data may be available and pertinent, methods for fitting models to data have generally focused on exponential growth rates in cases (84) or the distribution of the serial interval (85). Methods based on the latter class of data offer the advantage of illustrating real-time changes in reproductive numbers (86); however, the requisite information from patient line lists is seldom available. Reliance instead on ecological data exposes models to numerous vulnerabilities, notably the inability to discern individual risk factors (and thus the population meaningfully at risk). These shortcomings may prevent models from predicting reductions in transmission before depletion of the susceptible population.

A challenge thus lies ahead in determining the role of models in outbreak response and the best practices for communicating modeling results. Although the ability of models to evaluate prophylactic strategies may be considered a benefit, recommendations to act against remote future risks have sometimes triggered resistance among stakeholders (87). During the West African EBOV epidemic, for example, attention to worst-case model-based projections prompted some to question the reliability of the models (88), reflecting an important discrepancy between public understanding of modeling as a forecasting tool and the intended uses of models for scenario-based comparisons (89, 90). This use of modeling has been better understood in attempts to communicate the impact of interventions after the fact (91).

Microbial sequencing

The ease of sequencing pathogen genomes has afforded a new view into transmission during outbreaks. Use of sequence data to identify transmission clusters in the presence of unsure epidemiologic links dates to the early years of the human immunodeficiency virus and acquired immunodeficiency syndrome epidemic (92). In recent years, sequencing has aided efforts to track the sources of unexplained epidemics of cholera in Haiti (93) and EBOV in West Africa (94), and has shown increasing utility for reconstructing the geographic spread of pathogens (95, 96). A particular advantage of phylogenetic analysis is the possibility of estimating unobserved epidemiologic quantities, such as the reporting fraction (97) and reproductive numbers for subcritical transmission (98), which remain difficult to assess from traditional case-notification data.

Beyond reconstructing the demographic history of pathogen lineages, recent years have seen progress toward joint analysis of epidemiologic and sequencing data (99). Such “phyldynamic” approaches have shown particular relevance for emerging infections, including distinguishing the role of repeated introductions and subsequent local transmission (100–103). Whereas most applications have been tailored to specific data sets and assumptions, the development of generalized methods for joint inference of epidemiologic and phylogenetic parameters remains a priority (104) to support real-time analysis.

EVALUATION OF INFECTIOUS DISEASE INTERVENTIONS

Efficacy evaluations in emergencies

The ability to rapidly develop and deploy countermeasures to mitigate the threat posed by emerging infections has received increasing recognition as a component of public health preparedness. However, outbreaks are difficult environments in which to evaluate interventions. During the West African EBOV epidemic, the feasibility of a new paradigm for development and evaluation of interventions in emergencies was demonstrated by accelerated vaccine safety, immunogenicity, and efficacy studies (105). The Coalition for Epidemic Preparedness Innovations was established in 2017, with an initial focus on vaccines against Nipah virus, Middle East respiratory syndrome coronavirus, and Lassa fever virus, in addition to adaptable vaccine platforms for novel threats (106).

Lessons learned in EBOV vaccine trials will have an influential bearing on evaluations during future emergencies. Despite efforts to accelerate evaluation of candidate vaccines, incidence had reached low levels by the time phase III efficacy trials were ready to begin, posing a threat to their statistical power: A planned trial
in Liberia was canceled due to declining transmission (107), and no cases of disease occurred in a second trial in Sierra Leone (108), preventing efficacy assessments. In a stepped-wedge trial in Guinea, clusters of primary and secondary contacts of EBOV disease cases were randomly assigned to immediate or delayed vaccination; no cases were reported among vaccine recipients during the trial (109) or in subsequent field deployments of the vaccine, supporting a conclusion of near 100% vaccine efficacy.

Debates surrounding design of these trials highlight methodological questions requiring additional attention. In the Guinean trial, a “ring” vaccination scheme helped maximize power by enrolling contacts of known cases (110). However, the choice of individual- or cluster-level randomization within rings was debated. Because members of a vaccinated cluster are exposed to direct protection through vaccination and indirect protection due to reduced transmission within their clusters (28), cluster-randomized trials have weaker statistical power than individually randomized trials of the same size (111). Moreover, the direct effect measured in individually randomized studies may be a preferred, transportable efficacy measure (112). Uses of simulation helped in planning vaccine trials tailored to the real-world circumstances of the EBOV outbreak (113) and enabled trialists to compare alternative designs in terms of ethical mandates (114). Simulation-guided design further presents the opportunity for applying adaptive trial methods (115) in the context of infectious disease outbreaks, where dynamic trends in incidence may highlight the benefits of such approaches.

Observational designs

In addition to efficacy trials for new interventions, observational studies are needed to assess licensed interventions against evolving and re-emerging pathogens. Most commonly applied in evaluations of influenza vaccines, test-negative designs have become popular in routine (116) and exploratory (117) studies of vaccine effectiveness. By measuring vaccine effectiveness from the exposure odds ratio of vaccination among individuals seeking care who test positive or negative for a pathogen of interest, this design seeks to overcome associations of health-care seeking care who test positive or negative for a pathogen of interest, this design seeks to overcome associations of health-care seeking behavior, whereas at the population level, age-related confounding may occur when vaccine recommendations are unique challenges arise in vaccine safety studies; at the individual level, vaccination and adverse-event detection may be confounded due to health-care-seeking behavior, whereas at the population level, age-related confounding may occur when vaccine recommendations are based on the individual’s age.

Vaccine safety

Because vaccine studies are typically powered for primary clinical endpoints, long-term observational studies are needed to monitor for rare vaccine-attributable adverse events. Such studies have been crucial to identifying safety concerns such as intussusception after rotavirus vaccination (132) and to refuting spurious links, such as autism onset after measles-mumps-rubella vaccination (133). However, unique challenges arise in vaccine safety studies; at the individual level, vaccination and adverse-event detection may be confounded due to health-care-seeking behavior, whereas at the population level, age-related confounding may occur when vaccine recommendations are based on the individual’s age.

Ecological designs taking advantage of natural experiments have proven useful in numerous studies of vaccine safety (134, 135) but inconclusive for certain classes of rare events, including those that also result from the vaccine-targeted infection (136, 137). Recent years have seen growing interest in case-only methods offering the ability to reduce or rule out individual-level sources of confounding. Case-crossover methods are common among such approaches (138, 139) and resemble matched case-control studies by sampling “control” periods from the person-time contributed by case individuals before an adverse event. Self-controlled case-series methods similarly benefit from the use of cases as their own controls, following a cohort...
logic in estimating the relative incidence of adverse events after vaccination within specified risk periods (140); with adequate sample size, researchers may be able to use such analyses to eliminate or greatly reduce potential time- or age-related confounding.

**Antimicrobial drugs**

In response to the growing threat posed by antimicrobial resistance, the World Health Organization and national governments have prioritized bringing novel antimicrobial drugs to market (141). These plans will necessitate phase III trials in which the efficacy of new therapeutic agents is addressed and possibly phase IV studies, in which the optimal use of new and existing drugs, either singly or in combination, can be determined. Whereas patients traditionally have been enrolled in antimicrobial treatment studies on the basis of target bacterial species infections or clinical syndromes, it is uncertain within what strata such trials may yield transportable inferences. Rather than merely the infecting pathogen’s baseline resistance or susceptibility phenotype, strata may be defined by factors such as pathogen lineage, mutational barriers to resistance development (142), and presence of horizontally transferable resistance elements in cocolonizing agents or environmental sources (143). Stratification based on interpatient and even intrapatient tumor heterogeneity is an emerging feature of cancer therapy trials and may provide a template for such designs (144).

In addition to clinical endpoints, carriage of susceptible and resistant bacteria, including commensal agents not purposefully targeted by treatment, can inform the impact of treatment on resistance selection in targeted and bystander species (145). Whereas between-group differences in the absolute prevalence of colonization with resistant organisms (146) are tested for routinely, stratified measurements (see the reports of Shrag et al. (147) and Feikin et al. (148), for example) of the effect of treatment on acquisition and clearance of susceptible and resistant pathogens are more informative of the underlying biology (149) and may detect signals of selection masked by simpler between-group comparisons (150, 151).

Studies are also needed to address optimal deployment of new and existing antimicrobial drugs in clinical practice. The tradeoff between maximizing a drug’s impact and minimizing resistance selection has led policymakers to ration certain new drugs as last-resort treatments. However, in recent experience, such decisions have ignited ethical debates (152). Coupling mathematical modeling with field-based studies has proven useful for understanding the effectiveness of antimicrobial use policies, as highlighted in recent evaluations of the risk of resistance under population-wide access of the tuberculosis drug bedaquiline (153, 154) and antimicrobial cycling to limit resistance selection in hospital settings (155, 156).

**ETIOLOGIC UNDERSTANDING**

Whereas certain efforts we have discussed have been made to identify novel microorganisms able to cause human infection (40, 157), most epidemics caused by emerging pathogens have been recognized first by clusters of anomalous syndromes—such as cardiopulmonary syndrome caused by New World hantaviruses, severe acute respiratory syndrome caused by a coronavirus, and congenital abnormalities caused by Zika virus—before the role or even existence of the etiologic microorganism had been characterized. The problem of ascribing a clinical syndrome to an etiologic agent is among the oldest in epidemiology, dating at least to the 19th century, when Robert Koch laid out criteria for such inference (i.e., Koch’s, or more correctly, Henle-Koch’s, postulates (158)). However, these postulates have long been recognized as inadequate, particularly for illnesses caused by viruses, and thus of largely historical interest (159); for instance, the notion that a pathogen should be absent from healthy individuals is incompatible with the prominence of carriage and asymptomatic infection in the natural history of numerous pathogens. A growing appreciation of the complexity of the human microbiome, and the likelihood that intricate mixtures of microorganisms at diverse body sites may be either the cause or consequence of both detrimental and beneficial physiological states, has further highlighted the difficulty of linking a given health outcome to infection by a single microorganism.

The now-recognized critically important role of persistent infection and the inflammation it can produce in diverse cancers and possibly other chronic diseases has further diminished the relevance of Koch’s postulates and the age-old distinction between infectious and chronic diseases, as illustrated in the well-known example of human papillomavirus causing cervical, anal, and oral cancers (160). The best causal understanding of such relationships has come from randomized trials demonstrating that infection-preventing interventions are efficacious against downstream chronic illness, such as peptic ulcers due to *Helicobacter pylori* and chronic wheeze due to respiratory syncytial virus (161, 162). Natural experiments following the same intuition have provided additional evidence of such relationships, such as measles-induced immunosuppression (163), malnutrition and stunting due to enteric infection (164), and complex or chronic otitis media due to tissue damage from acute early-life disease (165). Such relationships have proven difficult to probe in the absence of a randomized or natural experiment, because of the likelihood that confounding factors influence individuals’ risk for initial infection as well as chronic sequelae. New paradigms for ascribing an etiologic role to microorganisms and resulting host responses are clearly needed and may prove important in efforts to quantify the health impacts of infectious disease interventions.

**SUMMARY**

The recognition in the 1970s and 1980s that infectious diseases were not, in fact, disappearing as important causes of morbidity and mortality in human populations has been followed by renewed interest in these conditions, especially in the emergence or re-emergence of diverse infectious diseases and in the role of infection in various chronic diseases. At the same time, advances in epidemiologic and statistical methods, together with the growing availability of data from diverse sources, have provided new tools and approaches for studying infectious diseases. The infectious disease epidemiologists of the future will need a solid grounding in the biology of infection and the host immune response, as well as training in the increasingly sophisticated approaches to causal inference; the manipulation and analysis of large-scale data sets, including pathogen genome sequences; and
mathematical modeling, together with the behavioral and social determinants of health. Integration of these elements into epidemiology training programs (e.g., through coursework in Bayesian statistics and phylogenetics) represents an increasingly important consideration for academic departments.

ACKNOWLEDGMENTS

Authors’ affiliations: Division of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, Berkeley, California (Joseph A. Lewnard, Arthur L. Reingold).

The authors received no specific funding for this article. Conflict of interest: none declared.

REFERENCES


