



Mathematical modeling of infectious disease dynamics

Constantinos I. Siettos & Lucia Russo

To cite this article: Constantinos I. Siettos & Lucia Russo (2013) Mathematical modeling of infectious disease dynamics, *Virulence*, 4:4, 295-306, DOI: [10.4161/viru.24041](https://doi.org/10.4161/viru.24041)

To link to this article: <https://doi.org/10.4161/viru.24041>



Copyright © 2013 Landes Bioscience



Published online: 03 Apr 2013.



Submit your article to this journal [↗](#)



Article views: 3501



View related articles [↗](#)



Citing articles: 49 View citing articles [↗](#)

Mathematical modeling of infectious disease dynamics

Constantinos I. Siettos^{1,*} and Lucia Russo²

¹School of Applied Mathematics and Physical Sciences; National Technical University of Athens; Athens, Greece; ²Consiglio Nazionale di Ricerche; Napoli, Italy

Keywords: mathematical epidemiology, statistical models, dynamical models, agent-based models, machine learning models

Over the last years, an intensive worldwide effort is speeding up the developments in the establishment of a global surveillance network for combating pandemics of emergent and re-emergent infectious diseases. Scientists from different fields extending from medicine and molecular biology to computer science and applied mathematics have teamed up for rapid assessment of potentially urgent situations. Toward this aim mathematical modeling plays an important role in efforts that focus on predicting, assessing, and controlling potential outbreaks. To better understand and model the contagious dynamics the impact of numerous variables ranging from the micro host–pathogen level to host-to-host interactions, as well as prevailing ecological, social, economic, and demographic factors across the globe have to be analyzed and thoroughly studied. Here, we present and discuss the main approaches that are used for the surveillance and modeling of infectious disease dynamics. We present the basic concepts underpinning their implementation and practice and for each category we give an annotated list of representative works.

Introduction

No doubt, the history of mankind has been shaped by the pitiless outbreaks of infectious disease pandemics. Whole nations and civilizations have been wiped off the map through the ages. The list is long: biblical pharaonic plagues that hit Ancient Egypt in the middle of Bronze Age around 1715 BC,¹ the “λοιμός” in Athens from 430 to 425 BC set the end of the Periclean golden era, the “cocoliztli” epidemics, which occurred during the 16th century, resulted in some 13 million deaths, decimating the Mesoamerican native population,² the Black Death bubonic plague burst in Europe in 1348, and is estimated to have killed over 25 million people in just five years. The pandemic influenza virus of 1918–1919 swept through America, Europe, Asia, and Africa smashing the globe: the death toll was around 40 million people. Two one-year, less severe influenza pandemics followed in the next decades: the 1957 and the 1963 influenza pandemics resulted to two and one million deaths respectively (World Health Organization: <http://apps.who.int/iris/handle/10665/68985>). In the last decades emerging and re-emerging epidemics such as

AIDS, measles, malaria, and tuberculosis cause death to millions of people each year. According to the UNAIDS report on the global AIDS epidemic, an estimated 34 million people, including 3.4 million children, were living with HIV worldwide at the end of 2010, while the related deaths and new infections were 1.8 and 2.7 million, respectively.³

The rapid technological and theoretical progress has dramatically enhanced our arsenal in fighting epidemics and we are getting better on it. The global surveillance network is growing under an intensive worldwide effort. We are now able to produce effective vaccines and antiviral drugs and knowledge goes deep in details such as the molecular structure of a variety of viruses. A large and intensive research is evolving for the design of better drugs and vaccines. Yet, studies warn us that a new pandemic—influenza-type is the most worrisome one—is sooner or later on the way.⁴ The critical question(s) is not whether but when it will arise, how it is going to spread, how deadly it will be, who should get the vaccine when not all can, how likely are multiple waves of re-emergence and what type of intervention may be applied to stop the spread. Unfortunately, even with all the advances, we still don't have robust answers.

The problem stems mainly from two reasons: (1) the continuous and ever-lasting mutations of the viruses, and (2) the complexity in the disease transmission mechanism. Unfortunately, the odds are that in a real crisis, even if researchers succeed to come up with a vaccine tailor-made for an emerged virus strain, it is doubtful that it would stop a pandemic.⁵

The complex multi-scale interplay between a host of factors ranging from the micro host–pathogen and individual-scale host–host interactions to macro-scale ecological, social, economic, and demographic conditions across the globe complicated by technical issues such as the time lag between vaccine prototype development and commercial production and distribution imposes a real impediment to our control strategy potential.

Mathematical, statistical models and computational engineering are playing a most valuable role in shedding light on the problem and for helping make decisions.

The Beginning of Mathematical Modeling in Epidemiology

The very first publication addressing the mathematical modeling of epidemics dates back in 1766. In this seminal paper, *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole*,⁶

*Correspondence to: Constantinos I. Siettos; Email: ksiet@mail.ntua.gr
Submitted: 12/22/12; Revised: 02/15/13; Accepted: 02/18/13
<http://dx.doi.org/10.4161/viru.24041>

Daniel Bernoulli developed a mathematical model to analyze the mortality due to smallpox in England, which at that time was one in 14 of the total mortality. Bernoulli used his model to show that inoculation against the virus would increase the life expectancy at birth by about three years. A translation in English and review of this work can be found in Sally Blower (2004),⁷ while a revision of the main findings and a presentation of the criticism by D'Alembert appears in Dietz and Heesterbeek (2002).⁸ Lambert, in 1772, followed up the work of Bernoulli extending the model by incorporating age-dependent parameters.⁹ Laplace has also worked on the same concept.¹⁰ However this line of research has not been developed systematically until the benchmark paper of Ross in 1911, which actually establishes modern mathematical epidemiology.¹¹ In this work, Ross addressed the mechanistic a priori modeling approach using a set of equations to approximate the discrete-time dynamics of malaria through the mosquito-borne pathogen transmission (for a discussion and a review of this model see also Smith et al. [2012]¹²).

Following up the work of Ross, Kermack and McKendrick published three seminal papers which founded the deterministic compartmental epidemic modeling.^{13–15} In these papers, they addressed the mass-action incident in disease transmission cycle, suggesting that the probability of infection of a susceptible (virgin from illness) is analogous to the number of its contacts with infected individuals. Hence, the rate at which susceptibles become infected is given by kSI where S and I represent population densities of susceptible and infected people, respectively. In this context, the rate at which infected individuals become recovered is given by λI , while the rate at which recovered individuals become again susceptible is given by μR ; k , λ and μ are analogy constants. This mechanistic-deterministic representation holds strong analogy to the Law of Mass Action¹⁶ introduced by Guldberg and Waage in 1864 and is called the SIR model, implying homogeneous mixing of the contacts and conservation of the total mass (population) as well as relatively low rates of interaction. Forty years after the paper of Ross, MacDonald extended Ross's model to explain in depth the transmission process of malaria and propose methods for eradicating the disease on an operational level. Due to the importance of MacDonald's contribution to the field by exploiting the use of computers, mathematical models for the dynamics and the control of mosquito-transmitted pathogens are known as Ross–MacDonald models.¹²

At this point it would be remiss of us not to mention the work of Enko,^{17–19} who in 1889 published a remarkable probabilistic model for describing the epidemic of measles in discrete time. With the use of the model, Enko evaluated the number of contacts between infectives and susceptibles in the population. The model of Enko is the precursor of the famous Reed-Frost chain binomial model introduced by W.H. Frost in 1928 in biostatistics lectures at Johns Hopkins University (not published then in a journal, but published in 1979²⁰). This model assumes that the infection spreads from an infected to a susceptible individual through discrete time Markov chain events. This representation set the basis of contemporary stochastic epidemic modeling.

Mathematical Modeling Methodologies in Epidemiology

Mathematical modeling and simulation allows for rapid assessment. Simulation is also used when the cost of collecting data is prohibitively expensive, or there are a large number of experimental conditions to test. Over the years, a vast number of approaches have been proposed looking at the problem from different perspectives. These encompass three general categories (see Fig. 1): (1) statistical methods for surveillance of outbreaks and identification of spatial patterns in real epidemics, (2) mathematical models within the context of dynamical systems (also called state-space models) used to forecast the evolution of a “hypothetical” or on-going epidemic spread, and (3) machine learning/expert methods for the forecasting of the evolution of an ongoing epidemic. For all three of these categories there are again different approaches weaving a big and diverse literature. Here, we try to draw the map of these approaches and try to describe their basic underpinning concepts.

Statistical-Based Methods for Epidemic Surveillance

One of the most important aspects in epidemics revolves around the surveillance, early detection of possible outbreaks and patterns that may help controlling a spread. One of the very first success stories in the area is the modeling of cholera epidemic that swept through London in 1854. At that time John Snow, a physician, collected spatiotemporal data and by visualizing them in a map found that there was a particular pattern around the Broad Street water pump,²¹ which actually was the zero point of transmission. His analysis helped eradicate the disease. In the dawn of 20th century Greenwood an epidemiologist and statistician was the first Professor of Epidemiology and Statistics at the London School of Hygiene and Tropical Diseases establishing a rigorous mathematical connection between fields.²²

Today, global initiatives to combat epidemics require effective domestic action mechanisms and preparedness through the globe. An intensive worldwide effort led by World Health Organization and Centers for Disease Control is speeding up the developments for the establishment of a global surveillance network. New emerged pandemics such as the AIDS, the severe acute respiratory syndrome (SARS) of 2002–2003 and the H1N1 swine flu of 2009 pandemics reminds us about the importance of surveillance and prompt outbreak detection. Toward this aim, statistical methods have enhanced our potential in fighting epidemics allowing for rapid assessment of emerging situations. Obviously, the correctness of the data and the selection of the appropriate methodology are crucial for the construction of statistical models that can capture in an efficient robust way the communicable disease characteristics.

To date, several statistical methods have been proposed (see also Unkel et al. [2012]²³ for a review of statistical methods for the detection of disease outbreaks). In the website of Centers for Disease Control and Prevention (CDC) (<http://www.cdc.gov/>) one can find a list of references in the field. Here we present

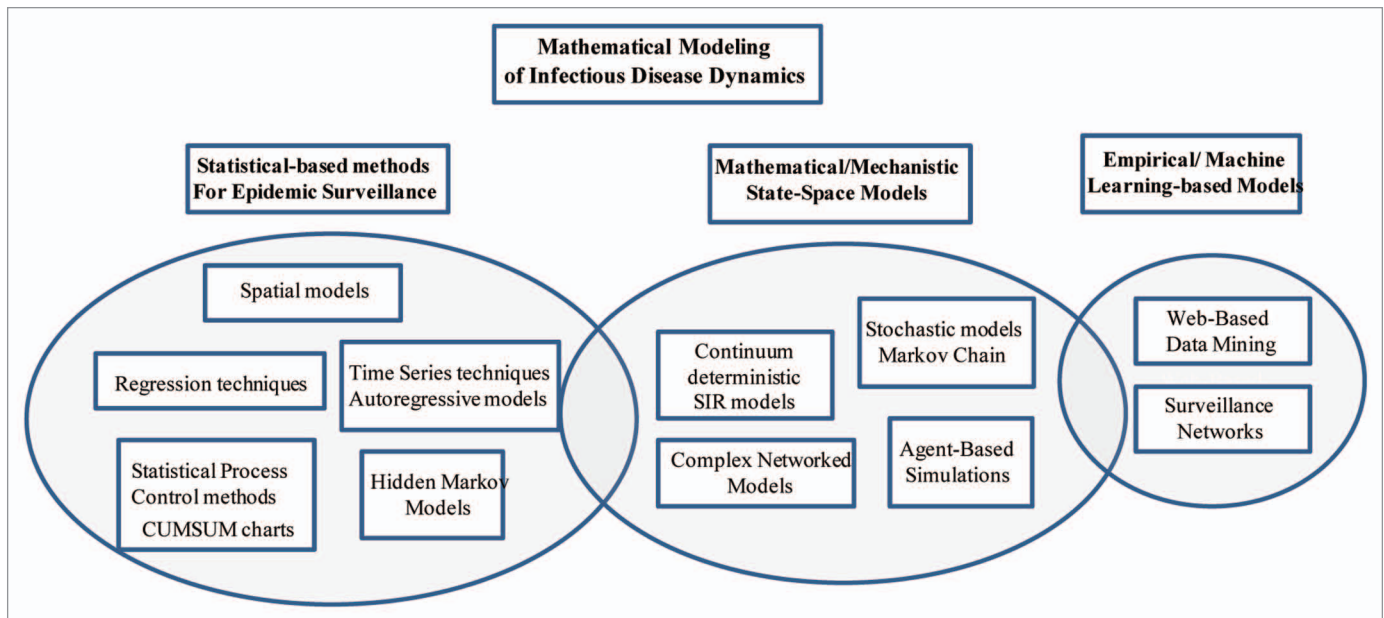


Figure 1. An overview of mathematical models for infectious diseases.

and discuss the most common schemes that can be classified as follows:

Regression methods.²⁴⁻²⁹ Regression models try to detect an outbreak from time-series of epidemic-free periods by monitoring a statistic of reported infected cases, say $y(t)$. An epidemic alert is raised when a certain threshold, say k , is surpassed, defined by

$|y(t) - \mu| > k$, (μ being the mean value of the time-series distribution) within a confidence interval (usually of 95%).

A basic regression model is that proposed from Serfling which was initially constructed to monitor the deaths of influenza based on the seasonal pattern of pneumonia and influenza deaths.²⁴ Due to the seasonal behavior of the disease the following cyclic regression model has been addressed:

$$y(t) = a + bt + \sum_{i=1}^m c_i \cos \theta + \sum_{i=1}^m d_i \sin \theta + e(t)$$

θ is a linear function of time t while the coefficients are to be determined by a parameter identification technique. The cosine and sine terms are used to approximate cyclical seasonal patterns; $e(t)$ is the noise (assumed that is Gaussian distributed with mean zero and variance σ^2) which is estimated from the time-series. In the original paper of Serfling, $y(t)$ was the expected mean value of total deaths due to pneumonia and influenza in units of 4-weeks periods. The model was fitted using data from 108 US cities for a 3 year period starting in September of 1955.

Using least squares estimation Serfling ended up to the following model:

$$y(t) = 300.5 + 2.1t + 97.6 \cos \left(\frac{2\pi t}{13} - 2.67 \right)$$

Other models including square terms, t^2 , to account long-term changes due to factors such as the population growth or disease

assessment have been also proposed.²⁷ Today, the above approach is used by the Centers for Disease Control in the US, Australia, France, and Italy for the detection of influenza outbreaks.

While this approach is very popular among epidemiologists for predicting and surveillance purposes, one has to be cautious about their use as the form of the equations relies usually on ad hoc assumptions on the dependence between the dynamics of a disease and the independent factors (variables) that determine its spread. In addition, the choice of the model (linear/nonlinear), assumptions on the statistical properties (for example independence, normal distribution and fixed variance) of the unmodeled dynamics (represented by $e(t)$) flash a “note of caution” in their use especially for the surveillance and prediction of outbreaks of new emerging epidemics.

Times series analysis based on autoregressive models such as the autoregressive integrated moving average model (ARIMA) and seasonal ARIMA (SARIMA)³⁰⁻³³ as well as **neural networks**.³⁴ These models relax the hypothesis of autocorrelation of regression models as well as the hypothesis of simple autoregressive models such as AR (autoregressive) and ARMA (autoregressive moving) in which past disturbances are not modeled. In this category, ARIMA models are the most commonly used. Their general form reads:

$$A(z^{-1})[\Delta^d y(t) - \mu] = B(z^{-1})e(t)$$

where $y(t)$ denotes a stationary stochastic process at time t with mean value $E(y(t)) = \mu$; z^{-1} is the backward shift operator defined by $z^{-k}y(t) = y(t - k)$ and Δ^d is the differencing operator of order d defined by $\Delta^d = (1 - z^{-1})^d$; $A(z^{-1})$ is the autoregressive operator

defined as $A(z^{-1}) = 1 + a_1 z^{-1} + \dots + \alpha_n z^{-n_a}$; $B(z^{-1})$ is the moving-average operator defined by $B(z^{-1}) = 1 + b_1 z^{-1} + \dots + b_{n_b} z^{-n_b}$;

$e(t)$ is the residual (noise) at time t representing the part of the measurement that cannot be predicted from previous measurements. For $d = 0$ and $n_a = 0$ one gets the moving average model, while for $d = 1$, $n_a = n_b = 0$ one gets the random walk with drift. Seasonal differencing enters naturally in the above framework by considering the seasonal differencing operator

$\Delta_s^d \equiv (1 - z^{-1})^d (1 - z^{-k})^S$ where k is the length of seasonal cycle and S is the degree of seasonal differencing producing series of changes from one season to the next.

The time-series is then split in two sets: one containing the times-series serving as a training set, and another one containing the remaining data serving as a test (validation) set. The Akaike Information Criterion³⁵ is usually applied to identify the optimal model order by compromising between the goodness-of-fit and number of parameters. The fitted model is then used for the forecasting of disease evolution. The reliability of such approaches is limited mostly by (1) the statistical uncertainty related to the estimation of the values of the unknown parameters and (2) the hypotheses related to the statistical properties of the corresponding time series.

Statistical process control methods including cumulative sum (CUSUM) charts³⁶⁻⁴¹ and exponentially weighted moving average (EWMA)^{42,43}-based methods. CUSUM is probably the most common used technique for the detection of disease outbreaks. This is achieved by monitoring a cumulative performance measure over time. Let us consider the number of infected cases $y(t_i)$ as observed at different time instances t_i , $i = 1, 2, \dots, n$. In its simple representation, for a single parameter process, CUSUM is defined as

$$CUSUM(i) = \sum_{l=1}^i (y(t_l) - k)$$

or in a recursive form as
 $CUSUM(0) = 0$

$$CUSUM(i) = \max(0, CUSUM(i-1) + y(t_i) - k), i \geq 0$$

where k is a reference value corresponding to the difference between to the in-control and the out-of-control mean. The process is considered to be in-control if $CUSUM(i) < h$ with h denoting a threshold (its value is usually taken to be three times the standard deviation from the baseline/mean value of in-control-observations). An alarm is raised at time t_i if $CUSUM(i)$ exceeds h ; the process is considered to be out-of-control. The reference value k is determined by likelihood ratio based methods.⁴⁴⁻⁴⁹ Hence, denoting by $f(\theta_0)$ and $f(\theta_1)$ the probability function of the in-control and out-of-control processes with parameters θ_0 and θ_1 respectively, the reference value reads:

$$k = -\log \frac{f(\theta_1 | y(t_i))}{f(\theta_0 | y(t_i))}$$

The probability functions $f(\theta_0)$ and $f(\theta_1)$ and their parameters can be estimated using data from past periods. For Poisson distributions the above relation reads:

$$k = \frac{\mu_1 - \mu_0}{\log(\mu_1) - \log(\mu_0)}$$

where μ_0 and μ_1 are the mean values of the in-control and out-of-control Poisson distributions.

For an epidemic that involves time-varying characteristics, such as seasonality, the reference parameter is now time-varying itself, i.e., $k = k(t)$.

The EWMA control chart method monitors infectious disease dynamics using the following recursive statistical estimator, which in its simple form reads:

$$z(t_0) = z(0) = 0$$

$$z(t_i) = (1 - \gamma)z(t_{i-1}) + \gamma y(t_i), i \geq 1$$

γ is a “forgetting” factor, a number between 0 and 1 which weights the significance of past values. Actually this factor reduces the importance of past observed information in estimating future. Again, an alarm is raised at time t_i if $z(t_i) > h$.

Other statistical process control methods such as temporal scan statistics have been also used.^{46,50,51}

Hidden Markov models (HMM) used to explain statistical correlation in time series.^{52,53} The question that the HMMs come to answer in epidemiology is the following: how can we infer about the dynamics of a particular infectious disease and forecast its outbreak when we cannot monitor/record explicitly the characteristics of the disease but we can observe some possible indicators of the disease? For example, can we forecast the evolution of an influenza epidemic by monitoring for example the number of reported cases as recorded through a surveillance network of physicians or in hospital units?^{52,54} HMM models are exploited exactly under these limitations/ constraints. Within this context, let us denote by $Y(t)$ the stochastic process of the unobserved (hidden) state, e.g., the number of cases of the disease in the population at time t and with $O(t)$ the stochastic process of the observable states.

Formally, HMMs are Markov processes, i.e., stochastic processes which satisfy the so called Markov property (here for the sake of presentation we assume discrete in time Markov processes) defined by:

$$P(Y(t) = y(t) | Y(t-1) = y(t-1), Y(t-2) = y(t-2), \dots, Y(1) = y(1)) = P(Y(t) = y(t) | Y(t-1) = y(t-1))$$

along with the time-invariant transition probability between two realizations, say $y_i(\cdot), y_j(\cdot)$:

$$a_{ij} = P(Y(t) = y_j(t) | Y(t-1) = y_i(t-1)), \sum_{j=1}^n a_{ij} = 1, j=1, 2, \dots, n$$

The above relations simply state that all the necessary information for predicting the distribution of $Y(t)$ at time $Y(t)$ with a certain probability defined by $P(\cdot)$ is contained within $Y(t-1)$; $y(\cdot)$ denotes a realization of the stochastic process $Y(\cdot)$.

In HMMs, the following conditional independence assumption holds:

$$P(O(t)=o(t)|Y(t)=y(t))= \\ = P\left(O(t)=o(t)|O(t-1)=o(t-1), O(t-2)=o(t-2), \dots, O(1)=o(1), \right. \\ \left. Y(t)=y(t), Y(t-1)=y(t-1), Y(t-2)=y(t-2), \dots, Y(1)=y(1)\right)$$

Here, the transition probability between an observed, say $o_j(\cdot)$, and a hidden state, say $y_i(\cdot)$, is defined as

$$b_{ij} = P(O(t)=o_j(t)|Y(t)=y_i(t)), \sum_{i=1}^n b_{ij} = 1, j=1, 2, \dots, m$$

There are three basic questions that have to be answered here: (1) what is the likelihood of the observed sequence, (2) what is the most likely hidden sequence given a_{ij}, b_{ij} and the observation sequence, and (3) given the observation sequence, which are the HMM parameters, i.e., a_{ij}, b_{ij} and initial distribution of observed states that maximize the likelihood of the observation sequence and/or hidden sequence. The first problem is usually tackled with the use of the forward-backward algorithm,⁵⁵ the second problem with the use of the Viterbi algorithm,⁵⁶ and the third problem with the use of the so-called Expectation-Maximization (EM) algorithm.⁵⁷

Spatial models for monitoring, identifying and forecasting disease outbreaks in different locations.⁵⁸⁻⁶⁴ Most of the infectious diseases result to strong spatio-temporal patterns whose systematic analysis is of utmost importance for better understanding, predicting and combating outbreaks. Spatial surveillance requires the use of multivariate techniques.⁶⁵ Most of the multivariate methods can be viewed as extensions of standard univariate methods—as the ones described above—; however, there are others such as clustering, principal component analysis (PCA) based methods that do not have a common ancestor with univariate ones.⁶⁶ Kleinshmidt et al. (2000) used a two tier approach for the surveillance of malaria.⁶⁷ They used regression analysis on the larger scale and kriging⁶⁸ to interpolate the count data at an unobserved location in order to forecast the prevalence of the disease in the local scale. Cohen et al. (2010) exploited PCA to create a single surveillance index that can be used to summarize temporal and spatial trends of malaria in India.⁶⁹ Coleman et al. (2009) used the SatScan freeware software (<http://www.satscan.org/>) to identify malaria outbreaks to a province of South Africa by detecting time and space clusters.⁷⁰ The SatScan software is based on the spatial scan statistic^{71,72} and the Bernoulli spatial model.⁷³ SatScan has been also exploited by Gaudart et al. (2006) to identify spatio-temporal clusters of high risk incidence of malaria in a Mali village.⁷⁴ A temporal analysis using ARIMA technique was also undertaken.

To this end, we should also mention the use of copulas^{75,76} (joint distribution functions used to model the dependencies between random variables based on given/known marginal distributions of the individual variables) for parametric multivariate analysis. Copulas can be integrated naturally within the HMM framework and hazard analysis approaches such as the Cox^{77,78} and Plackett–Dale⁷⁹ survival models to better understand and ultimately design more efficient intervention policies such as vaccination on targeted parts of the population and project future trends for risk assessment especially for fatal diseases such as AIDS.⁸⁰⁻⁸⁴ Such models are used to quantify the relation of

demographic variables (such as age, gender, social status, spatial characteristics) on the survival rates, i.e., occurrence rates of events such as death or infection in the population.^{85,86}

Mathematical/Mechanistic State-Space Models

According to the level of the approximation of the reality and increasing complexity mathematical models may be categorized in the following categories:

“Continuum” models in the form of differential and/or (integro)-partial differential equations. Continuum models describe the coarse-grained dynamics of the epidemics in the population.⁸⁷⁻⁹⁰ One might, for example, study a model for the evolution of the disease as a function of the age and the time since vaccination^{91,92} or investigate the influence of quarantine or isolation of the infected part of the population.^{93,94} Such models can be explored using powerful analysis techniques for ordinary or partial differential equations. However, due to the complexity and the stochasticity of the phenomena, most available continuum models are often only qualitative caricatures that cannot capture all of the details, therefore compromising epidemiological realism.

Within this context, the population is divided in compartments in accordance to the state of their health, such as susceptible (S), infected (I), and recovered (R). Other states of the population linked with control policies such as vaccinated (V) and quarantined (Q) are also used.

The compartmental SIR mass-action model of Kermack and McKendrick (1922) is the basis of such models. In this representation, it is assumed that an infected individual infects a susceptible with a probability $p_{S \rightarrow I}$ and that an infected individual recovers with a probability $p_{I \rightarrow R}$. The systems dynamics under the mass-balance formulation can be approximated by the following three ordinary differential equations:^{95,96}

$$\frac{dP_i(S)}{dt} = -p_{S \rightarrow I} \sum_{N(S)} P_i(S, I), \quad \frac{dP_i(I)}{dt} = -p_{I \rightarrow R} P_i(I), \quad \frac{dP_i(R)}{dt} = p_{I \rightarrow R} P_i(I)$$

where $P_i(\{S, I, R\})$ denotes the probability that an individual is on one of the states $\{S, I, R\}$ at time t and $P_i(A, B)$ is the pair joint probability to have states A and B communicating at time t ; $N(S)$ denotes the set of links of a susceptible individual. The above equations are not in a closed form. Assuming Markovian behavior of the underlying process, $P_i(S, I) = P_i(S)P_i(I)$. Under the mean field approximation, assuming that the population is perfectly mixed and that every susceptible has the same probability of becoming infected the probabilities are equated to the expected (mean) values of the corresponding variables in the population. These assumptions lead to the following set of equations:

$$\frac{dS}{dt} = -\alpha SI, \quad \frac{dI}{dt} = \alpha SI - \beta I, \quad \frac{dR}{dt} = \beta I$$

where S, I, R denote expected (mean) values; α and $1/\beta$ denote mean values of the disease transmission probability and length of the period for which an individual can transmit the disease before recovering. The above set of equations is the celebrated Kermack and McKendrick model. When a recovered individual

becomes again susceptible after a period of time $1/\gamma$ then the *SIRS* mean field model becomes:

$$\frac{dS}{dt} = -\alpha SI + \gamma R, \quad \frac{dI}{dt} = \alpha SI - \beta I, \quad \frac{dR}{dt} = \beta I - \gamma R$$

In the Kermack and McKendrick model, the disease becomes epidemic, i.e., $\frac{dI}{dt} > 0$, $\frac{dS}{dt} < 0$ if and only if $S > \frac{\alpha}{\beta}$. Hence, the number of infective will increase as long as $S > \frac{\alpha}{\beta}$. At $S = \frac{\alpha}{\beta}$ the number of infected cases reach a maximum and after this it decreases to zero. The threshold $\frac{\beta}{\alpha} S(t=0)$ is called the basic reproduction number (R_0) and indicates whether the disease will become epidemic (if $R_0 > 1$) or it will die out (if $R_0 < 1$).

Generally speaking, R_0 represents the average number of secondary infections produced from a single infected individual introduced into a completely susceptible population. A transmission potential index that relaxes the hypothesis of the fully susceptible population is the effective reproduction number defined as the average number of secondary infections produced from a single infected individual in a population which is already infected from a disease. The parameters of these models can be estimated using epidemic data from past periods. Within this context Coburn et al. (2009) give a review on simulating influenza including swine flu (H1N1) with SIR models.⁹⁷ Nichol et al. (2010) used a SIR model to simulate influenza dynamics in a college campus and through this to assess the impact of various scenarios of vaccinations.⁹⁸ Correia et al. (2011) used a SIR model to study the measles and hepatitis C in Portugal using data from 1996 until 2007.⁹⁹

SIR-type models have also been extended to incorporate demographics such as age distributions, mortality and spatial dependence of the spread to account for diffusion and migration effects as well as genetic mutations in the interacting populations, thus enhancing their realism.

Gaudart et al. (2009) addressed a modified McDonald's SIR model to approximate the dynamics of malaria in the region Bancoumana of Mali that deployed from June 1996 to June 2001.¹⁰⁰ The McDonald's model has been extended to incorporate the state of contagious children as well as the state of susceptible Anopheles and the state of contagious Anopheles. Magal et al. (2010)¹⁰¹ presented an age-dependent infection model with a mass action law, and analyze its stability using a Lyapunov function. Metcalf et al. (2011) developed a metapopulation SIR-based model including the probability of infection by age to predict the rubella dynamics in Peru.¹⁰² Ajelli et al. (2011) developed an SIR-based metapopulation model that incorporates a spatial contact matrix describing the mixing level between Italian regions.¹⁰³ The authors used the model to predict the spatiotemporal dynamics of hepatitis A in the south regions of Italy. The model was fitted using weekly time series of reported rubella cases from 1997 to 2009. The same type of models have been also used to model nosocomial epidemics modeled both at the level of pathogen and

host-host interactions (see, e.g., Webb et al., 2005¹⁰⁴). In Gaudart et al. (2010), the Ross and McKendrick model has been extended to incorporate demographics and genetic changes in the populations to simulate the spread of malaria in Mali and the plague in the Middle Ages.¹⁰⁵ The authors have employed the Archimedean copula approach to relate the risk of infection and biological age. In another study the authors have augmented the model by age classes and with a diffusion term to account for spatial effects in order to approximate the epidemic front wave dynamics of the Black Death between 1348 and 1350.¹⁰⁶ In Demongeot et al. (2012) the Ross and McKendrick SIR model has been revised to incorporate demographic and spatial dynamics introducing continuous age classes and diffusion of both human and vectors species subpopulations within the infected zones.¹⁰⁷ The model has been used to simulated the spread of malaria in Bancoumana, Mali.

Stochastic models including discrete and continuous-time individual based Markov-chain models.¹⁰⁸⁻¹¹⁰ These are usually individual-level models that relax the hypothesis of the mean field approximations of infinite population and perfect mixing introducing the uniqueness of the individual behavior including multiple heterogeneous characteristics. The main representative in the category is the discrete Markov chains (DMC). In DMC both time and states are defined on a discrete set of values. The states of the individuals change at every discrete time step in a probabilistic manner according to simple rules involving their own states and the states of their links satisfying the Markov property, i.e., that that the future values of the states at time $t + \Delta t$ depend only on the values of the states at the previous time step t , i.e.

For example, for a stochastic *SIRS*-like model these transition rules may read:

- Rule #1: An infected individual (I) infects a susceptible (S) link with a probability $p_{S \rightarrow I} = \lambda$ if an active physical communication exists between them.
- Rule #2: An infected individual (I) recovers with a probability $p_{I \rightarrow R} = \delta$.
- Rule #3: A recovered individual (R) becomes susceptible (S) with a probability $p_{R \rightarrow S} = \gamma$. This condition expresses the case of temporal immunity.

When these transition probabilities remain constant in time, the Markov process is then called time homogenous Markov process. The links between individuals form the contact network through which the disease spreads. For simple DMC models this network is assumed to be a fully connected graph resulting to homogeneous mixing of individuals. For this case and in the limit of infinite number of individuals, the stochastic model can be regarded as a mean field deterministic model.

For a uniform distribution with z links per individual and in the limit of an infinite size population the governing equations read:

$$\frac{dS}{dt} = -\lambda z SI + \gamma R, \quad \frac{dI}{dt} = \lambda z SI - \delta I, \quad \frac{dR}{dt} = \delta I - \gamma R$$

However the above deterministic mean field approximations may impose important bias when the assumptions about infinite

size population, homogeneous individuals, homogenous or random regular networks do not hold. Therefore, they may miss important quantitative and/or qualitative information at the coarse-grained/emergent (continuum) level. This situation worsens as the heterogeneity becomes stronger (e.g., interactions on more complex networks with finite size populations).

Within this context, a comparison between stochastic and the analogous deterministic models is given in Allen and Burgin (2000).¹¹¹ Lekone et al. (2006) used a stochastic SEIR model (E stands for exposed to the disease individuals) to simulate the dynamics of Ebola outbreak in the Democratic Republic of Congo in 1995.¹¹² Bishai et al. (2011) used a stochastic SIR model with age structure and two additional states (compartments) to describe heterogeneity in vaccination.¹¹³ The authors combined the epidemic model with an economic model incorporating the costs of the control disease policies to study the cost effectiveness of supplemental immunization activities for measles in Uganda. Wang et al. (2012) developed a stochastic model within the SIR concept to simulate and better understand the multi-periodic patterns in outbreaks of avian flu in North America.¹¹⁴ The model assumes random contact between individuals as well as environmental transmission of the virus.

Non-markovian SIR-like models have been also proposed. These models incorporate “memory” in transmission dynamics. For example, Streftaris and Gibson (2004) propose a non-markovian SIR model for the foot-and-mouth disease outbreaks.¹¹⁵ In their model they assume that individuals remain infected for a time drawn randomly from a two-parameter Weibull distribution. Randomization of classical deterministic SIR-like models, coming from the random, chemical kinetics to account for non-constant population with age classes due to birth and death processes and spatial demographics have been also been proposed.¹¹⁶ Within this context, Allen and Burgin (2000) compare the dynamics of deterministic and their counterparts stochastic epidemics models for populations with constant and variable.¹¹⁷

Complex network models¹¹⁸⁻¹²⁰ that are relaxing the hypothesis of the above stochastic models that the interactions between individuals are instantaneous and homogeneous.¹²¹⁻¹²³ One of the most critical problems in epidemics concerns the dynamic effects of the contact network heterogeneity. Contacts between individuals evolve under numerous complicated and strongly heterogeneous modes that are influenced by a broad spectrum of factors, ranging from the pathogen inherent variability and host–pathogen interaction stochasticity characterizing the transmission mechanisms of a particular disease, to the population-level ones complicated by environmental, seasonal, economic, and demographic conditions. Furthermore, in many situations the spread of an epidemic is shaped by the topology of the contact social network, and, vice versa, the dynamic evolution of the transmission network depends on the emergent dynamics of the epidemic. For example, in a severe epidemic outbreak, a change in the state of endemicity of a particular part of the population can cause a significant change in the characteristics of the transmission network (due to, e.g., link-cutting due hospitalization). Understanding this complex behavior is of utmost importance to public-health measures and policies for controlling diseases

outbreaks. Vaccination, quarantine, and/or use of antiviral drugs on targeted parts of the population have to be carefully designed for the efficient combat of an emerged epidemic. Poor understanding of the infectious disease dynamics as these emerge due to heterogeneous contact interactions may result to serious negative consequences. Over the last years, there has been an intense effort in studying the interplay between the emergent dynamics of infectious diseases and the underlying topology of transmission network.

Within this context, Kuperman and Abramson (2001) showed how changes in the rewiring probability used to construct small-world networks influence the dynamics of a simple epidemic model.¹²⁴ It was shown that there exists a critical value of the rewiring probability that marks the onset of a phase transition from stationary endemic situations to self-sustained oscillations. Hwang et al. (2005) studied the influence of the clustering coefficient and average path length on epidemic outbreaks evolving on scale free networks.¹²⁵ Shirley and Rushton studied the impact of four different types of network topologies, namely Erdős–Rényi, regular lattices, small-world, and scale free on epidemic dynamics.¹²⁶ Reppas et al. (2012) studied the influence of the path length of small world networks on the dynamics of a simple SIRS stochastic epidemic model.¹²⁷ Studies on adaptive networks have only very recently begun to appear in the physics literature¹²⁸ indicating that adaptation can trigger effects that are not present in other types of networks.

Regarding real-world cases, Read et al. (2008) studied the impact of social networking to the spread of a communicable disease by constructing the underlying contact network from a diary-based survey from 49 adults who recorded 8661 encounters with 3528 different individuals over 14 non-consecutive days.¹²³ Christakis and Fowler (2010) studied a flu outbreak at Harvard University in 2009.¹²⁹ Following 744 students they mapped the transmission network following their friends and contacts and detected the critical nodes and links that were responsible for rapid spread and could be used as early warning detectors. In particular, by measuring several statistics of the underlying network topology, they quantified the centrality of individuals in the network, i.e., how much likely is for the disease to pass and transmitted from an individual to other individuals through the network. Salathé et al. (2010) used wireless sensors to obtain close proximity interactions during a typical day at an American high school.¹³⁰ Based on these measures, they constructed the transmission network and studied the potential of the disease to spread in terms of topological characteristics such as transitivity and average-path-length with respect to the duration of contact between students. Keeling et al. (2010) constructed two metapopulation networks based on information available from 2001 on the commuter movements between 10 000 wards in Great Britain.¹²¹ From the cattle trading system they also constructed the movement network between 150 000 farms. They consider four infectious diseases, namely influenza and small-pox in humans and foot-and-mouth disease or tuberculosis in cattle. Comparing simulations with actual data the authors raised the question if simple network models can eventually catch the influence of movements in an epidemic. Furthermore

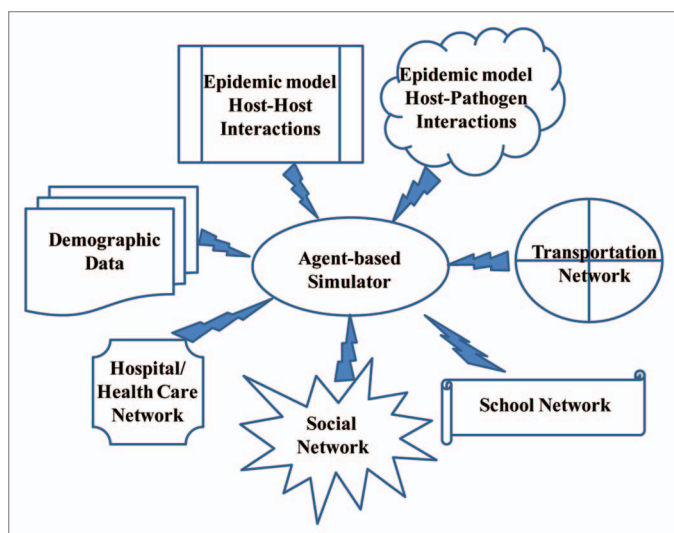


Figure 2. Schematic of the components of an agent-based epidemic simulator.

they showed that the identity of individuals in contrast to random-mover assumption can significantly influence the emergent infection dynamics. Rocha et al. (2011) simulated the spread of sexual transmitted infections using SI and SIR models evolving over the transmission network constructed from data extracted from a Brazilian Internet community where sex buyers rate their encounters with escorts.¹³¹ The network was extended over 12 cities. They showed that due to the high clustering and the distinct communities of the underlying topology, the network slows down outbreaks.

Agent-based simulations. In contemporary mathematical epidemiology, agent-based modeling represents the state-of-the-art for reasoning about and simulating complex epidemic systems. These take into account details such as the transportation infrastructure of the simulated area, the mobility of the population, demographics, and epidemiological aspects such as the evolution of the disease within a host and transmission between hosts (Fig. 2). Public-health epidemiologists, researchers, and policy makers are turning to these detailed models for reasons of ethics, cost, timeliness, and appropriateness. In epidemic systems, testing experimental conditions would put the safety of people at risk, creating an ethical problem. In other cases, real-time evaluation of an existing system may be prohibitively long. For example, in a disaster, simulation can be used to rapidly evaluate many previously unexamined alternatives. In all of these cases, since the real-world system under study is a complex system, multi-agent simulations are used as they are considered to incorporate the appropriate level of complexity. For example the Models of Infectious Disease Agent Study (MIDAS, <https://www.epi-models.org/midas/pubglobamodel.do>), a network launched on May 1, 2004 and funded by the US National Institutes of Health has as its pilot effort the detailed modeling of the dynamics of a hypothetical flu pandemic.

Within this context, Eubank et al. (2004) addressed the use of EpiSims, a detailed agent-based simulator which incorporates data

from population mobility based on TRANSIMS and epidemic models of host-pathogen and host-host interactions.¹³² EpiSims, developed at Los Alamos National Laboratory creates a synthetic population based on the Transportation Analysis and Simulation System (TRANSIMS, <http://code.google.com/p/transims/>). The authors simulated the spread of an infectious disease in the area of Portland, Oregon, US whose network involves 1.5 million people (nodes), 180 000 locations and a total of 1.6 million vertices. Ferguson et al. (2005) developed and presented the simulations results concerning the H5N1 influenza A pandemic in Southeast Asia.¹³³ Their simulations involved 85 million agents residing in Thailand and a 100 km-wide zone of neighboring countries. Demographic data involving details about households, location of schools and workplaces, and population mobility where taken into account. Using the detailed agent-based simulations they evaluated the containment strategies with respect to the potential of preventing a pandemic and the distribution of drugs necessary to eradicate the spread. Burke et al. (2006) presented an agent-based model for the spread of smallpox. The model considered hypothetical towns of 6000- and 50 000 inhabitants.¹³⁴ A distribution of households, workplaces, schools, and hospital units was constructed based on US demographic data. The authors investigated the efficiency of various contagion control scenarios such as vaccination of households, children at schools, isolation of infected persons and vaccination of medical staff in hospitals. Balcan et al. (2009) investigated how short-scale and long-scale contacts due to air travel can influence the spatiotemporal pattern of a pandemic.¹³⁵ The authors made use of the GLEaM agent-based computational platform (<http://www.gleamviz.org/>) consisting of three data layers: the demographic/population, the mobility-related, and the epidemic modeling layer. In this study, real-world data from 29 countries around the globe as well as air travel flowing from 3362 airports indexed by IATA were integrated into a spatial metapopulation epidemic model.

Empirical/Machine Learning-Based Models

Over the last years, machine learning using data extracted from internet-based communication platforms and search engines have been used to extract early indicators of social trends. Microblogging socializing services and web searching platforms have revolutionized the way private and publicly available information diffuses. Such emerging technology appears promising to data mining agents' personal behavior. For example, such services with the aid of search queries have been exploited as tools to stock-market prediction and movie box-office revenue.

Within this context Ginsberg et al. (2009) exploited the aid of search queries on the Google platform for early detection of influenza epidemic in the US.¹³⁶ The authors used around 50 million Google web queries related to influenza symptoms between 2003 and 2008. A linear model using the log-odds of a visit of a physician in a certain region and the log-odds of a related search query submitted from the same region was fitted using publicly available data from the CDC's US Influenza Sentinel Provider Surveillance Network (<http://www.cdc.gov/flu/>). This approach

has now been realized as a surveillance web-based tool (<http://www.google.org/flutrends/>).

Hulth et al. (2009) processed web queries submitted in a Swedish website related to influenza between 2005 and 2007.¹³⁷ The authors fitted two models, one for relating web queries volume with the total number of laboratory verified influenza and the number of persons exhibiting influenza-like symptoms treated by physicians in Sweden. The models were used in turn to estimate outbreaks of the disease in time as well as to predict the influenza evolution. In Chan et al. (2011) a linear model was used to relate Google search queries related to dengue in Bolivia, Brazil, India, Indonesia, and Singapore using publicly available dengue cases between 2003 and 2010.¹³⁸

Conclusion

In this paper, we discussed and presented key modeling methods used for the surveillance and forecasting of infectious disease outbreaks. Generally speaking, epidemiological models can be categorized in three classes: statistical, mathematical-mechanistic state space, and machine-learning based ones. Public-health organizations throughout the world use such models to evaluate and develop intervention disease outbreak policies for ever-emerging epidemics. Simulation allows for rapid assessment and decision making, providing quantification and insight into the spatio-temporal dynamics of a spread. An intensive inter- and multi-disciplinary research effort is speeding up the developments in the field integrating advances from epidemiology, molecular biology, computational engineering and science, and applied mathematics as well as sociology. Nowadays, molecular, sociological, demographic, and epidemiologic data are exploited to develop state-of-the-art detailed very large-scale bottom-up agent-based

models aspiring to approximate the dynamics of real-world cases. Within this context, along with the available information ranging from the host-pathogen interaction level to the host-host, city, country, and globe level, complex network theory has provided the necessary “glue” for the systematic link between epidemiology demographics and sociology.

On one hand, for the bridging of the scales of modeling, one has to first find the appropriate observable variables for which deterministic or stochastic models can be expressed. To this direction, data mining techniques that have flourished over the last few years can be employed to extract such information. On the other hand, due to the complexity of the underlying multi-scale interactions, such models are built on incomplete knowledge imported, e.g., as parameter, rule evolution, and contact network inaccuracies. Thus far, simple brute-force temporal simulations are used to study the behavior of very large scale detailed agent-based simulators in the presence of such inaccuracies. For example some of the rules and model’s parameters, such as the virus pathogenicity—as this may be expressed in terms of the reproduction number—and different social network topologies, are examined in order to assess how such factors may influence the spread of an outbreak. However, such simple simulations are inefficient for the systematic analysis of the emergent epidemic in the parameter space. New rigorous computational methodologies, such as the equation-free multiscale framework,^{96,139-142} that can be used to address this issue have the potential to expedite novel computational modeling and analysis as well as to enhance our understanding and forecasting capability to combat epidemic outbreaks.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Trevisanato SI. The biblical plague of the Philistines now has a name, tularemia. *Med Hypotheses* 2007; 69:1144-6; PMID:17467189; <http://dx.doi.org/10.1016/j.mehy.2007.02.036>
2. Acuna-Soto RD, Stahle DW, Therrell MD, Gomez Chavez S, Cleveland MK. Drought, epidemic disease, and the fall of classic period cultures in Mesoamerica (AD 750-950). Hemorrhagic fevers as a cause of massive population loss. *Med Hypotheses* 2005; 65:405-9; PMID:15922121; <http://dx.doi.org/10.1016/j.mehy.2005.02.025>
3. World Health Organization, UNAIDS, Global Report 2011.
4. World Health Organization. “Avian influenza: assessing the pandemic threat.” January, 2005. WHO/CDS/2005.29
5. Stöhr K, Esveld M. Public health. Will vaccines be available for the next influenza pandemic? *Science* 2004; 306:2195-6; PMID:15618505; <http://dx.doi.org/10.1126/science.1108165>
6. Bernoulli D. Essai d’une nouvelle analyse de la mortalité causée par la petite vérole. *Mém. Math Phys Acad Roy Sci Paris* 1766; 1:1-45.
7. Blower S, Bernoulli D. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. 1766. *Rev Med Virol* 2004; 14:275-88; PMID:15334536; <http://dx.doi.org/10.1002/rmv.443>
8. Dietz K, Heesterbeek JAP. Daniel Bernoulli’s epidemiological model revisited. *Math Biosci* 2002; 180:1-21; PMID:12387913; [http://dx.doi.org/10.1016/S0025-5564\(02\)00122-0](http://dx.doi.org/10.1016/S0025-5564(02)00122-0)
9. Lambert JH. Die Toedlichkeit der Kinderblattern. *Beiträge zum Gebrauche der Mathematik und deren Anwendung. Buchhandlung der Realschule* 1772; 3:568.
10. Laplace PS. *Théorie analytique des probabilités*. Courcier, Paris, 1812.
11. Ross R. *The prevention of malaria*. London: John Murray, 1911; 651-86.
12. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog* 2012; 8:e1002588; PMID:22496640; <http://dx.doi.org/10.1371/journal.ppat.1002588>
13. Kermack WO, McKendrick AG. Contribution to the mathematical theory of epidemics. *Proc R Soc Lond, A Contain Pap Math Phys Character* 1927; 115:700-21; <http://dx.doi.org/10.1098/rspa.1927.0118>
14. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics, part II. *Proc R Soc Lond* 1932; 138:55-83; <http://dx.doi.org/10.1098/rspa.1932.0171>
15. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics, part III. *Proc R Soc Lond* 1933; 141:94-112; <http://dx.doi.org/10.1098/rspa.1933.0106>
16. Guldberg CM, Waage P. Studies Concerning Affinity. C. M. Forhandlinger: Videnskabs-Selskabet i Christiana 1864; 111.
17. En’ko PD. On the course of epidemics of some infectious diseases. *Vrach. St. Petersburg* 1889; 1008-1010. 1039-1042, 1061-1063.
18. En’ko PD. On the course of epidemics of some infectious diseases. *Int J Epidemiol* 1989; 18:749-55; PMID:2695472; <http://dx.doi.org/10.1093/ije/18.4.749>
19. Dietz K. The first epidemic model: a historical note on P.D. En’ko. *Aust J Stat* 1988; 30A:56-65; <http://dx.doi.org/10.1111/j.1467-842X.1988.tb00464.x>
20. Frost WH. Some conceptions of epidemics in general by Wade Hampton Frost. *Am J Epidemiol* 1976; 103:141-51; PMID:766618.
21. McLeod KS. Our sense of Snow: the myth of John Snow in medical geography. *Soc Sci Med* 2000; 50:923-35; PMID:10714917; [http://dx.doi.org/10.1016/S0277-9536\(99\)00345-7](http://dx.doi.org/10.1016/S0277-9536(99)00345-7)
22. Greenwood M. The application of mathematics to epidemiology. *Nature* 1916; xcvi:243-4; <http://dx.doi.org/10.1038/097243a0>
23. Unkel S, Farrington PC, Paul H, Garthwaite PH, Robertson C, Andrew N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. *J R Stat Soc Ser A Stat Soc* 2012; 175:49-82; <http://dx.doi.org/10.1111/j.1467-985X.2011.00714.x>
24. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Rep* 1963; 78:494-506; PMID:19316455; <http://dx.doi.org/10.2307/4591848>

25. Stroup DF, Thacker SB, Herndon JL. Application of multiple time series analysis to the estimation of pneumonia and influenza mortality by age 1962-1983. *Stat Med* 1988; 7:1045-59; PMID:3264612; <http://dx.doi.org/10.1002/sim.4780071006>
26. Costagliola D, Flahault A, Galinec D, Garnerin P, Menares J, Valleron AJ. When is the epidemic warning cut-off point exceeded? *Eur J Epidemiol* 1994; 10:475-6; PMID:7843360; <http://dx.doi.org/10.1007/BF01719680>
27. Greenland S. Regression Methods for Epidemiologic Analysis. *Handbook of Epidemiology*. Part II, 2005, 625-91; http://dx.doi.org/10.1007/978-3-540-26577-1_17
28. Pelat C, Boëlle PY, Cowling BJ, Carrat F, Flahault A, Ansart S, et al. Online detection and quantification of epidemics. *BMC Med Inform Decis Mak* 2007; 7:29; PMID:17937786; <http://dx.doi.org/10.1186/1472-6947-7-29>
29. Dafni UG, Tsiodras S, Panagiotakos D, Gkolfinopoulou K, Kouvatseas G, Tsourti Z, et al. Algorithm for statistical detection of peaks-syndromic surveillance system for the Athens 2004 Olympic Games. *MMWR Morb Mortal Wkly Rep* 2004; 24(Suppl):86-94.
30. Choi K, Thacker SB. An evaluation of influenza mortality surveillance, 1962-1979. I. Time series forecasts of expected pneumonia and influenza deaths. *Am J Epidemiol* 1981; 113:215-26; PMID:6258426.
31. Abeku TA, de Vlas SJ, Borsboom G, Teklehaimanot A, Kebede A, Olana D, 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; PMID:12358620; <http://dx.doi.org/10.1046/j.1365-3156.2002.00924.x>
32. Soebiyanto RP, Adimi F, Kiang RK. Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLoS One* 2010; 5:e9450; PMID:20209164; <http://dx.doi.org/10.1371/journal.pone.0009450>
33. Nunes B, Viboud C, Machado A, Ringholz C, Rebelo-de-Andrade H, Nogueira P, et al. Excess mortality associated with influenza epidemics in Portugal, 1980 to 2004. *PLoS One* 2011; 6:e20661; PMID:21713040; <http://dx.doi.org/10.1371/journal.pone.0020661>
34. Bai Y, Jin Z. Prediction of SARS epidemic by BP neural networks with online prediction strategy. *Chaos Solitons Fractals* 2005; 26:559-69; <http://dx.doi.org/10.1016/j.chaos.2005.01.064>
35. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974; 19:716-23; <http://dx.doi.org/10.1109/TAC.1974.1100705>
36. Page ES. Continuous inspection scheme. *Biometrika* 1954; 41:100-15.
37. Raubertas RF. An analysis of disease surveillance data that uses the geographic locations of the reporting units. *Stat Med* 1989; 8:267-71, discussion 279-81; PMID:2711060; <http://dx.doi.org/10.1002/sim.4780080306>
38. Cowling BJ, Wong IOL, Ho LM, Riley S, Leung GM. Methods for monitoring influenza surveillance data. *Int J Epidemiol* 2006; 35:1314-21; PMID:16926216; <http://dx.doi.org/10.1093/ije/dyl162>
39. Höhle M, Paul M. Count data regression charts for the monitoring of surveillance time series. *Comput Stat Data Anal* 2008; 52:4357-68; <http://dx.doi.org/10.1016/j.csda.2008.02.015>
40. Watkins RE, Eagleson S, Veenendaal B, Wright G, Plant AJ. Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia. *BMC Med Inform Decis Mak* 2008; 8:37; PMID:18700044; <http://dx.doi.org/10.1186/1472-6947-8-37>
41. Spanos A, Theocharis G, Karageorgopoulos DE, Peppas G, Fouskakis D, Falagas ME. Surveillance of community outbreaks of respiratory tract infections based on house-call visits in the metropolitan area of Athens, Greece. *PLoS One* 2012; 7:e40310; PMID:22905091; <http://dx.doi.org/10.1371/journal.pone.0040310>
42. Roberts SW. Control chart tests based on geometric moving averages. *Technometrics* 1959; 1:239-50; <http://dx.doi.org/10.1080/00401706.1959.10489860>
43. Elbert Y, Burkom HS. Development and evaluation of a data-adaptive alerting algorithm for univariate temporal biosurveillance data. *Stat Med* 2009; 28:3226-48; PMID:19725023; <http://dx.doi.org/10.1002/sim.3708>
44. Shiryaev AN. On optimum methods in quickest detection problems. *Theory Probab Appl* 1963; 8:22-46; <http://dx.doi.org/10.1137/1108002>
45. Frisen M, de Mare J. Optimal surveillance. *Biometrika* 1991; 78:271-80.
46. Kulldorff M. Prospective time periodic geographical disease surveillance using a scan statistic. *J Royal Stat Soc A* 2001; 164:61-72; <http://dx.doi.org/10.1111/1467-985X.00186>
47. Bock D, Andersson E, Frisen M. Statistical surveillance of epidemics: peak detection of influenza in Sweden. *Biom J* 2008; 50:71-85; PMID:17849383; <http://dx.doi.org/10.1002/bimj.200610362>
48. Zhou Q, Luo Y, Wang Z. A control chart based on likelihood ratio test for detecting patterned mean and variance shifts. *Comput Stat Data Anal* 2010; 54:1634-45; <http://dx.doi.org/10.1016/j.csda.2010.01.020>
49. Frisen M, Andersson E, Schiöler L. Sufficient reduction in multivariate surveillance. *Comm Statist Theory Methods* 2011; 40:1821-38; <http://dx.doi.org/10.1080/03610921003714162>
50. Naus JJ. Clustering of random points in two dimensions. *Biometrika* 1965; 52:263-7.
51. Kulldorff M, Heffernan R, Hartman J, Assunção RM, Mostashari F. A space-time permutation scan statistic for the early detection of disease outbreaks. *PLoS Med* 2005; 2:e59; PMID:15719066; <http://dx.doi.org/10.1371/journal.pmed.0020059>
52. Le Strat Y, Carrat F. Monitoring epidemiologic surveillance data using hidden Markov models. *Stat Med* 1999; 18:3463-78; PMID:10611619; [http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19991230\)18:24<3463::AID-SIM409>3.0.CO;2-I](http://dx.doi.org/10.1002/(SICI)1097-0258(19991230)18:24<3463::AID-SIM409>3.0.CO;2-I)
53. Rath TM, Carreras M, Sebastiani P. Automated detection of influenza epidemics with hidden Markov models. Berthold MR, Lenz HJ, Bradley E, Kruse R, Borgelt C, editors. Berlin: Springer, 2003.
54. Cooper B, Lipsitch M. The analysis of hospital infection data using hidden Markov models. *Biostatistics* 2004; 5:223-37; PMID:15054027; <http://dx.doi.org/10.1093/biostatistics/5.2.223>
55. Rabiner LR, Juang BH. An introduction to hidden Markov models. *IEEE ASSP Mag* 1986; 3:4-16; <http://dx.doi.org/10.1109/MASSP.1986.1165342>
56. Viterbi AJ. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans Inf Theory* 1967; 13:260-9; <http://dx.doi.org/10.1109/TIT.1967.1054010>
57. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM Algorithm. *J R Stat Soc, B* 1977; 39:1-38.
58. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res* 2005; 14:35-59; PMID:15690999; <http://dx.doi.org/10.1191/0962280205sm388oa>
59. Knorr-Held L, Rasser G. Bayesian detection of clusters and discontinuities in disease maps. *Biometrics* 2000; 56:13-21; PMID:10783772; <http://dx.doi.org/10.1111/j.0006-341X.2000.00013.x>
60. Gangnon RE, Clayton MK. Bayesian detection and modeling of spatial disease clustering. *Biometrics* 2000; 56:922-35; PMID:10985238; <http://dx.doi.org/10.1111/j.0006-341X.2000.00922.x>
61. Denison DGT, Holmes CC. Bayesian partitioning for estimating disease risk. *Biometrics* 2001; 57:143-9; PMID:11252589; <http://dx.doi.org/10.1111/j.0006-341X.2001.00143.x>
62. MacNab YC. Hierarchical Bayesian modeling of spatially correlated health service outcome and utilization rates. *Biometrics* 2003; 59:305-16; PMID:12926715; <http://dx.doi.org/10.1111/1541-0420.00037>
63. Frisen M. Spatial outbreak detection based on inference principles for multivariate surveillance. *Research Report* 2012; 1: ISSN 0349-8034.
64. Robertson C, Nelson TA, MacNab YC, Lawson AB, Lawson AB. Review of methods for space-time disease surveillance. *Spat Spatiotemporal Epidemiol* 2010; 1:105-16; PMID:22749467; <http://dx.doi.org/10.1016/j.sste.2009.12.001>
65. Lawson AB, Kleinman K, eds. *Spatial and Syndromic Surveillance for Public Health*. Chichester: Wiley, 2005.
66. Rencher AC. *Methods of Multivariate Analysis*. John Wiley and Sons, 2002.
67. Kleinschmidt I, Bagayoko M, Clarke GPY, Craig M, Le Sueur D. A spatial statistical approach to malaria mapping. *Int J Epidemiol* 2000; 29:355-61; PMID:10817136; <http://dx.doi.org/10.1093/ije/29.2.355>
68. Carrat F, Valleron AJ. Epidemiologic mapping using the "kriging" method: application to an influenza-like illness epidemic in France. *Am J Epidemiol* 1992; 135:1293-300; PMID:1626546.
69. Cohen AA, Dhingra N, Jotkar RM, Rodriguez PS, Sharma VP, Jha P. The Summary Index of Malaria Surveillance (SIMS): a stable index of malaria within India. *Popul Health Metr* 2010; 8:1; PMID:20181218; <http://dx.doi.org/10.1186/1478-7954-8-1>
70. Coleman M, Coleman M, Mabuza AM, Kok G, Coetzee M, Durrheim DN. Using the SaTScan method to detect local malaria clusters for guiding malaria control programmes. *Malar J* 2009; 8:68; PMID:19374738; <http://dx.doi.org/10.1186/1475-2875-8-68>
71. Kulldorff M. Prospective time-periodic geographical disease surveillance using a scan statistic. *J R Stat Soc Ser A Stat Soc* 2001; A164:61-72; <http://dx.doi.org/10.1111/1467-985X.00186>
72. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Stat Med* 1995; 14:799-810; PMID:7644860; <http://dx.doi.org/10.1002/sim.4780140809>
73. Kulldorff M, Heffernan R, Hartman J, Assunção R, Mostashari F. A space-time permutation scan statistic for disease outbreak detection. *PLoS Med* 2005; 2:e59; PMID:15719066; <http://dx.doi.org/10.1371/journal.pmed.0020059>
74. Gaudart J, Poudiougu B, Dicko A, Ranque S, Toure O, Sagara I, et al. Space-time clustering of childhood malaria at the household level: a dynamic cohort in a Mali village. *BMC Public Health* 2006; 6:286; PMID:17118176; <http://dx.doi.org/10.1186/1471-2458-6-286>
75. Sklar A. Random variables, joint distributions, and copulas. *Kybernetika* 1973; 9:449-60.
76. Nelsen RB. *An Introduction to Copulas*. 2nd edition. NewYork: Springer, 2006.
77. Cox DR. Regression Models and Life Tables (with Discussion). *J R Stat Soc, B* 1972; 34:187-220.
78. Cox DR, Oakes D. *Analysis of Survival Data*. London: Chapman and Hall, 1984.
79. Molenberghs G, Lessafre E. Marginal modeling of correlated ordinal data using a multivariate Plackett distribution. *J Am Stat Assoc* 1994; 89:633-44; <http://dx.doi.org/10.1080/01621459.1994.10476788>
80. Hawkins C, Chalamilla G, Okuma J, Spiegelman D, Hertzmark E, Aris E, et al. Gender differences in antiretroviral treatment outcomes among HIV-infected adults in Dar es Salaam, Tanzania. *AIDS* 2011; 25:1189-97; PMID:21505309; <http://dx.doi.org/10.1097/QAD.0b013e3283471deb>

81. Louzada F, Suzuki AK, Cancho VG, Prince FL, Pereira GA. The long-term bivariate survival FGM copula model: an application to a Brazilian HIV Data. *J Data Sci* 2012; 10:511-35.
82. Hogan DR, Salomon JA, Canning D, Hammit JK, Zaslavsky AM, Bärnighausen T, et al. National HIV prevalence estimates for sub-Saharan Africa: controlling selection bias with Heckman-type selection models. *Sex Transm Infect* 2012; 88(Suppl 2):i17-23; PMID:23172342; <http://dx.doi.org/10.1136/sextrans-2012-050636>
83. Cornell M, Schomaker M, Garone DB, Giddy J, Hoffmann CJ, Lessells R, et al. International Epidemiologic Databases to Evaluate AIDS Southern Africa Collaboration. Gender differences in survival among adult patients starting antiretroviral therapy in South Africa: a multicentre cohort study. *PLoS Med* 2012; 9:e1001304; PMID:22973181; <http://dx.doi.org/10.1371/journal.pmed.1001304>
84. Tibaldi F, Molenberghs G, Burzykowski T, Geys H. Pseudo-likelihood estimation for a marginal multivariate survival model. *Stat Med* 2004; 23:947-63; PMID:15027082; <http://dx.doi.org/10.1002/sim.1664>
85. Kor CT, Cheng KF, Chen YH. A method for analyzing clustered interval-censored data based on Cox's model. *Stat Med* 2013; 32:822-32; PMID:22911905; <http://dx.doi.org/10.1002/sim.5562>
86. El Adlouni S, Beaulieu C, Ouarda TBMJ, Gosselin PL, Saint-Hilaire A. Effects of climate on West Nile Virus transmission risk used for public health decision-making in Quebec. *Int J Health Geogr* 2007; 6:40-7; PMID:17883862; <http://dx.doi.org/10.1186/1476-072X-6-40>
87. Anderson RM, May RM. Population biology of infectious diseases: Part I. *Nature* 1979; 280:361-7; PMID:460412; <http://dx.doi.org/10.1038/280361a0>
88. Brauer F, Castillo-Chavez C. Mathematical models in population biology and epidemiology. New York: Springer-Verlag, 2001.
89. Murray JD. Mathematical Biology II. New York: Springer-Verlag, 2002.
90. Feng Z, Dieckmann U, Levin S, eds. Disease evolution: models, concepts and data analysis. AMS 2006.
91. Keeling MJ. The effects of local spatial structure on epidemiological invasions. *Proc Biol Sci* 1999; 266:859-67; PMID:10343409; <http://dx.doi.org/10.1098/rspb.1999.0716>
92. Greenhalgh D, Das R. Modelling epidemics with variable contact rates. *Theor Popul Biol* 1995; 47:129-79; PMID:7740440; <http://dx.doi.org/10.1006/tpb.1995.1006>
93. Liu WM, Levin SA, Iwasa Y. Influence of non-linear incidence rates upon the behavior of SIRS epidemiological models. *J Math Biol* 1986; 23:187-204; PMID:3958634; <http://dx.doi.org/10.1007/BF00276956>
94. Hethcote HW, Zhihen M, Shengbing L. Effects of quarantine in six endemic models for infectious diseases. *Math Biosci* 2002; 180:141-60; PMID:12387921; [http://dx.doi.org/10.1016/S0025-5564\(02\)00111-6](http://dx.doi.org/10.1016/S0025-5564(02)00111-6)
95. Joo J, Lebowitz JL. Pair approximation of the stochastic susceptible-infected-recovered-susceptible epidemic model on the hypercubic lattice. *Phys Rev E Stat Nonlin Soft Matter Phys* 2004; 70:036114; PMID:15524594; <http://dx.doi.org/10.1103/PhysRevE.70.036114>
96. Reppas A, De Decker Y, Siettos CI. On the Efficiency of the Equation-Free Closure of Statistical Moments: Dynamical properties of a Stochastic Epidemic Model on Erdos-Renyi networks. *J Stat Mech* 2012; 8:P08020; <http://dx.doi.org/10.1088/1742-5468/2012/08/P08020>
97. Coburn BJ, Wagner BG, Blower S. Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC Med* 2009; 7:30; PMID:19545404; <http://dx.doi.org/10.1186/1741-7015-7-30>
98. Nichol KL, Tummers K, Hoyer-Leitzel A, Marsh J, Moynihan M, McKelvey S. Modeling seasonal influenza outbreak in a closed college campus: impact of pre-season vaccination, in-season vaccination and holidays/breaks. *PLoS One* 2010; 5:e9548; PMID:20209058; <http://dx.doi.org/10.1371/journal.pone.0009548>
99. Correia AM, Mena FC, Soares AJ. An Application of the SIR Model to the Evolution of Epidemics in Portugal, Dynamics. *Games and Science II. Springer Proceedings in Mathematics* 2011; 2:247-50; http://dx.doi.org/10.1007/978-3-642-14788-3_19
100. Gaudart J, Touré O, Dessay N, Dicko AL, Ranque S, Forest L, et al. Modelling malaria incidence with environmental dependency in a locality of Sudanese savannah area, Mali. *Malar J* 2009; 8:61; PMID:19361335; <http://dx.doi.org/10.1186/1475-2875-8-61>
101. Magal P, McCluskey CC, Webb GF. Lyapunov functional and global asymptotic stability for an infection-age model. *Appl Anal* 2010; 89:1109-40; <http://dx.doi.org/10.1080/00036810903208122>
102. Metcalf CJE, Munayco CV, Chowell G, Grenfell BT, Bjørnstad ON. Rubella metapopulation dynamics and importance of spatial coupling to the risk of congenital rubella syndrome in Peru. *J R Soc Interface* 2011; 8:369-76; PMID:20659931; <http://dx.doi.org/10.1098/rsif.2010.0320>
103. Ajelli M, Fumanelli L, Manfredi P, Merler S. Spatiotemporal dynamics of viral hepatitis A in Italy. *Theor Popul Biol* 2011; 79:1-11; PMID:20883708; <http://dx.doi.org/10.1016/j.tpb.2010.09.003>
104. Webb GF, D'Agata EM, Magal P, Ruan S. A model of antibiotic-resistant bacterial epidemics in hospitals. *Proc Natl Acad Sci U S A* 2005; 102:13343-8; PMID:16141326; <http://dx.doi.org/10.1073/pnas.0504053102>
105. Gaudart J, Ghassani M, Mints J, Rachdi M, Waku J, Demongeot J. Demographic and spatial factors as causes of an epidemic spread, the copule approach. Application to the retro-prediction of the Black Death epidemic of 1346. *IEEE Advanced Information Networking and Application* 2010: 751-758.
106. Gaudart J, Ghassani M, Mints J, Rachdi M, Waku J, Demongeot J. Demography and diffusion in epidemics: malaria and black death spread. *Acta Biotheor* 2010; 58:277-305; PMID:20706773; <http://dx.doi.org/10.1007/s10441-010-9103-z>
107. Demongeot J, Gaudart J, Mints J, Rachdi M. Demography in epidemics modelling. *Communications Pure & Appl. Analysis* 2012; 11:61-82.
108. Severo N. Generalizations of some stochastic epidemic models. *Math Biosci* 1969; 4:395-402; [http://dx.doi.org/10.1016/0025-5564\(69\)90019-4](http://dx.doi.org/10.1016/0025-5564(69)90019-4)
109. Prosperi MCF, D'Autilia R, Incardona F, De Luca A, Zazzi M, Ulivi G. Stochastic modelling of genotypic drug-resistance for human immunodeficiency virus towards long-term combination therapy optimization. *Bioinformatics* 2009; 25:1040-7; PMID:18977781; <http://dx.doi.org/10.1093/bioinformatics/btn568>
110. Vardavas R, Blower S. The emergence of HIV transmitted resistance in Botswana: "when will the WHO detection threshold be exceeded?" *PLoS One* 2007; 2:e152; PMID:17225857; <http://dx.doi.org/10.1371/journal.pone.0000152>
111. Allen LJ, Burgin AM. Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Math Biosci* 2000; 163:1-33; PMID:10652843; [http://dx.doi.org/10.1016/S0025-5564\(99\)00047-4](http://dx.doi.org/10.1016/S0025-5564(99)00047-4)
112. Lekone PE, Finkenstädt BE. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics* 2006; 62:1170-7; PMID:17156292; <http://dx.doi.org/10.1111/j.1541-0420.2006.00609.x>
113. Bishai D, Johns B, Nair D, Nabyonga-Orem J, Fiona-Makmor B, Simons E, et al. The cost-effectiveness of supplementary immunization activities for measles: a stochastic model for Uganda. *J Infect Dis* 2011; 204(Suppl 1):S107-15; PMID:21666151; <http://dx.doi.org/10.1093/infdis/jir131>
114. Wang RH, Jin Z, Liu QX, van de Koppel J, Alonso D. A simple stochastic model with environmental transmission explains multi-year periodicity in outbreaks of avian flu. *PLoS One* 2012; 7:e28873; PMID:22363397; <http://dx.doi.org/10.1371/journal.pone.0028873>
115. Streftaris G, Gibson GJ. Bayesian inference for stochastic epidemics in closed populations. *Stat Model* 2004; 4:63-75; <http://dx.doi.org/10.1191/1471082X04st0650a>
116. Mints J, Rachdi M, Demongeot J. Stochastic approach in modelling epidemic spread. In: *IEEE AINA' 11 & BLSMC' 11, IEEE Proceedings, Piscataway*, 478-482 (2011).
117. Allen LJS, Burgin AM. Comparison of deterministic and stochastic SIS and SIR models in discrete time. *Math Biosci* 2000; 163:1-33; PMID:10652843; [http://dx.doi.org/10.1016/S0025-5564\(99\)00047-4](http://dx.doi.org/10.1016/S0025-5564(99)00047-4)
118. Newman MEJ. The structure and function of networks. *SIAM Rev* 2003; 45:167-256; <http://dx.doi.org/10.1137/S003614450342480>
119. Barabási AL. Statistical mechanics of complex networks. *Rev Mod Phys* 2002; 74:47-97; <http://dx.doi.org/10.1103/RevModPhys.74.47>
120. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. *Phys Rep* 2006; 424:175-308; <http://dx.doi.org/10.1016/j.physrep.2005.10.009>
121. Keeling MJ, Danon L, Vernon MC, House TA. Individual identity and movement networks for disease metapopulations. *Proc Natl Acad Sci U S A* 2010; 107:8866-70; PMID:20421468; <http://dx.doi.org/10.1073/pnas.1000416107>
122. Keeling MJ, Eames KT. Networks and epidemic models. *J R Soc Interface* 2005; 2:295-307; PMID:16849187; <http://dx.doi.org/10.1098/rsif.2005.0051>
123. Read JM, Eames KTF, Edmunds WJ. Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface* 2008; 5:1001-7; PMID:18319209; <http://dx.doi.org/10.1098/rsif.2008.0013>
124. Kuperman M, Abramson G. Small world effect in an epidemiological model. *Phys Rev Lett* 2001; 86:2909-12; PMID:11290070; <http://dx.doi.org/10.1103/PhysRevLett.86.2909>
125. Hwang DU, Boccaletti S, Moreno Y, López-Ruiz R. Thresholds for epidemic outbreaks in finite scale-free networks. *Math Biosci Eng* 2005; 2:317-27; PMID:20369925; <http://dx.doi.org/10.3934/mbe.2005.2.317>
126. Shirley MDF, Rushton SP. The impacts of network topology on disease spread. *Ecol Complex* 2005; 2:287-99; <http://dx.doi.org/10.1016/j.ecocom.2005.04.005>
127. Reppas AI, Spiliotis KG, Siettos CI. On the effect of the path length of small-world networks on epidemic dynamics. *Virulence* 2012; 3:146-53; PMID:22460641; <http://dx.doi.org/10.4161/viru.19131>
128. Gross T, Sayama H. Adaptive Networks: Theory, Models, and Data. Springer, Heidelberg, 2009.
129. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. *PLoS One* 2010; 5:e12948; PMID:20856792; <http://dx.doi.org/10.1371/journal.pone.0012948>
130. Salathé M, Kazandjeva M, Lee JW, Levis P, Feldman MW, Jones JH. A high-resolution human contact network for infectious disease transmission. *Proc Natl Acad Sci U S A* 2010; 107:22020-5; PMID:21149721; <http://dx.doi.org/10.1073/pnas.1009094108>
131. Rocha LEC, Liljeros F, Holme P. Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLoS Comput Biol* 2011; 7:e1001109; PMID:21445228; <http://dx.doi.org/10.1371/journal.pcbi.1001109>
132. Eubank S, Guclu H, Kumar VS, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. *Nature* 2004; 429:180-4; PMID:15141212; <http://dx.doi.org/10.1038/nature02541>

133. Ferguson NM, Cummings DA, 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; PMID:16079797; <http://dx.doi.org/10.1038/nature04017>
134. Burke DS, Epstein JM, Cummings DA, Parker JI, Cline KC, Singa RM, et al. Individual-based computational modeling of smallpox epidemic control strategies. *Acad Emerg Med* 2006; 13:1142-9; PMID:17085740; <http://dx.doi.org/10.1111/j.1553-2712.2006.tb01638.x>
135. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci U S A* 2009; 106:21484-9; PMID:20018697; <http://dx.doi.org/10.1073/pnas.0906910106>
136. 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; PMID:19020500; <http://dx.doi.org/10.1038/nature07634>
137. Hulth A, Rydevik G, Linde A. Web queries as a source for syndromic surveillance. *PLoS One* 2009; 4:e4378; PMID:19197389; <http://dx.doi.org/10.1371/journal.pone.0004378>
138. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 2011; 5:e1206; PMID:21647308; <http://dx.doi.org/10.1371/journal.pntd.0001206>
139. Kevrekidis IG, Gear CW, Hyman JM, Kevrekidis PG, Runborg O, Theodoropoulos C. Equation-free coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis. *Comm Math Sciences* 2003; 1:715-62.
140. Cisternas J, Gear CW, Levin S, Kevrekidis IG. Equation-free modelling of evolving diseases: coarse-grained computations with individual-based models. *Proc R Soc Lond A* 2004; 460:1-19.
141. Reppas AI, Spiliotis KG, Siettos CI. Epidemionics: from the host-host interactions to the systematic analysis of the emergent macroscopic dynamics of epidemic networks. *Virulence* 2010; 1:338-49; PMID:21178467; <http://dx.doi.org/10.4161/viru.1.4.12196>
142. Siettos CI. Equation-Free multiscale computational analysis of individual-based epidemic dynamics on networks. *Appl Math Comput* 2011; 218:324-36; <http://dx.doi.org/10.1016/j.amc.2011.05.067>