Curr Epidemiol Rep (2016) 3:212–222
DOI 10.1007/s40471-016-0078-4

EPIDEMIOLOGIC METHODS (D WESTREICH, SECTION EDITOR)

Trends in the Mechanistic and Dynamic Modeling of Infectious Diseases

Justin Lessler[1] · **Andrew S. Azman**[1] · **M. Kate Grabowski**[1] · **Henrik Salje**[1] · **Isabel Rodriguez-Barraquer**[1]

Published online: 2 July 2016

Abstract The dynamics of infectious disease epidemics are driven by interactions between individuals with differing disease status (e.g., susceptible, infected, immune). Mechanistic models that capture the dynamics of such "dependent happenings" are a fundamental tool of infectious disease epidemiology. Recent methodological advances combined with access to new data sources and computational power have resulted in an explosion in the use of dynamic models in the analysis of emerging and established infectious diseases. Increasing use of models to inform practical public health decision making has challenged the field to develop new methods to exploit available data and appropriately characterize the uncertainty in the results. Here, we discuss recent advances and areas of active research in the mechanistic and dynamic modeling of infectious disease. We highlight how a growing emphasis on data and inference, novel forecasting methods, and increasing access to "big data" are changing the field of infectious disease dynamics. We showcase the application of these methods in phylodynamic research, which combines mechanistic models with rich sources of molecular data to tie genetic data to population-level disease dynamics. As dynamics and mechanistic modeling methods mature and are increasingly tied to principled statistical approaches, the historic separation between the infectious disease dynamics and "traditional" epidemiologic methods is beginning to erode; this presents new opportunities for cross pollination between fields and novel applications.

Keywords Infectious disease · Modeling · Dynamics · Phylodynamics · Dynamic models · Mechanistic models

"…the 'how' precedes the 'why';"—Galileo

Introduction: a Brief History of Dependent Happenings and Infectious Disease Modeling

In 1916, Ronald Ross coined the term "dependent happenings" to capture the fundamental difference between the study of infectious diseases in populations and other health phenomena [1]. Because infectious diseases are, for the most part, acquired from the people around us, our own future health status depends on that of our neighbors (e.g., the more people we know who are infected, the more likely we are to become infected ourselves). For acute infectious diseases, the health status of the population often changes quickly over time, with the number of people infectious, susceptible to being infected, and immune to the disease changing substantially over the course of an epidemic. Further, the membership in each of these groups does not vary arbitrarily over time but is driven by often well-understood biological processes (Box 1). For instance, in the simple example of a permanently immunizing infection spread through person-to-person transmission such as measles, new susceptible individuals only enter the population through birth and immigration; these individuals can

This article is part of the Topical Collection on *Epidemiologic Methods*

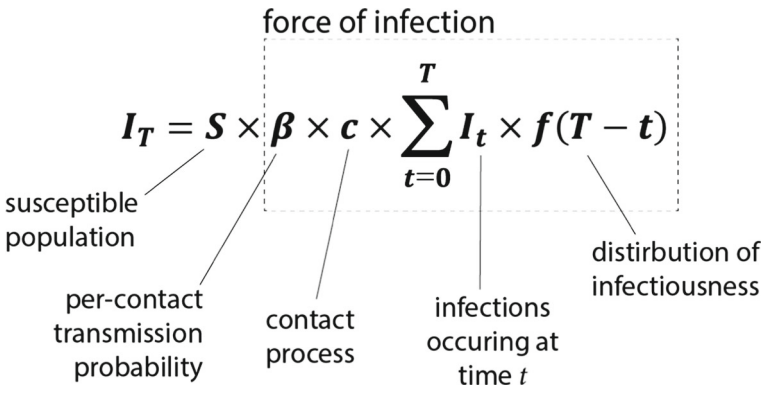
Andrew S. Azman, M. Kate Grabowski, Henrik Salje and Isabel Rodriguez-Barraquer contributed equally to this work.

✉ Justin Lessler
justin@jhu.edu

[1] Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

then only become infected by contact with existing infectious individuals, who, in turn, will eventually become immune or die and be removed forever removed from participation in the epidemic process.

Box 1. Drivers of infectious disease dynamics



The epidemic dynamics of infectious diseases are driven by similar mechanistic relationships between the current and future health states of the population. The expected number of infections at some time T is illustrated for a directly transmitted disease in the above equation. Dynamic and mechanistic models of disease spread, regardless of complexity, capture these relationships in order to improve inference or predict the disease dynamics. The study of infectious disease dynamics encompasses the study of any of the shared drivers of the mechanistic processes of disease spread with an eye towards better understanding disease transmission. As illustrated above, these include:

The size of the susceptible population ( ): The number of people available to be infected. The dynamics of susceptibility is not shown here, but can itself can be complex, as new susceptibles enter the population through birth, immigration and loss of immunity. For many diseases (e.g., dengue, influenza), susceptibility is not a binary state, and complex models may be needed.

The force of infection: The force of infection is the probability that any individual who is susceptible at a given time becomes infected (analogous to the hazard of infection). The size of the susceptible population times the force of infection is the reproductive number ( ). When this value is above 1, the epidemic will grow. When it falls below 1, it will recede.

The infectious process ( ): The infectious process dictates the chances of becoming infected on a direct or indirect contact with an infectious individual. Here represented as a per contact probability of infection, this itself can be a complex, multi-faceted process.

The contact process ( ): The process by which infectious contacts are made, whether directly or mediated by a some vector or the environment, is one of the most complex parts of the infectious process. Much modern research focuses on accounting the role of space and population structure in the contact process.

Previous infections ( ): Fundamental to the nature of infectious diseases is the number of previous infections, however, these may not as directly lead to current infections as illustrated here if transmission is mediated by a vector or the environment.

The natural history of disease ( ): How infectious people are at particular times after their infection determines their contribution to ongoing disease transmission, and fundamentally drives the speed at which epidemics move through the population. Other aspects of disease natural history (e.g., the incubation period) may determine our ability to control a disease and its ultimate health impact.

Over the course of the twentieth century, the main body of epidemiologic research became increasingly reliant on models of statistical association, often with strong assumptions of independence between observations (hereafter referred to as associative models) [2]. However, as a result of the need to deal with dependent happenings, there remained a strong subpopulation within infectious disease epidemiology that used models of an entirely different type. Variously referred to as "mathematical," "dynamic," or "mechanistic" models, these models are characterized by having a mechanistic representation of the dynamic epidemic process that determines how the population's state at time $t+1$ depends on its state at time t (hereafter referred to as mechanistic models). Historically, these models have more often been deterministic and built top-down from first principles rather than based on patterns in any particular dataset. However, as increasing computational power has caused an explosion in the types of models that can be subject to rigorous statistical analysis, there has been a shift toward more data-driven and statistical approaches and a greater focus on stochasticity and uncertainty. This confluence between principled statistical inference and mechanistic processes is paying huge dividends in the quality of the work being produced and the types of questions being answered across disciplines within infectious disease epidemiology. Infectious disease models are being given a firmer empirical footing, while the use of generative mechanistic approaches allows us to use models as tools for forecasting, strategic planning, and other activities in ways that would not be possible with models that do not represent the underlying dynamic epidemiologic processes.

In this manuscript, we review current research into dynamic and mechanistic models of infectious disease with a focus on how the confluence of mechanistic approaches, new statistical methods, and novel sources of data related to disease spread are opening up new avenues in infectious disease research and public health. For those interested in further pursuing the topic, we provide a list of key resources in Box 2.

Box 2. Key tools and resources

There exist a number of freely available resources that aid infectious disease modelling efforts.

Courses

1. **MOOCs** - There are several Massive Open Online Courses (MOOCs) that focus on the basic concepts around modeling of infectious diseases. One of the most popular is the *Epidemics - the Dynamics of Infectious Diseases* MOOC created by Penn State and available on Coursera (www.coursera.org), that offers an introduction to basic concepts on infectious disease dynamics

Software/technical resources

2. **BEAST2** - Powerful open source platform for phylodynamic inference (beast2.org).
3. **The R-epi project.** A project that includes a number of packages to perform inference in infectious disease outbreaks. Operates on the freely available and open source R platform (sites.google.com/site/therepiproject).
4. **GLEAM** - Flexible platform that combines human mobility data (e.g., from flight path data) with stochastic transmission models to allow epidemic forecasting (www.gleamviz.org).

Data resources

5. **WorldPop** - Population and demographic maps at 100m2 grid cells throughout the globe (available at www.worldpop.org.uk).

A Focus on Data and Inference

Recent work in infectious disease dynamics has been characterized by an increasing focus on data and principled approaches to inference. Traditionally, deterministic models were a dominant tool for studying the theoretical and practical basis of disease transmission in humans and animals. This approach yielded important practical and theoretical results

that form the basis of our understanding of disease dynamics [3•, 4] but was limited in approach. Deterministic models are usually parameterized through some combination of trajectory matching (i.e., minimizing the distance between observed and simulated data) and specifying parameters based on previous literature. This approach may be sufficient to describe the expected behavior of an infectious disease in a large population, but an increasing focus on how stochasticity and parameter uncertainty impact public health decision making, combined with the growing availability of computational power, has driven a move toward more statistically principled and data-driven likelihood-based approaches.

Illustrative of this evolution is the contrast between early descriptions of the key dynamic properties of HIV transmission with more recent dynamic characterizations of pandemic H1N1 influenza (H1N1pdm), Middle Eastern Respiratory Syndrome (MERS-CoV), and Ebola. In the late 1980s, several papers were published laying out the essential properties of HIV transmission dynamics that would govern the course of the epidemic (at least in the near term) [5–7]. These papers presented deterministic epidemic models that captured the processes driving the epidemic and highlighted the key parameters, such as the speed of progression to AIDS, that needed to be investigated. Uncertainty was largely addressed through scenario-based approaches (e.g., different future epidemic trajectories were presented for different plausible sets of parameters), and for the most part, different aspects of the transmission dynamics were derived from independent studies, with only the growth rate (i.e., doubling time) estimated from incidence data. While the parameters essential to characterizing epidemic dynamics remain largely unchanged for recently emerging pathogens, the approach to data and estimation is qualitatively different. Integrated statistical frameworks built on Markov chain Monte Carlo (MCMC) techniques are used to estimate all, or most, parameters from different datasets and to produce posterior distributions both for parameter estimates and forecasts of future incidence [8, 9•, 10]. These methods allow innovative use of unconventional data sources, such as disease incidence among travelers [8, 9•], to estimate population incidence of the disease, and molecular data can supplement incidence data providing independent estimates of the same parameters (see discussion of phylodynamics below) [8, 9•].

These recent attempts to quickly characterize the properties of emerging diseases are emblematic of an increasing focus on developing statistical methods, grounded in dynamical models, to estimate key epidemic parameters based on diverse data sources. Surveillance data is often used to estimate the reproductive number (R_t, the number of secondary infections that a primary infection is expected to infect at any point, t, in an epidemic), incubation period, and serial interval (the expected time between symptom onset in a case and the people that case infects), as was done in recent outbreaks of MERS-CoV [9•, 11] and Ebola [10]. Surveillance data has also been paired with serological data to estimate force of infection (i.e., the hazard of infection) and basic reproductive number (R_t, when the population is fully susceptible, designated R_0) of several pathogens, including dengue and chikungunya [12–14].

Dynamic modeling approaches can also aid in the interpretation of surveillance data. State-space models (e.g., hidden Markov models) have been used to pair our mechanistic understanding of disease transmission with a statistical inference framework by linking observed incidence and dynamics with underlying population disease burden and susceptibility (i.e., the population's state). Notably, this approach has been used to estimate global reductions in mortality due to measles in the face of incomplete reporting [15•]. Likewise, Valle and colleagues used hybrid associative and mechanistic models to account for biases that treatment of detected malaria cases might have on estimates of key values such as the incidence rate [16].

Perhaps, the biggest limitation when attempting to characterize the parameters driving disease transmission remains data availability. Data on disease transmission often comes from incomplete surveillance data or represents one aspect of a partially observed epidemic process. For example, epidemic curves are usually limited to symptomatic cases. Similarly, key events in the transmission process, such as the exact time of infection, are generally not observable and have to be inferred from observed data. Methods, such as the use of MCMC-based data augmentation and known transmission processes to infer the possible distribution of transmission trees, have been developed to deal with partially observed data and have been used to reconstruct outbreaks [17], characterize risk factors for transmission [18], and quantify the impact of interventions [19].

A limitation of likelihood-based approaches, such as those mentioned above, is that it is often impossible or impractical to evaluate the data likelihood, in particular, for complex models and large datasets. To deal with this challenge, several "likelihood-free approaches" have been developed, including approximate Bayesian Computation (ABC) and sequential Monte Carlo (SMC) [20, 21]. An advantage of these approaches is that they only require the ability to simulate from candidate models (i.e., if data can be simulated, calculation of the likelihood is unnecessary) and therefore can be more easily applied than methods that require iterative evaluation of the likelihood. ABC has been used to integrate phylodynamic and epidemic models of influenza and other pathogens [22], and SMC methods have been used to parametrize dengue transmission models using data from multi-centric vaccine clinical trials [23•].

Despite important advances over the past decades in linking data and transmission models as tools of inference, many challenges remain and are the topic of continued

research. Inference for complex models and using large datasets remain challenging, in part, due to computational burden. Mechanistic models offer promise as a way to simultaneously link data from diverse, heterogenous data sources (as in [24]), but this promise has yet to be realized, though some phylodynamics methods come close (see below). Further, rapid inference on emergent epidemics remains a tool only used in high-profile epidemics [9•, 10], and these inferential techniques remain inaccessible to field epidemiologists.

Forecasting

Scientists and physicians have tried to forecast the course of epidemics since the time of Hippocrates. Associations between incidence and extrinsic factors such as time of year, climate, and weather can and have been used to forecast infectious diseases [25–27]. However, mechanistic models that capture the natural history of the disease (e.g., duration of immunity and cross protection) [28], mode of transmission [29], and movement patterns [30, 31] can improve forecasts, particularly when associations with extrinsic drivers of incidence, such as climate, are weak or unknown (e.g., for emerging pathogens).

In recent years, forecasts based on models that capture the underlying mechanistic processes of transmission and pathogenesis have become common. Uses range from forecasting the peak timing and magnitude of an influenza season [32], to forecasting the spread and spatial extent of emerging pathogens such as Zika, Ebola, and chikungunya [33–35]. The mechanistic underpinning of these models allows forecasts to take into account dynamic processes that may, otherwise, be impossible to capture, including changes in behavior and resource availability in response to an epidemic [36].

Approaches adopted from computer science, machine learning, and climate science have enhanced our ability to provide reliable forecasts with quantified uncertainty. Particularly important are ensemble approaches [37], which integrate forecasts from multiple imperfect models or different parameterizations of the same model to calculate a distribution of potential courses of the epidemic [38, 39]. Ensemble approaches have been used for forecasting influenza in temperate regions [40], where influenza is highly seasonal, and more recently in subtropical areas such as Hong Kong, where the seasonal pattern is less distinct [32]. Similarly, ensemble-based climate models have been incorporated with infectious disease models to forecast climate-related disease including plague and malaria [41]. These examples use multiple parameterizations of a single model. Ensemble approaches can also be used to accommodate uncertainty in model structure by comparing estimates from parameterizations across different models, as is work by Smith and colleagues where an ensemble of 14 different individual-based models was used to estimate the impact of a future malaria vaccine [42].

There has been an explosion in the number of forecasts being made to aid public health decision making, including a number of government-sponsored contests to forecast the progression epidemics of diseases ranging from influenza to chikungunya and dengue [43–45]. As forecasts become more widely used, care must be given to ensure that the purposes of the model and uncertainty (both structural and statistical) are well communicated. In a recent outbreaks of emerging infectious diseases, like Ebola, groups raced to make forecasts of the evolution and spatial spread of the outbreak [33, 34], with some predicting an epidemic size orders of magnitude greater than what was actually observed. While some of these extreme forecasts were made as worst-case planning scenarios, they were interpreted as likely scenarios, raising alarm and casting doubt on the validity of model-based forecasts, thus highlighting the importance of clear communication of a model's purpose and its limitations [33, 46].

The quality of infectious disease forecasts and standards for their interpretation are far from the gold standard of methods and conventions used in the meteorology. Improvement in both the methods used and their practical use remain critical areas of future research.

Big Data

The advent of "big data" has opened up new avenues in how we parameterize and understand models of infectious disease spread. Big data refers to massive datasets that are too large or complex to be processed using conventional approaches [47]. However, advances in computing increasingly allow their use without large delays in processing time or unrealistic computing capacity requirements.

One of the most successful attempts to use big data to understand disease dynamics has been the use of call data records (CDRs) to capture human mobility. For each call that is made or received, mobile phone operators capture the mobile phone tower through which the call is made. By tracking tower locations for a subscriber, we can capture where he or she is moving. In practice, to ensure confidentiality, CDRs are usually averaged over millions of subscribers to provide estimates of flux between different locations in a country. Transmission models built upon CDR-based estimates of seasonal patterns of human movement have been used to explain patterns of rubella disease in Kenya [48] and dengue in Pakistan [31]. In both instances, models built on empirical human movements seen in CDRs outperformed alternative parametric models of population movement based on our theoretical understanding of human travel patterns (e.g., gravity models where movement is based on community size and distance [49]) and models where movement was not considered. CDR-based models have also been used to understand the dynamics of large-scale outbreaks such

as Ebola in West Africa [50] and challenges in malaria elimination [51]. Questions remain as to the generalizability of CDR-based analyses in settings where mobile phone ownership is low [52], and problems capturing flows between countries remain. However, the large-scale penetration of mobile phones, even in resource-poor settings, makes CDRs a hugely valuable data source for informing infectious disease models.

Another type of big data that has enormous potential for furthering our understanding of disease dynamics is satellite imagery. Detailed satellite imagery can provide high-spatial-resolution estimates of key determinants of many infectious disease processes, including environmental factors (e.g., land cover), climatic conditions (e.g., precipitation, temperature), and population density throughout the globe [53, 54]. In infectious disease epidemiology, such datasets have recently been used as the basis for statistical models that produce fine scale maps of disease incidence, prevalence, and derived transmission parameters (e.g., force of infection, basic reproductive number) for a large number of diseases. Early efforts focused on mapping the global distribution of key drivers of malaria transmission [55, 56]. These approaches have since been used to estimate the burden from a wide range of pathogens [57–60], vectors [61], and host reservoirs [62]. These analyses have allowed disease burden and risk to be estimated in areas with limited surveillance capabilities, expanding our understanding of the global burden of many pathogens. High-resolution geographic data can gain additional power when paired with mechanistic models that capture changes in disease risk, as in recent analyses that accounted for the effect of birth, natural infection, and vaccine disruptions driving increases in measles susceptibility and epidemic risk in the wake of the Ebola outbreak [63].

Finally, big data are increasingly being used with mechanistic models to more directly estimate disease burden in real time [64]. For example, patterns in the usage of different Google search terms have been shown to correlate well with incidence trends for diseases such as influenza [65, 66] and dengue [67]. It is worth noting that big data alone can typically only explain part of trends in incidence, and models that incorporate seasonal dynamics typically outperform models that rely solely on search terms. Similar approaches have been used with Wikipedia updates [68] and social media sites such as Twitter and Facebook. Electronic medication sales data and electronic medical records have also been proposed as novel data sources of disease trends [69]. These approaches can provide estimates much faster than traditional surveillance systems, where it often takes weeks or months for results of cases to be aggregated and analyzed. Mechanistic models can then be fit to this data to better understand seasonal or spatial parameters. For example, Yang et al. used mechanistic models fit to Google Flu Trend data to estimate epidemiological parameters such as the basic reproductive number and the attack rate for 115 cities in the USA over a 10-year period [69, 70].

Phylodynamic Inference

Phylodynamics, the study of how epidemiological, immunological, and evolutionary processes interact to shape pathogen genealogies, is among the newest and fastest-growing areas in infectious disease research [71]. The term phylodynamics was coined in 2004 by Grenfell et al., who observed that the structure of pathogen phylogenies reveals important features of epidemic dynamics in populations and within hosts [72]. This relationship provides a theoretical framework for linking molecular data with population-level disease patterns using dynamic models.

Early methodological work in phylodynamics concentrated on the formal integration of the Kingman's coalescent and birth death models from population genetics with standard deterministic epidemic models. The coalescent model provides a framework for estimating the probability of coalescent events (lineages converging at a common ancestor) as we move back in time given changes in population size [73]. The branching patterns in a phylogenetic tree describe the ancestral history of sequenced pathogens, such that nodes closer to the root of the tree represent historical coalescent events while nodes near the tip represent recent events. The strong relationship between the genetic divergence of pathogens and time allows us to estimate the timing of coalescent events and estimate the rate of growth (or decline) of pathogen populations. These estimates are the critical link between genetic and epidemic models [74].

The formal statistical integration of population genetic and epidemic models allows us to estimate the critical epidemiological parameters such as the basic reproductive number directly from pathogen sequence data [75–77]. For example, Magiorkinis et al. used sequence data from viruses collected over a 12-year period in Greece to estimate subtype-specific reproductive numbers and generation times for hepatitis C [77]. Using data from the ATHENA HIV cohort, which samples ~60 % of HIV-infected persons in the Netherlands, Bezemer et al. used viral sequence data to estimate reproductive numbers for hundreds of circulating transmission chains, showing that large chains persisted within the Netherlands for years near the threshold for sustaining an epidemic ($R = 1$) [77, 78].

Other phylodynamic applications have focused on elucidating the spatial dispersal pattern diseases such as influenza and HIV. In an analysis of nearly 10,000 influenza genomes, Bedford et al. showed fundamental differences in the global circulation patterns of H3N2, H1N1,

and influenza B viruses and that these were potentially driven by differences in the force of infection and rates of immune escape (i.e., antigenic drift) [79]. Likewise, Faria et al. used HIV sequence data from central Africa to reconstruct the early epidemic dynamics of HIV-1 using phylodynamic methods and showed that Kinshasa in the Democratic Republic of Congo likely served as the focal point for global HIV spread [79, 80].

Phylodynamics plays an important role in real-time infectious disease surveillance and targeted control [81]. In recent epidemics of MERS-CoV and Ebola, genomic data was used to assess transmission patterns, monitor viral evolution in populations, and inform epidemic control [9•, 82–84]. Analyses of HIV epidemics among US and European men who have sex with men demonstrate that the amalgamation of epidemiologic and genomic data can be used to identify high-risk transmitters and optimal targeted intervention packages [85•, 86]. However, the utility of real-time phylodynamic analysis in many settings remains hindered by inadequate infrastructure, few viral sequence data, and limited analytic capacity at local levels.

Initial phylodynamic models could only deal with simple epidemic patterns (e.g., exponential growth), and recent methodological work has focused on extending the phylodynamic framework to account for complex nonlinear population dynamics [87, 88]. For instance, Rasmussen and colleagues showed how phylodynamic models could be extended to integrate more complex stochastic and structured epidemic models using Bayesian MCMC and particle filtering [89•]. Others have focused on resolving transmission network structure from phylogenies [90, 91] or integrating data across multiple scales by incorporating information on intra-host pathogen diversity and ecological processes directly into phylodynamic models [92]. However, equally important recent work has shown that phylodynamic inferences can be highly sensitive to sampling and unmeasured factors. Simulation studies show that the relationship between phylogenetic trees and the underlying transmission networks is a complex function of the sampling fraction and underlying epidemic dynamics [93, 94] and that failure to account for intra-host viral diversity may bias phylodynamic inference [95].

Other Uses of Dynamic and Mechanistic Models

Here, we have focused on areas where we feel that there has been the most innovation in the use of dynamic epidemic models in recent years. This is not to imply that innovation has stopped in other areas where dynamic models play a key role. Dynamic models have long been key to our understanding of epidemic theory. Innovative models continue to be developed to deal with the challenges posed by pathogen evolution [96], complex immunological interactions [97], and host heterogeneity [98]. There has been increasing emphasis on the use of dynamic models in informing public health policy since the early 2000s when they played a key role in the response to the foot-and-mouth disease outbreak in the UK [99] and assessment of the risk from a smallpox-based bioterrorist attack [100, 101]. These uses have extended to endemic disease, such as a 2009 modeling analysis by Granich and colleagues [102] that highlighted the potential of "test-and-treat" strategies for HIV control. Recently, dynamic models have played an important role in guiding the response to emerging disease threats, from pandemic influenza [103], to multi-drug resistant organisms [104, 105] to MERS-CoV [106]. Many of the themes discussed throughout this manuscript have had a profound impact on these efforts, as does the need to report results and assumptions in a way accessible to policy makers.

Mechanistic models also crop up in other areas of epidemiology, often in less obvious ways. Nearly all of the key methods of genetic epidemiology are based on a mechanistic understanding of the underlying processes inheritance, mutation, and selective pressure. Social epidemiology at its core is based on the idea that our health depends on the behavior and health of those around us and, hence, has its own approaches to dependent happenings (though the terminology differs). Recently, there has been increasing interest in using mechanistic modeling approaches similar to those used for infectious disease to understand health phenomena that are, in part, socially driven, such as obesity [107]. Physiological measurements are often founded on mechanistic models of processes within the body (e.g., use of serum creatinine to approximate the glomerular filtration rate, a key measure of kidney function [108]).

Conclusion

Infectious disease dynamics is, perhaps, unique in epidemiology in the number of researchers that it brings from non-health related disciplines, particularly physics, computer science, and ecology. This, combined with the unique aspects of infectious disease systems, has contributed to the use of models that are distinct from "traditional" epidemiologic methods. However, the field is being transformed by the same forces that are transforming epidemiology in general: increasing access to technological tools and computational power; an explosion in the availability of data at the molecular, individual, and population levels; and a shift in what the important epidemiologic questions are as we eliminate old health threats and change our environment. Increasing emphasis on

principled statistical analysis in infectious disease modeling combined with an increasing need to deal with dynamic phenomena in epidemiologic inference opens up new opportunities for the cross pollination of ideas and the erosion of the historical barriers between epidemiologic fields.

Compliance with Ethical Standards

Conflict of Interest Justin Lessler, Andrew S. Azman, M. Kate Grabowski, and Isabel Rodriguez-Barraquer declare that they have no conflict of interest.

Henrik Salje declares grant funding from the NIH.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- • Of importance

1. Ross R. An application of the theory of probabilities to the study of a priori pathometry. Part I. Proceedings of the Royal Society A: Mathematical. Phys Eng Sci. 1916;92:204–30.
2. McCullagh P, Nelder JA. Generalized linear models, second edition. 2nd ed. Boca Raton, FL: Chapman and Hall/ CRC Press; 1989.
3. • Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford University Press, USA; 1991. **One of the key, must have references in the field. Lays out framework to think about dynamics of infectious diseases and presents basic theory and key results, with a focus on designing control and eradication strategies**.
4. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton, NJ: Princeton University Press; 2008.
5. Lui KJ, Darrow WW, Rutherford 3rd GW. A model-based estimate of the mean incubation period for AIDS in homosexual men. Science. 1988;240:1333–5.
6. May RM, Anderson RM. Transmission dynamics of HIV infection. Nature. 1987;326:137–42.
7. McEvoy M, Tillett HE. Some problems in the prediction of future numbers of cases of the acquired immunodeficiency syndrome in the UK. Lancet. 1985;2:541–2.
8. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009;324:1557–61.
9. • Cauchemez S, Fraser C, Van Kerkhove MD, Donnelly CA, Riley S, Rambaut A, et al. Middle East respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. Lancet Infect Dis. 2014;14:50–6. **Analysis of MERS outbreak that incorporated a wide array of different data sources including human mobility, phylogenetic and case data into mechanistic models to allow inference on key transmission parameters**.
10. WHO Ebola Response Team. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. N Engl J Med. 2014;371:1481–95.
11. Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DAT, et al. Hospital outbreak of Middle East respiratory syndrome coronavirus. N Engl J Med. 2013;369:407–16.
12. Salje H, Cauchemez S, Alera MT, Rodriguez-Barraquer I, Thaisomboonsuk B, Srikiatkhachorn A, et al. Reconstruction of 60 years of chikungunya epidemiology in the Philippines demonstrates episodic and focal transmission. J Infect Dis. 2016;213:604–10.
13. Imai N, Dorigatti I, Cauchemez S, Ferguson NM. Estimating dengue transmission intensity from sero-prevalence surveys in multiple countries. PLoS Negl Trop Dis. 2015;9:e0003719.
14. Rodriguez-Barraquer I, Buathong R, Iamsirithaworn S, Nisalak A, Lessler J, Jarman RG, et al. Revisiting Rayong: shifting seroprofiles of dengue in Thailand and their implications for transmission and control. Am J Epidemiol. 2014;179:353–60.
15. • Simons E, Ferrari M, Fricks J, Wannemuehler K, Anand A, Burton A, et al. Assessment of the 2010 global measles mortality reduction goal: results from a model of surveillance data. Lancet. 2012;379:2173–8. **Uses mechanistic models to make key inferences about the global burden of disease in the presence of imperfect data**.
16. Valle D, Clark J. Improving the modeling of disease data from the government surveillance system: a case study on malaria in the Brazilian Amazon. PLoS Comput Biol. 2013;9:e1003312.
17. Cauchemez S, Bhattarai A, Marchbanks TL, Fagan RP, Ostroff S, Ferguson NM, et al. Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. Proc Natl Acad Sci U S A. 2011;108:2825–30.
18. Cauchemez S, Carrat F, Viboud C, Valleron AJ, Boëlle PY. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. Stat Med. 2004;23:3469–87.
19. Lau MSY, Cowling BJ, Cook AR, Riley S. Inferring influenza dynamics and control in households. Proc Natl Acad Sci U S A. 2015;112:9094–9.
20. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. J R Soc Interface. 2009;6:187–202.
21. Sisson SA, Fan Y, Tanaka MM. Sequential Monte Carlo without likelihoods. Proc Natl Acad Sci U S A. 2007;104:1760–5.
22. Ratmann O, Donker G, Meijer A, Fraser C, Koelle K. Phylodynamic inference and model assessment with approximate Bayesian computation: influenza as a case study [Internet]. Kosakovsky Pond SL, editor. 2012. Available from: http://dx.plos.org/10.1371/journal.pcbi.1002835.
23. • Coudeville L, Baurin N, Vergu E. Estimation of parameters related to vaccine efficacy and dengue transmission from two large phase III studies. Vaccine [Internet]. 2015; Available from: http://dx.doi.org/10.1016/j.vaccine.2015.11.023. **Example of how sequential Monte Carlo methods can be used to parametrize complex transmission models**.
24. Lessler J, Metcalf CJE, Grais RF, Luquero FJ, Cummings DAT, Grenfell BT. Measuring the performance of vaccination programs using cross-sectional surveys: a likelihood framework and retrospective analysis [Internet]. Wallinga J, editor. 2011. Available from: http://dx.plos.org/10.1371/journal.pmed.1001110.
25. Abeku TA, de Vlas SJ, Gerard B, Awash T, Asnakew K, Dereje O, et al. Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: a simple seasonal adjustment method performs best. Trop Med Int Health. 2002;7:851–7.

26. Chaves LF, Pascual M. Climate cycles and forecasts of cutaneous leishmaniasis, a nonstationary vector-borne disease. PLoS Med. 2006;3:e295.
27. Pascual M, Rodó X, Ellner SP, Colwell R, Bouma MJ. Cholera dynamics and El Niño-Southern Oscillation. Science. 2000;289:1766–9.
28. Reich NG, Shrestha S, King AA, Rohani P, Lessler J, Kalayanarooj S, et al. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. J R Soc Interface. 2013;10:20130414.
29. Gatto M, Mari L, Bertuzzo E, Casagrandi R, Righetto L, Rodriguez-Iturbe I, et al. Generalized reproduction numbers and the prediction of patterns in waterborne disease. Proc Natl Acad Sci U S A. 2012;109:19703–8.
30. Reiner Jr RC, Stoddard ST, Scott TW. Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. Epidemics. 2014;6:30–6.
31. Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. Proc Natl Acad Sci U S A. 2015;112:11887–92.
32. Yang W, Cowling BJ, Lau EHY, Shaman J. Forecasting influenza epidemics in Hong Kong. PLoS Comput Biol. 2015;11:e1004383.
33. Meltzer MI, Atkins CY, Santibanez S, Knust B, Petersen BW, Ervin ED, et al. Estimating the future number of cases in the Ebola epidemic—Liberia and Sierra Leone, 2014-2015. MMWR Suppl. 2014;63:1–14.
34. Gomes MFC, Pastore Y Piontti A, Rossi L, Chao D, Longini I, Halloran ME, et al. Assessing the international spreading risk associated with the 2014 West African Ebola outbreak. PLoS Curr. [Internet]. 2014;6. Available from: http://dx.doi.org/10.1371/currents.outbreaks.cd818f63d40e24aef769dda7df9e0da5.
35. Perkins A, Alex P, Amir S, Ruktanonchai CW, Moritz K, Andrew T. Model-based projections of Zika virus infections in childbearing women in the Americas [Internet]. 2016. Available from: http://dx.doi.org/10.1101/039610.
36. Drake JM, Kaul RB, Alexander LW, O'Regan SM, Kramer AM, Pulliam JT, et al. Ebola cases and health system demand in Liberia. PLoS Biol. 2015;13:e1002056.
37. Lindström T, Tildesley M, Webb C. A Bayesian ensemble approach for epidemiological projections. PLoS Comput Biol. 2015;11:e1004187.
38. Ruiz D, Daniel R, Cyrille B, Connor SJ, Omumbo JA, Bradfield L, et al. Testing a multi-malaria-model ensemble against 30 years of data in the Kenyan highlands. Malar J. 2014;13:206.
39. Thomson MC, Doblas-Reyes FJ, Mason SJ, Hagedorn R, Connor SJ, Phindela T, et al. Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. Nature. 2006;439:576–9.
40. Shaman J, Karspeck A, Yang W, Tamerius J, Lipsitch M. Real-time influenza forecasts during the 2012-2013 season. Nat Commun. 2013;4:2837.
41. Moore SM, Andrew M, Griffith KS, Titus A, Mead PS, Eisen RJ. Improvement of disease prediction and modeling through the use of meteorological ensembles: human plague in Uganda. PLoS One. 2012;7:e44431.
42. Smith T, Ross A, Maire N, Chitnis N, Studer A, Hardy D, et al. Ensemble modeling of the likely public health impact of a pre-erythrocytic malaria vaccine. PLoS Med. 2012;9:e1001157.
43. Epidemic Prediction Initiative [Internet]. [cited 2016 Mar 29]. Available from: http://predict.phiresearchlab.org/flu/index.html.
44. Chikungunya threat inspires new DARPA challenge [Internet]. 2014 [cited 2016 Mar 29]. Available from: http://www.sciencemag.org/news/2014/08/chikungunya-threat-inspires-new-darpa-challenge.
45. US Department of Commerce, NOAA, National Weather Service. Dengue Forecasting. NOAA's National Weather Service; [cited 2016 Mar 29]; Available from: http://dengueforecasting.noaa.gov/.
46. Yasmin S. Ebola infections fewer than predicted by disease models [Internet]. Scientific American. [cited 2016 Mar 29]. Available from: http://www.scientificamerican.com/article/ebola-infections-fewer-than-predicted-by-disease-models/.
47. Dumbill E. Planning for big data. "O'Reilly Media, Inc."; 2012.
48. Wesolowski A, Metcalf CJE, Eagle N, Kombich J, Grenfell BT, Bjørnstad ON, et al. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. Proc Natl Acad Sci U S A. 2015;112:11114–9.
49. Erlander S, Stewart NF. The gravity model in transportation analysis: theory and extensions. Utretch: VSP; 1990.
50. 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. [Internet]. 2014;6. Available from: http://dx.doi.org/10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e.
51. Tatem AJ, Huang Z, Narib C, Kumar U, Kandula D, Pindolia DK, et al. Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. Malar J. 2014;13:52.
52. Wesolowski A, Eagle N, Noor AM, Snow RW, Buckee CO. The impact of biases in mobile phone ownership on estimates of human mobility. J R Soc Interface. 2013;10:20120986.
53. Eischeid JK, Bruce Baker C, Karl TR, Diaz HF. The quality control of long-term climatological data using objective data analysis. J Appl Meteorol. 1995;34:2787–95.
54. Sorichetta A, Hornby GM, Stevens FR, Gaughan AE, Linard C, Tatem AJ. High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. Sci Data. 2015;2:150045.
55. Hay SI, Guerra CA, Gething PW, Patil AP, Tatem AJ, Noor AM, et al. A world malaria map: Plasmodium falciparum endemicity in 2007. PLoS Med. 2009;6:e1000048.
56. Hay SI, Snow RW. The malaria Atlas Project: developing global maps of malaria risk. PLoS Med. 2006;3:e473.
57. Pigott DM, Bhatt S, Golding N, Duda KA, Battle KE, Brady OJ, et al. Global distribution maps of the leishmaniases. Elife [Internet]. 2014;3. Available from: http://dx.doi.org/10.7554/eLife.02851.
58. Rogers DJ, Suk JE, Semenza JC. Using global maps to predict the risk of dengue in Europe. Acta Trop. 2014;129:1–14.
59. Pigott DM, Golding N, Mylne A, Huang Z, Weiss DJ, Brady OJ, et al. Mapping the zoonotic niche of Marburg virus disease in Africa. Trans R Soc Trop Med Hyg. 2015;109:366–78.
60. Garni R, Tran A, Guis H, Baldet T, Benallal K, Boubidi S, et al. Remote sensing, land cover changes, and vector-borne diseases: use of high spatial resolution satellite imagery to map the risk of occurrence of cutaneous leishmaniasis in Ghardaïa. Algeria Infect Genet Evol. 2014;28:725–34.
61. Kraemer MUG, Sinka ME, Duda KA, Mylne AQN, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife. 2015;4:e08347.
62. Pigott DM, Golding N, Mylne A, Huang Z, Henry AJ, Weiss DJ, et al. Mapping the zoonotic niche of Ebola virus disease in Africa. Elife. 2014;3:e04395.
63. Takahashi S, Metcalf CJE, Ferrari MJ, Moss WJ, Truelove SA, Tatem AJ, et al. Reduced vaccination and the risk of measles and other childhood infections post-Ebola. Science. 2015;347:1240–2.
64. Althouse BM, Scarpino SV, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. EPJ Data Sci. [Internet].

2015;4. Available from: http://www.epjdatascience.com/content/4/1/17.
65. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature. 2009;457:1012–4.
66. Patwardhan A, Bilkovski R. Comparison: flu prescription sales data from a retail pharmacy in the US with Google Flu trends and US ILINet (CDC) data as flu activity indicator. PLoS One. 2012;7:e43611.
67. Althouse BM, Ng YY, Cummings DAT. Prediction of dengue incidence using search query surveillance. PLoS Negl Trop Dis. 2011;5:e1258.
68. McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. PLoS Comput Biol. 2014;10:e1003581.
69. Azman AS, Lessler J, Satter SM, Mckay MV, Khan A, Ahmed D, et al. Tracking cholera through surveillance of oral rehydration solution sales at pharmacies: insights from urban Bangladesh. PLoS Negl Trop Dis. 2015;9:e0004230.
70. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci U S A. 2015;112:14473–8.
71. Volz EM, Koelle K, Bedford T. Viral phylodynamics. PLoS Comput Biol. 2013;9:e1002947.
72. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. Science. 2004;303:327–32.
73. Kingman JFC. The coalescent. Stochastic Process Appl. 1982;13:235–48.
74. Volz EM, Kosakovsky Pond SL, Ward MJ, Leigh Brown AJ, Frost SDW. Phylodynamics of infectious disease epidemics. Genetics. 2009;183:1421–30.
75. Magiorkinis G, Sypsa V, Magiorkinis E, Paraskevis D, Katsoulidou A, Belshaw R, et al. Integrating phylodynamics and epidemiology to estimate transmission diversity in viral epidemics. PLoS Comput Biol. 2013;9:e1002876.
76. Frost SDW, Volz EM. Viral phylodynamics and the search for an "effective number of infections". Philos Trans R Soc Lond B Biol Sci. 2010;365:1879–90.
77. Stadler T, Kouyos R, von Wyl V, Yerly S, Böni J, Bürgisser P, et al. Estimating the basic reproductive number from viral sequence data. Mol Biol Evol. 2012;29:347–57.
78. Bezemer D, Cori A, Ratmann O, van Sighem A, Hermanides HS, Dutilh BE, et al. Dispersion of the HIV-1 epidemic in men who have sex with men in the Netherlands: a combined mathematical model and phylogenetic analysis. PLoS Med. 2015;12:e1001898. discussion e1001898.
79. Bedford T, Riley S, Barr IG, Broor S, Chadha M, Cox NJ, et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. Nature. 2015;523:217–20.
80. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. Science. 2014;346:56–61.
81. Li LM, Grassly NC, Fraser C. Genomic analysis of emerging pathogens: methods, application and future trends. Genome Biol. 2014;15:541.
82. Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science. 2014;345:1369–72.
83. Volz E, Pond S. Phylodynamic analysis of Ebola virus in the 2014 Sierra Leone epidemic. PLoS Curr. [Internet]. 2014;6. Available from: http://dx.doi.org/10.1371/currents.outbreaks.6f7025f1271821d4c815385b08f5f80e.
84. Stadler T, Kühnert D, Rasmussen DA, du Plessis L. Insights into the early epidemic spread of Ebola in Sierra Leone provided by viral sequence data. PLoS Curr. [Internet]. 2014;6. Available from: http://dx.doi.org/10.1371/currents.outbreaks.02bc6d927ecee7bbd33532ec8ba6a25f.
86.• Ratmann O, van Sighem A, Bezemer D, Gavryushkina A, Jurriaans S, Wensing A, et al. Sources of HIV infection among men having sex with men and implications for prevention. Sci Transl Med. 2016;8:320ra2. **Key example showing that clinical and phylogenetic data can be used to identify predominant sources of ongoing viral transmission and to assess the potential effectiveness of intervention packages**.
86. Volz EM, Ionides E, Romero-Severson EO, Brandt M-G, Mokotoff E, Koopman JS. HIV-1 transmission during early infection in men who have sex with men: a phylodynamic analysis. PLoS Med. 2013;10:e1001568. **discussion e1001568**.
87. Rasmussen DA, Ratmann O, Koelle K. Inference for nonlinear epidemiological models using genealogies and time series. PLoS Comput Biol. 2011;7:e1002136.
88. Volz EM. Complex population dynamics and the coalescent under neutrality. Genetics. 2012;190:187–201.
90.• Rasmussen DA, Volz EM, Koelle K. Phylodynamic inference for structured epidemiological models. PLoS Comput Biol. 2014;10:e1003570. **Uses innovative methods to combine mechanistic models within phylodynamics to estimate transmission parameters**.
90. Frost SDW, Volz EM. Modelling tree shape and structure in viral phylodynamics. Philos Trans R Soc Lond B Biol Sci. 2013;368:20120208.
91. Leventhal GE, Kouyos R, Stadler T, von Wyl V, Yerly S, Böni J, et al. Inferring epidemic contact structure from phylogenetic trees. PLoS Comput Biol. 2012;8:e1002413.
92. Rasmussen DA, Boni MF, Koelle K. Reconciling phylodynamics with epidemiology: the case of dengue virus in southern Vietnam. Mol Biol Evol. 2014;31:258–71.
93. O'Dea EB, Wilke CO. Contact heterogeneity and phylodynamics: how contact networks shape parasite evolutionary trees. Interdiscip. Perspect. Infect. Dis. [Internet]. Hindawi Publishing Corporation; 2010 [cited 2016 Mar 29];2011. Available from: http://www.hindawi.com/journals/ipid/2011/238743/.
94. Robinson K, Fyson N, Cohen T, Fraser C, Colijn C. How the dynamics and structure of sexual contact networks shape pathogen phylogenies. PLoS Comput Biol. 2013;9:e1003105.
95. Worby CJ, Lipsitch M, Hanage WP. Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. PLoS Comput Biol. 2014;10:e1003549.
96. Chan CH, McCabe CJ, Fisman DN. Core groups, antimicrobial resistance and rebound in gonorrhoea in North America. Sex Transm Infect BMJ Publishing Group Ltd. 2012;88:200–4.
97. Kucharski AJ, Gog JR. Age profile of immunity to influenza: effect of original antigenic sin. Theor Popul Biol. 2012;81:102–12.
98. House T, Keeling MJ. Insights from unifying modern approximations to infections on networks. J R Soc Interface. 2011;8:67–73.
99. Keeling MJ, Woolhouse ME, Shaw DJ, Matthews L, Chase-Topping M, Haydon DT, et al. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science. 2001;294:813–7.
100. Halloran ME, Longini Jr IM, Nizam A, Yang Y. Containing bioterrorist smallpox. Science. 2002;298:1428–32.
101. Kaplan EH, Craft DL, Wein LM. Emergency response to a smallpox attack: the case for mass vaccination. Proc Natl Acad Sci U S A. 2002;99:10935–40.
102. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet Elsevier. 2009;373:48–57.
103. Ferguson NM, Cummings DAT, Cauchemez S, Fraser C, Riley S, Meeyai A, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature. 2005;437:209–14.

104. Grad YH, Goldstein E, Lipsitch M, White PJ. Improving control of antibiotic-resistant gonorrhea by integrating research agendas across disciplines: key questions arising from mathematical modeling. J Infect Dis. 2016;213:883–90.
105. Cohen T, Murray M. Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness. Nat Med. 2004;10:1117–21.
106. Lessler J, Rodriguez-Barraquer I, Cummings DAT, Garske T, Van Kerkhove M, Mills H, et al. Estimating potential incidence of MERS-CoV associated with Hajj pilgrims to Saudi Arabia, 2014. PLoS Curr. [Internet]. 2014;6. Available from: http://dx.doi.org/10.1371/currents.outbreaks.c5c9c9abd636164a9b6fd4dbda974369.
107. El-Sayed AM, Scarborough P, Seemann L, Galea S. Social network analysis and agent-based modeling in social epidemiology. Epidemiol Perspect Innov. 2012;9:1.
108. Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med. 2012;367:20–9.