

# UCLA

## UCLA Previously Published Works

### Title

Modeling infectious disease dynamics in the complex landscape of global health

### Permalink

<https://escholarship.org/uc/item/3mt4k1vd>

### Journal

Science, 347(6227)

### ISSN

0036-8075

### Authors

Heesterbeek, Hans  
Anderson, Roy M  
Andreasen, Viggo  
et al.

### Publication Date

2015-03-13

### DOI

10.1126/science.aaa4339

Peer reviewed



Published in final edited form as:

Science. 2015 March 13; 347(6227): aaa4339. doi:10.1126/science.aaa4339.

## Modeling infectious disease dynamics in the complex landscape of global health

Hans Heesterbeek<sup>1,†,\*</sup>, Roy Anderson<sup>2</sup>, Viggo Andreasen<sup>3</sup>, Shweta Bansal<sup>4</sup>, Daniela De Angelis<sup>5</sup>, Chris Dye<sup>6</sup>, Ken Eames<sup>7</sup>, John Edmunds<sup>7</sup>, Simon Frost<sup>8</sup>, Sebastian Funk<sup>4</sup>, Deirdre Hollingsworth<sup>9,10</sup>, Thomas House<sup>11</sup>, Valerie Isham<sup>12</sup>, Petra Klepac<sup>8</sup>, Justin Lessler<sup>13</sup>, James Lloyd-Smith<sup>14</sup>, Jessica Metcalf<sup>15</sup>, Denis Mollison<sup>16</sup>, Lorenzo Pellis<sup>11</sup>, Juliet Pulliam<sup>17</sup>, Mick Roberts<sup>18</sup>, Cecile Viboud<sup>19</sup>, and Isaac Newton Institute IDD Collaboration<sup>‡</sup>

<sup>1</sup>Faculty of Veterinary Medicine, University of Utrecht, Utrecht, Netherlands <sup>2</sup>School of Public Health, Imperial College, London, UK <sup>3</sup>Roskilde University, Roskilde, Denmark <sup>4</sup>Georgetown University, Washington D.C. USA <sup>5</sup>MRC Biostatistics Unit, Cambridge, UK <sup>6</sup>WHO, Geneva, Switzerland <sup>7</sup>Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK <sup>8</sup>University of Cambridge, Cambridge, UK <sup>9</sup>School of Life Sciences, University of Warwick, UK <sup>10</sup>School of Tropical Medicine, University of Liverpool, UK <sup>11</sup>Warwick Mathematics Institute, University of Warwick, Coventry, UK <sup>12</sup>Department of Statistical Science, University College London, London, UK <sup>13</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA <sup>14</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, USA <sup>15</sup>Department of Zoology, University of Oxford, Oxford, UK & Department of Ecology and Evolutionary Biology, Princeton University, Princeton, USA <sup>16</sup>Heriot-Watt University, Edinburgh, UK <sup>17</sup>Department of Biology/Emerging Pathogens Institute, University of Florida, Gainesville, USA <sup>18</sup>Institute of Natural & Mathematical Sciences, Massey University, Auckland, New Zealand <sup>19</sup>DIEPS, Fogarty International Center, NIH, Bethesda, USA

### Abstract

Despite some notable successes in the control of infectious diseases, transmissible pathogens still pose an enormous threat to human and animal health. The ecological and evolutionary dynamics of infections play out on a wide range of interconnected temporal, organizational and spatial scales, which even within a single pathogen often span hours to months, cellular to ecosystem levels, and local to pandemic spread. Some pathogens are directly transmitted between individuals of a single species, while others circulate among multiple hosts, need arthropod vectors, or can survive in environmental reservoirs. Many factors, including increasing antimicrobial resistance, increased human connectivity, and dynamic human behavior, raise prevention and control from formerly national to international issues. In the face of this complexity, mathematical models offer

\*Correspondence to: Hans Heesterbeek, j.a.p.heesterbeek@uu.nl.

†Authors, apart from first and last author, are in alphabetical order.

‡All authors are members of this collaboration, for additional authors and affiliations see the end of the paper (134).

essential tools for synthesizing information to understand epidemiological patterns, and for developing the quantitative evidence base for decision-making in global health.

---

## Introduction

Thirty-five years ago, there was a belief that the health burden of infectious diseases was close to becoming insignificant, as our means of defense and control, including hygiene, nutrition, drugs and vaccines, had brought about a steady decline in overall mortality (1). In recent decades, however, it has become clear that the threat persists in our rapidly changing world, and human mortality attributed to infection is now projected to remain at current levels of 13–15 million deaths annually until at least 2030 (2). Successes in eradicating smallpox and rinderpest have been isolated events in a landscape of endemic and epidemic infections (3). Newly emerging infectious agents represent a continuing challenge, for example HIV in the 20<sup>th</sup> century, more recently SARS- and MERS-coronavirus, West Nile Virus, Nipah virus, drug resistant pathogens, novel influenza-A strains, and a major Ebola outbreak in 2014. Most new infections enter the human population from wildlife or livestock, and the possibilities for emergence and spread in the coming decades are likely to increase due to population growth, increased urbanization and land changes, greater travel, and increased livestock production to meet demands from the world's expanding population (4–8). In our modern world of instant communication the changing behavior of individuals, often in response to publicity about epidemics (as in the case of the SARS epidemic, H5N1 and H7N9 avian influenza, or recent Ebola and MERS outbreaks) or social media attention, is a key issue (9,10). Phylogenetic data shed light on an additional layer of complexity (11), as will an increased understanding of the human genome and how it influences factors such as susceptibility, infectiousness and its duration. At the same time, the development of effective new vaccines remains a difficult challenge, especially for antigenically very variable pathogens (e.g., HIV, falciparum malaria) and for pathogens that stimulate immunity that is only partly protective (e.g., *Mycobacterium tuberculosis*) or temporary (e.g., *Vibrio cholerae*).

In the face of this complexity, computational tools (Box 1) are essential for synthesizing information to understand epidemiological patterns, and for developing and weighing the evidence base for decision-making. Here, we review the contribution of these tools to our understanding of infectious disease dynamics for public health, concentrating both on selected illustrative examples and on current developments. We argue that experts on infectious disease dynamics and experts on prevention and control need to collaborate closely on a global scale, to improve decision making concerning veterinary and human public health as well as agricultural health, both for current infectious disease challenges and for those that lie ahead. Quantitative analysis should be part of all public health policy formulation when addressing the control of infectious diseases.

**Box 1****Quantitative tools in infectious disease dynamics**

Here, we use the words “computational tools” loosely. In infectious disease dynamics, there is a broad range of relevant quantitative tools and we refer to the entire collection. This comprises first of all statistical methods for inference directly from data, including methods to analyze sequencing and other genetic data. This leads to estimates of important epidemiological information such as length of latency, incubation and infectious periods and their statistical distributions, inferred transmission chains and trees early in outbreaks, the risks related to various transmission routes, or estimates of rates of evolution. Mathematical models in the strict sense refer to mathematical descriptions of processes thought to be associated with the dynamics of infection, for example in a population or within an individual. Such models take many forms, depending on the level of biological knowledge of processes involved and data available. Such models also have many different purposes, influencing the level of detailed required.

Several classes of model are used, spanning the spectrum of information available. At one end of the range are detailed individual-based simulation models, where large numbers of distinct individual entities (with their own characteristic traits such as age, spatial location, sex, immune status, risk profile, behavior pattern) are described in interaction with each other, possibly in a contact network, and with the infectious agent. At the other end, there are compartmental models where no individuals are recognized, but only states for individuals (for example: susceptible, infectious, immune) aggregated into compartments where everyone has the same average characteristics, and where interaction is typically random (everybody mixes with everybody else). Such models do not describe the disease history of single individuals, but rather the time evolution of aggregated variables, such as the number of individuals that are currently susceptible. Mathematical models can have both mechanistic parts in their description, based on assumptions about biological mechanisms involved, and more phenomenological parts, where there is a statistical or presumed relation between variables, without clear assumptions from which this relation can be derived. An example of the former is the assumption of mass action to describe interaction between individual hosts; an example of the latter is an empirical relation between the length of an infectious period in a mosquito and environmental temperature.

For infectious disease dynamics, our world is clearly stochastic, in that chance events play a role in many of the processes involved. Certainly at lower levels of biological aggregation, chance dominates, for example in infection of individual cells or in contacts individual hosts make. At higher aggregation levels, many cells or individuals interact, and chance effects may average out to allow deterministic descriptions. There are purely stochastic models, purely deterministic models, and models that are mixed. It is important to point out that, even though the world is stochastic, that stochastic descriptions are not by definition better than deterministic descriptions. Both are still models of reality, and the fact that chance plays a role may have a far less significant influence on model outcome and prediction than choices made in the relations between ingredients and variables.

Areas of rapid growth are statistical and numerical methods and tools to estimate model parameters from, often scarce, mismatched or incomplete, data, and to contrast model output with real world observations.

## Models and public health policy formulation

The value of mathematical models to investigate public health policy questions was recognized at least 250 years ago when, in 1766, Daniel Bernoulli addressed an important public health controversy at the time and published a mathematical analysis of the benefits of smallpox inoculation (then called variolation) (12). Since then, and particularly over the past 50 years, the study of infectious disease dynamics has grown into a rich interdisciplinary field. It is driven both by the desire for fundamental understanding and the need to use that understanding to aid public health decision making. Decision making on vaccination strategies, for example, increasingly depends on model analyses, where infection dynamics must be combined with cost data (Box 2, Influenza: prevention & control). In recent decades, response to major infectious disease outbreaks including HIV, BSE, Foot and Mouth Disease (FMD), SARS, and pandemic and avian influenza, have shown both the need for and capabilities of models (Box 3, HIV: Test & treat strategy). Model-based analysis of such outbreaks also continually pinpoints areas for improvement in methodology and data, emerging from comparison of predictions of models with observed patterns.

### Box 2

#### Influenza - prevention and control

Human influenza – pandemic and seasonal – remains a major issue in public health due to the continued emergence of novel genetic strains, and one where models have successfully addressed questions from basic biology to epidemiology and health policy. In recent years, modeling and other quantitative analysis has been used to study at least three major issues: pandemic preparedness and mitigation strategies (84–89), rethinking vaccination strategies for seasonal influenza (70), and improved methods in phylodynamics and influenza strain evolution (11). Recent models of influenza fitness have also been developed to predict viral evolution from one year to the next, providing a principled and more precise method for the vaccine selection required every year (90).

For seasonal influenza, models have played a key role in providing the scientific evidence base for vaccination policy, making use of the information in multiple, often unavoidably biased, data sources such as syndromic time-series, vaccine coverage and efficacy, economic costs, and contact patterns in the population. For example, a combined epidemic and economic model was fitted to fine-grained data from many sources to describe the dynamics of influenza in the United Kingdom, and the influence of previous vaccination programs (70). With confidence in the model's predictions based on its ability to capture past patterns, it was used to look at alternative vaccination policies and led to a new national policy to vaccinate school-age children (91). Targeting those individuals most likely to spread the virus, rather than only those most likely to

suffer the largest morbidity, is a marked departure from established practice in the UK and is currently under consideration elsewhere (92).

### Box 3

#### HIV – test & treat

Mathematical modeling has played a central role in our understanding of the HIV epidemic, and in informing policy from the outset of our recognition of the pandemic (93). Some of the many insights include a model-based analysis of viral load data from inhibition experiments, which revealed the rapid and on-going turnover of the within-host viral population (94), and the use of phylogenetic models to show that the HIV pandemic did not emerge in the 1980s, but had its roots in the early 20<sup>th</sup> century (95).

A key contribution of mathematical modeling has been to identify when viral transmission occurs over the course of infection, which determines the potential to halt spread by various measures. Models have shown that transmission of HIV depends on the epidemic phase and the sexual behavior of the population, and a large proportion of transmissions may occur late in infection (96). Model-based inference in the Netherlands also suggested that the effective reproduction number (Box 4) had fallen below one due to a combination of low risk behavior and a very effective diagnosis and treatment program (97). The debate was transformed in the mid 2000s, when eradication of HIV through a ‘test and treat’ strategy was hypothesized (98,99). Subsequent trial results showing that antiretroviral treatment (ART) of HIV-positive individuals could practically eliminate transmission within sexual partnerships when the index case is treated (100), have further supported the role of treatment as prevention. Although these findings have not dispelled concerns about transmission early in infection (93), or about extra-couple transmission (101), it is suggested that high population coverage of ART may have reduced the incidence of HIV infection in rural KwaZulu-Natal, South Africa (102).

These findings, combined with the prospect of cheaper, more effective drugs and delivery structures, underpin UNAIDS’ goal of ‘zero new infections’ for HIV and the initiation of a multi-million dollar cluster-randomized trial (103), which will have its outcome assessed against model predictions. In the meantime, the scientific discussion of the effectiveness of ART in preventing transmission continues, sparked by studies that fail to show a decline in incidence after increased treatment (104). Such debates are essential to elucidate areas for improvement of the models used, data needs for such improvement, and to highlight methodological limitations (105). Notably, one can argue that individual-level characteristics, behavior and decisions may be important in determining success. This will demand more refined stochastic models capturing many sources of heterogeneity concomitant with the acquisition of more refined data. As in many areas of science and engineering, a law of diminishing returns in understanding can operate with the representation of increased complexity, but as models increasingly aspire to provide quantitative as well as qualitative predictions, the representation of many heterogeneities may be necessary.

**Box 4****Some fundamental terms and concepts**

- **susceptible** – individual who is at risk of becoming infected if exposed to an infectious agent.
- **basic reproduction number,  $R_0$**  - average number of infections caused by a typical infected individual in a population consisting only of susceptibles; if  $R_0 > 1$ , the infectious agent can start to spread.
- **effective reproduction number,  $R_e$**  – average number of infections caused by a typical infected individual when only part of the population is susceptible; as long as  $R_e > 1$ , the agent can continue to spread.
- **herd immunity** – state of the population where the fraction protected is just sufficient to prevent outbreaks ( $R_e < 1$ ).
- **critical elimination threshold,  $p_c$**  - proportion of the susceptible population that needs to be successfully protected, for example by vaccination, to achieve herd immunity;  $p_c = 1 - 1/R_0$  is a rule of thumb from models when hosts are assumed to mix randomly.
- **force of infection** – per capita rate at which susceptible individuals acquire infection.
- **final size** – fraction of the initial susceptible population that eventually becomes infected during an outbreak.
- **prevalence** – proportion of the population with infection or disease at a given time point.
- **super-spreader/super-shedder** - infected individual that produces substantially more new cases than the average because of greater infectiousness, longer duration of infectiousness, many more transmission opportunities/contacts, or combinations of these. Even when the average  $R_0$  is relatively small, these individuals have large effects on outbreaks.
- **metapopulation** – collection of populations, separated in space, but connected through movement of individuals.
- **critical community size** – minimum number of individuals in a population that allow an infectious agent to persist without importation of cases.
- **case fatality ratio** – proportion of symptomatic infections that result in death.
- **SIR model** – most basic model metaphor for immunizing infections where each living individual is assumed to be in one of three epidemiological states at any given time: **S**usceptible, **I**nfected & infectious, and **R**ecovered & immune. The model specifies the rates at which individuals change their state. Individuals progress from **S** to **I** when infected, then from **I** to **R** upon recovery. Many

variants exist, for example recognizing different classes of S, I and R individuals, depending on individual traits such as age.

For infectious agents important to public health, ‘case law’ has built up, for example regarding important factors for infection dynamics, concepts to aid understanding and communication of quantitative and qualitative insight, and regarding computational tools and modeling frameworks (Table 1 and Box 4). The basic reproduction number  $R_0$  is a central concept characterizing the average number of secondary cases generated by one primary case in a susceptible population. Such concepts serve to highlight what must be measured to interpret observed patterns and to quantify the impact of control strategies in a cycle of interaction between policy and modeling (Fig. 1).

Two fundamental properties of the world that shape infectious disease dynamics make computational tools key for understanding reality. The world is essentially a stochastic and highly non-linear system. The non-linearity not only derives from the complex interaction between factors involved in transmission are complicated, but also from the influence that the infection process has on the distribution of important characteristics at various temporal and spatial scales. For example, the age-related nature of infection and mortality in HIV changes the age distribution of the population, and previous exposure to strains of influenza alters the distribution of influenza susceptibility. Such feedback mechanisms contribute to the high nonlinearity of infection processes. Non-linearity leads to counterintuitive phenomena (Fig. 2), and the inability simply to extrapolate experience from one situation to another (for example when deciding whether to implement a vaccination policy in different countries, Fig. 1). Mathematical tools, relating to data and processes on a large range of interacting scales, have become essential to explore, anticipate, understand and predict the effects of feedbacks within such complex systems, including changes caused by intervention.

## Current and future opportunities for models in public health

Over the last decade, key public health questions, ranging from emergence to elimination, have posed a range of challenges for modeling infectious disease dynamics, many of which rest on leveraging novel data sources, and integrating data from a range of scales - from sequence data to global circulation. Given commonalities in processes across pathogens, progress made in one area can lead to advances in another, and progress in the areas described above all build on and inform each other, making this a dynamic time for research in the discipline (13). A few themes are chosen to illustrate current trends in model development and public health application.

### The difficulties of real-time outbreak modeling: the Ebola 2014 outbreak

The ongoing (at the time of writing) outbreak of Ebola in West Africa serves to highlight both opportunities and challenges in modeling for public health. In the initial phase of an outbreak, real-time estimates of the reproduction number or simple exponential extrapolation (14) allow short-term predictions of epidemic growth that can be used, for example, to plan for necessary bed capacity. Also in an early phase, phylogenetic



quantitative tools can use samples from initial victims to provide important estimates of outbreak origin (15). Later in an outbreak, these methods are replaced by more detailed mechanistic models that explicitly take into account depletion of susceptibles, as well as specifics such as different transmission routes or settings, to provide more accurate and fine-grained predictions of the impact of public health measures. For Ebola, such methods make it possible to distinguish between transmission at funerals, in health-care settings and in the community (16,17) to predict the effect of increasing bed capacity or reducing funeral transmission. Determining which combinations of measures are most effective in bringing down the reproduction number call for close and fast interaction between modelers and policy makers (18). It is likely that disease-dynamic models will also be at the forefront of a debate on the optimal deployment of initially scarce Ebola vaccines, once such vaccines become available.

With the opportunities of real-time modeling for public health come specific challenges. The imperative to produce reliable and meaningful analysis for those at the forefront of battling the outbreak has to be balanced against the pressures and delays of scientific publication. In an ongoing outbreak, data can be patchy and reporting delayed, and different data sources are not always synthesized. Particularly, data are lacking on the effect on transmission dynamics of the various control measures operating simultaneously in the hectic circumstances of the most severely hit areas. Also, under-reporting is a critical challenge for ongoing assessment of the epidemic and has enormous impact on predictions of outbreak size, but also of outbreak impact, for example in terms of the case fatality ratio. This estimate of severity suffers early in an outbreak from imprecise information on both the numerator (deaths confirmed as being caused by the infection) and the denominator (under-reporting and not all suspected cases are laboratory-confirmed infections). This caused problems early in the H1N1 influenza outbreak starting in Mexico in 2009, as well as in the current Ebola outbreak. While levels of under-reporting can be estimated after the outbreak from retrospective serological studies, it is usually not identifiable in real-time data.

These limitations make it almost impossible to make reliable long-term predictions. Modeling results are therefore often based on scenarios in which a pathogen spreads unaltered by behavioral changes or the public health response. In reality this is rarely the case, especially in a devastating outbreak such as the one of Ebola in West Africa, where the situation constantly changes due to growing awareness in the community, as well as national and international intervention. Careful communication of findings is key, and data and methods of analysis (including code) must be made freely available to the wider research community. Only in this way can reproducibility of analyses and an open exchange of methods and results be ensured for maximal transparency and, consequently, benefit to public health.

### **Emergence of novel human pathogens**

There is an ever-present hazard that novel human pathogens emerge from livestock and wild mammal and bird reservoirs. Research on potential emerging zoonoses draws on concepts from across the spectrum of infectious disease dynamics, disease ecology, microbiology and phylogenetic analysis. Particular challenges include estimating human-to-human

transmissibility against a backdrop of on-going zoonotic spillover, detecting anomalous outbreaks, and assessing the risk that more dangerous strains may arise through pathogen evolution.

The recently identified gap in methodology for zoonoses with weak human-to-human transmission (6) is being filled, with new approaches to estimating  $R_0$  and other transmission-related quantities from subcritical outbreak data (19–21). These studies address key public health concerns, but rely on strong assumptions regarding the quality and completeness of case observations. Better information on surveillance program efficacy could be gained through serological surveys (where blood and saliva samples reveal evidence of past and present infections) or sociological study, and modeling studies can help to design and characterize efficient surveillance programs (22). Given the predominance of zoonotic pathogens among emerging infections, models for transmission dynamics and evolution in multispecies ecosystems and food webs, consisting of host species and non-host species interacting ecologically and epidemiologically, are a crucial area for future development (6,23). The greatest challenge – and the greatest prize – in modeling emerging zoonoses is to assess which diseases pose the most risk to humans and how these might change over time and in different localities (24). Such tasks, which will join molecular studies to experimental infections to epidemiological and ecological surveys, will drive empirical and theoretical efforts for decades to come.

The rising availability of pathogen genome sequence data, coupled with new computational methods, presents opportunities to identify with precision ‘who infects whom’ and the networks of infection between humans and reservoirs (25). Full realization of this potential, though, will require denser and more systematic whole genome sampling of pathogens coupled with associated epidemiological data, as well as baseline information on genetic diversity and evolutionary rates, especially in animal hosts (26).

### **Pathogen evolution and phylodynamics**

As pathogen genetic data become increasingly available, modelers are finding ways to synthesize these new data streams with more traditional epidemiological information in phylodynamic tools (27,28). However, current frameworks employ compartmental epidemiological models, which do not make efficient use of individual-level epidemiological data. Although sampling theory is well developed for standard surveillance data, the relationship between a set of pathogen sequences and the phylogeny inferred from a population sample is more complex (11). Many-to-one mapping possibilities between, on the one hand, combinations of epidemiological, immunological and evolutionary processes shaping sequences and, on the other hand the inferred phylogeny, demand the integration of diverse data sources and an increased focus on systematic sampling.

Phylodynamic studies to date have largely focused on fast-evolving RNA viruses, driven by the large amount of data generated for clinical (e.g. Hepatitis C Virus (HCV), HIV) or surveillance (e.g. influenza) purposes (11). Replicating these efforts on an expanded array of pathogens, including DNA viruses, bacteria, fungi, protozoa (e.g. malaria) and helminths is a promising avenue for future research (29). It is of particular importance in the context of the evolution and spread of drug resistant variants and escape mutants from vaccine

protection. However, genome-wide pathogen data also present challenges, in particular in relation to accommodating recombination, re-assortment, and mobile genetic elements. Analysis of bacterial genomes usually considers only those genes that are shared across taxa, but there are good reasons to believe that non-core genes play an important role in bacterial evolution, including the evolution of antibiotic resistance (30).

While sequence data are extremely valuable, in order to link these data fully to disease dynamics a key advance will be determining how sequence changes may affect functions related to pathogen fitness, such as replication rate, transmissibility, and immune recognition (genotype to phenotype). Molecular epidemiological studies often treat pathogen genetic variation as simply reflecting the underlying transmission process, whereas in reality such variation may play an important role in determining transmission dynamics, as exemplified by the escape from herd immunity by influenza A virus (31).

'Deep' sequencing of pathogens within individual hosts generates information on within-host diversity, resulting from evolution within the host (often in response to drug treatment), or multiple infections. To tackle within-host diversity, models that embed pathogen evolution within a transmission tree are needed. Such models, which cross the within- and between-host scales, are only just becoming analytically and computationally feasible despite being proposed several years ago (32). Similarly, while progress has been made in scaling inference from genes to genomes (33), scaling inference to large numbers of sequences is lagging far behind.

### Multiple infections

Infectious disease epidemiology evolved by focusing on interactions between a single host species and a single infectious agent. It is becoming increasingly clear that multiple agents simultaneously infecting the same host populations and individuals significantly add to public health burden and complicate prevention and control. Co-infections in relation to HIV, for example tuberculosis and HCV, or co-infection of different strains of influenza A virus raise important public health and evolutionary issues. Multiple agents infecting the same host individual have been shown to influence each other by increasing or decreasing susceptibility and/or infectivity of that individual, thereby influencing the population dynamics of these agents in ways that we have yet to explore and understand (34,35).

Multiple infections of the same individual with closely related pathogens occur when infection elicits no, or only a partial, immune response. Macroparasites including many of the important human helminth infections are good examples of pathogens that evade human immune responses permitting repeated infection of the same host (36). Biological mechanisms giving rise to such multiple infections include sequential reinfections caused by antigenic drift in influenza, antigenic variation in respiratory syncytial virus (RSV) and waning (slow loss of) immunity in pertussis while lack of cross-protection in many colonizing microparasites, for example pneumococcus and human papilloma virus (HPV), allows for multiple concurrent infections. Although the existence of reinfections is a clinical fact, population-level data are scarce as reinfections are often subclinical and individual-based longitudinal infection-histories only anecdotal. Results from new analytic approaches

relating to deep sequencing and neutralization tests covering multiple antigens are being utilized (37).

The immunodynamics of influenza have clear policy implications for the identification of high-risk groups in connection with pandemic planning (38) while the dynamics of waning immunity are key to the current concerns about immunization level for pertussis (39). Multivalent vaccines covering only a targeted subset from the circulating strains of pneumococcus and HPV pose important new applied problems (40). The spread of recombinant viruses implies the existence of multiple infections. An example is the Sydney 2012 strain of norovirus, but how this can occur in an acute infection remains to be understood, as the time window for multiple exposure is limited. Mathematical models may help to explore how, for example, a subpopulation with chronic infection or a hypothesized environmental reservoir may contribute to the dynamics of multiple infections.

### **Behavior of hosts**

Human behavior is a fundamental determinant of infectious disease dynamics, whether by affecting how people come in contact with each other, vaccination coverage, reporting biases, or adherence to treatment. Traditional epidemic models have tended to ignore heterogeneity in contact behavior (although early HIV models addressed heterogeneity in sexual behavior by necessity (41)), however increasing sophistication of contact network models (42), together with data on epidemiological contacts, creates opportunities for understanding and controlling transmission at a fundamental level (43), and opens up the possibility of independent study of relevant social factors (10). Recent years have seen exciting developments in the measurement of contact patterns and ‘who might infect whom’ through advances in individual electronic identification technology. This is a promising avenue for linking pathogen genetic data and human behavior.

Contact patterns are not static, and can shift during outbreaks as individuals change their behavior in response to perceived risk and public-health interventions (44). Modeling has illuminated this process, for example incorporating peer influence on vaccination behavior into models of infectious disease dynamics (45,46). Analysis of data from online social networks has also created promising opportunities to validate such approaches with empirical observations (47,48).

Movement and travel are tightly linked to the spread of infection, and have been explored through models to highlight commuting and agricultural migration driving local disease transmission (49), and global disease patterns through air travel (50). These processes are now being investigated to gain insights into the more complex case of vector-borne diseases such as malaria and dengue, where both host and vector movement can interact to drive local (51) and large-scale dynamics (52).

### **Elimination/eradication**

Modeling has long provided support for elimination efforts – vector control (53), critical community size (54), herd immunity and critical vaccination threshold (55,56) were all powerful insights from models framed in relatively simple and homogenous terms. Subtleties and complexities in many current eradication programs, as well as the availability

of novel data sources, have called for a range of extensions in the theory. As we approach elimination targets, disease dynamics have changed in ways that were largely predicted by models, but also in unanticipated ways due to ignorance about key epidemiological processes.

Incentives for control efforts also change, both at the individual level (passive or active refusal can develop (57)) and at the country level (58). This reinforces the call for development of models of human behavior and its interaction with infectious disease dynamics (9), potentially drawing on new data sources (e.g. from social media (59,60)), as well as for models that can capture national and non-governmental motivations, interactions and competition, economical or otherwise. Long-term control puts pathogens under strong selection for resistance, calling for evolution-proof control methods (61) and novel vaccine technologies and their optimized delivery (62).

Finally, since the era of smallpox eradication, patterns of global disease circulation have changed radically. Human mobility and migration are increasing global connectivity, strengthening the need for cooperation and international synchronization of efforts (as illustrated by polio). Techniques for analysis of novel data-sources are again key here, e.g., Call Data Records provide unique opportunities to understand disease source-sink dynamics (52).

### **Computational statistics, model fitting and big data**

By definition and design, models are not reality. The properties of stochasticity and non-linearity strongly influence the accuracy of absolute predictions over long time horizons. Even if the mechanisms involved are broadly understood and relevant data are available, predicting the exact future course of an outbreak is impossible due to changes in conditions in response to the outbreak itself, and due to the many chance effects in play. These stochastic effects dominate developments in situations with relatively few infected individuals (such as at emergence, approaching the threshold for sustained host-host spread, or approaching elimination/eradication). This makes it virtually impossible to predict which infectious disease agent is going to emerge/evolve next and where, or to predict when and where the next or last case in an outbreak will occur. There is, typically in complex systems, a fundamental horizon beyond which accurate prediction is impossible. The field has yet to explore where that horizon is and whether computational tools and additional data (and if so which data) can stretch predictions to this limit. In contrast, ‘what-if’ scenarios for public health intervention can provide qualitative (and increasingly semi-quantitative) insight into their population consequences.

With growing applications in public health there is an increasing demand to validate models by making model predictions consistent with observed data. The development of evermore-powerful computers is accompanied by new techniques utilizing this power, notably for statistically rigorous parameter estimation and model comparison. Techniques such as Markov Chain Monte Carlo (MCMC) have become firmly established tools for parameter estimation from data in infectious disease models (e.g. 63), and Monte Carlo based methods will play a pivotal role in addressing the challenges that lie in reconciling predictions and observations (64,65). Other techniques, such as so-called particle filters, approximate

Bayesian computation, emulation, and their combinations with MCMC (e.g. 66), are rapidly developing and allow matching stochastic models that explicitly account for incomplete observations to time series of cases, giving insights into scenarios as diverse as cholera in Bangladesh (67) and influenza (68,69). The need to integrate multiple data sources (70,71) as well as to include uncertainty in model parameters and/or structure has led to an increase in the popularity of Bayesian approaches.

The rapid expansion of infectious disease models and their application over the past decade has coincided with an increase in open access datasets available from a variety of sources, but progress in data capture needs to be accelerated. While some of these technologically advanced data streams have been incorporated into models, for example to track the incidence of influenza in the USA (72), to elucidate the spatial dynamics of measles and malaria in Africa (53,73) and to chart the spread of dengue globally (74), much more remains to be done to leverage data collected from different sources (e.g. demographic, genetic, epidemiological, treatment and travel patterns) and at different temporal and spatial scales.

## Concluding remarks

Infectious diseases are an important frontier in public health, and their prevention and control call for global, rather than national or regional, coordinated efforts (75–78). The success of smallpox and rinderpest eradication campaigns shows the possibilities; the global spread of newly emerged pathogens (recently avian influenza strains and MERS coronavirus), the difficulties in curbing the spread of antibiotic resistance, the upsurge of polio towards the ‘end-phase’ of its eradication, and the recent unprecedented spread of Ebola virus, are examples that show the need for international coordination and collaboration. Non-linearity in infectious disease dynamics and global connectivity cause sub-optimal national decisions on control and prevention to have regional and even global repercussions.

Given the mismatch with regions where most expertise on infectious disease dynamics is concentrated, it is important to empower local scientists and policy makers in regions where the burden of disease is heaviest about problems facing their own countries and the consequences of local actions. It is essential to make expertise, data, models, statistical methods and software widely available by open access. There are several initiatives (e.g. Thiswormyworld.org, Garkiproject.nd.edu, EDENextdata.com, and the Malaria Atlas Project), but much more needs to be done. Modeling tools and software for data analysis are beginning to become open source, such that findings can be replicated, additional scenarios can be evaluated, and others can incorporate methods for data analysis or simulation. Ultimately, sharing models guarantees more reproducible results, while maximizing model transparency.

Making data sets widely available is also crucial, for example to support replication of findings and broader comparative analyses (79). As models become open access, so should much of the data collected by governments, international agencies and epidemiology research groups. Two outbreaks never occur in exactly matching circumstances, even for the

same infectious agent, so there is potential to study many outbreaks in parallel to gain insight into the determinants of outbreak pattern and severity. Looking forward, there is a major opportunity to design experiments, clinical trials (for example for vaccines (80)), and surveillance protocols to test model predictions or assumptions, and to help reduce or better target the enormous costs involved. By integrating modeling approaches throughout the full life cycle of infectious disease policies, including economic considerations (58, 70, 81), health outcomes can be improved and scientific understanding can be advanced.

At present, the evidence provided by infectious disease models is not considered on the GRADE scale ([www.gradeworkinggroup.org](http://www.gradeworkinggroup.org)) alongside that of conventional studies such as clinical trials. Regardless, models are essential when diverse sources of data (including GRADE-scale evidence) need to be combined and weighed. In many cases, the definitive trial cannot be performed and models are needed to catalyze insight and extract maximum value from data that *are* available. In recent years, uniformity of practice and quality control for models has received more attention, resulting in initial attempts to characterize Good Modelling Practice for infectious diseases (82,83).

The optimal use of models to inform policy decisions requires a continuous dialogue between the multidisciplinary infectious disease dynamics community and decision makers. This is increasingly understood by governments in developed countries, in non-governmental agencies and by large funding bodies such as the Bill and Melinda Gates Foundation. This dialogue will help to reduce the burden from infectious diseases by providing better-informed control strategies. Mathematical models will allow us to capitalize on new data streams, and lead to an ever-greater ability to generate robust insight and collectively shape successful local and global public health policy.

## Acknowledgments

HH conceived and wrote the paper; RA, VA, SB, DDA, CD, KE, JE, SFr, SFu DH, TH, VI, PK, JL JLS, JM, DM, JP, LP, MR and CV provided text and edited the manuscript; JL, JM and DH produced the figures; the Isaac Newton Institute IDD Collaboration jointly produced and discussed ideas for the outline and content (all of the above plus NA, FB, TB, JG, BG, AL, AM, PON, CP, SR, GST, PT, JW). This paper was conceived and developed at a program on Infectious Disease Dynamics at the Isaac Newton Institute, Cambridge, UK, August 19 – September 13, 2013, and May 19 – June 6, 2014 ([www.newton.ac.uk](http://www.newton.ac.uk)). We gratefully acknowledge financial and infrastructural support from the Isaac Newton Institute, fundamental to the success of this program. We are also grateful for the financial support the program received from the Research and Policy for Infectious Disease Dynamics (RAPIDD) program of the Science and Technology Directorate, U.S. Department of Homeland Security, and the Fogarty International Center, NIH.

## References and Notes

1. World Health Organization. Global strategy for health for all by the year 2000. WHO; Geneva: 1981. <http://whqlibdoc.who.int/publications/9241800038.pdf>
2. World Health Organization. Mortality and global health estimates 2013. WHO; 2013. <http://apps.who.int/gho/data/node.main.686?lang=en>
3. Klepac P, Metcalf JE, Hampson K. Towards the endgame and beyond: complexities and challenges for the elimination of infectious diseases. *Phil Trans R Soc B*. 2013; 386(1623)
4. Daszak P, Cunningham AA, Hyatt AD. Emerging infectious diseases of wildlife--threats to biodiversity and human health. *Science*. 2000; 287:443–449. [PubMed: 10642539]
5. Jones KE, et al. Global trends in emerging infectious diseases. *Nature*. 2008; 451:990–993. [PubMed: 18288193]

6. Lloyd-Smith JO, et al. Epidemic dynamics at the human-animal interface. *Science*. 2009; 326:1362–1367. [PubMed: 19965751]
7. Keesing F, et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature*. 2010; 468:647–652. [PubMed: 21124449]
8. Smith KF, et al. Global rise in human infectious disease outbreaks. *J R Soc Interface*. 2014; 11:20140950. [PubMed: 25401184]
9. Funk S, Salathé M, Jansen VAA. Modeling the influence of human behaviour on the spread of infectious diseases: a review. *J R Soc Interface*. 2010; 7:1247–56. [PubMed: 20504800]
10. Bauch CT, Galvani AP. Social factors in epidemiology. *Science*. 2013; 342:47–49. [PubMed: 24092718]
11. Volz EM, Koelle K, Bedford T. Viral Phylodynamics. *PLoS Comput Biol*. 2013; 9:e1002947. [PubMed: 23555203]
12. Bernoulli D. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Rev Med Virol*. 2004; 14:275–288. reprinted translation by L. Bradley with review by S. Blower; original in French 1766. [PubMed: 15334536]
13. Lloyd-Smith JO, Mollison D, Metcalf JE, Klepac P, Heesterbeek JAP. Challenges in Modelling Infectious Disease Dynamics. *Epidemics*. Mar.2015 10(special issue)
14. 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; 371:1481–1495. [PubMed: 25244186]
15. Gire SK, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014; 345:1369–1372. [PubMed: 25214632]
16. Camacho A, et al. Potential for large outbreaks of Ebola viral disease. *Epidemics*. 2014; 9:70–78. [PubMed: 25480136]
17. Pandey A, et al. Strategies for containing Ebola in West Africa. *Science*. 2014; 346:991–995. [PubMed: 25414312]
18. Whitty CJM, et al. Tough choices to reduce Ebola transmission. *Nature*. 2014; 515:192–194. [PubMed: 25391946]
19. Blumberg S, Lloyd-Smith JO. Inference of  $R_0$  and transmission heterogeneity from the size distribution of stuttering chains. *PLoS Comp Biol*. 2013; 9:e1002993.
20. Cauchemez S, et al. Using routine surveillance data to estimate the epidemic potential of emerging zoonoses: application to the emergence of US swine origin influenza A H3N2v virus. *PLoS Med*. 2013; 10:e1001399. [PubMed: 23472057]
21. Kucharski A, et al. Distinguishing Between Reservoir Exposure and Human-to-Human Transmission for Emerging Pathogens Using Case Onset Data. *PLoS Currents Outbreaks*. 2014; 1
22. Restif O, et al. Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecol Letters*. 2012; 15:1083.
23. Selakovic S, De Ruiter P, Heesterbeek JAP. Infectious disease agents mediate interaction in food webs and ecosystems. *Proc R Soc B*. 2014; 281:20132709.
24. Plowright RK, et al. Ecological dynamics of emerging bat virus spillover. *Proc R Soc B*. 2015; 282:20142124.
25. Cotten M, et al. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *MBio*. 2014; 5:e01062. [PubMed: 24549846]
26. Pepin KM, et al. Identifying genetic markers of adaptation for surveillance of viral host jumps. *Nature Rev Microbiol*. 2010; 8:802. [PubMed: 20938453]
27. Grenfell BT, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*. 2004; 303:327–332. [PubMed: 14726583]
28. Rambaut A, et al. The genomic and epidemiological dynamics of human influenza A virus. *Nature*. 2008; 453:615–619. [PubMed: 18418375]
29. Gardi JL, et al. Whole-genome sequencing and social network analysis of a tuberculosis outbreak. *N Engl J Med*. 2011; 364:730–739. [PubMed: 21345102]
30. Stokes H, Gillings M. Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiology Reviews*. 2011; 35:790–819. [PubMed: 21517914]



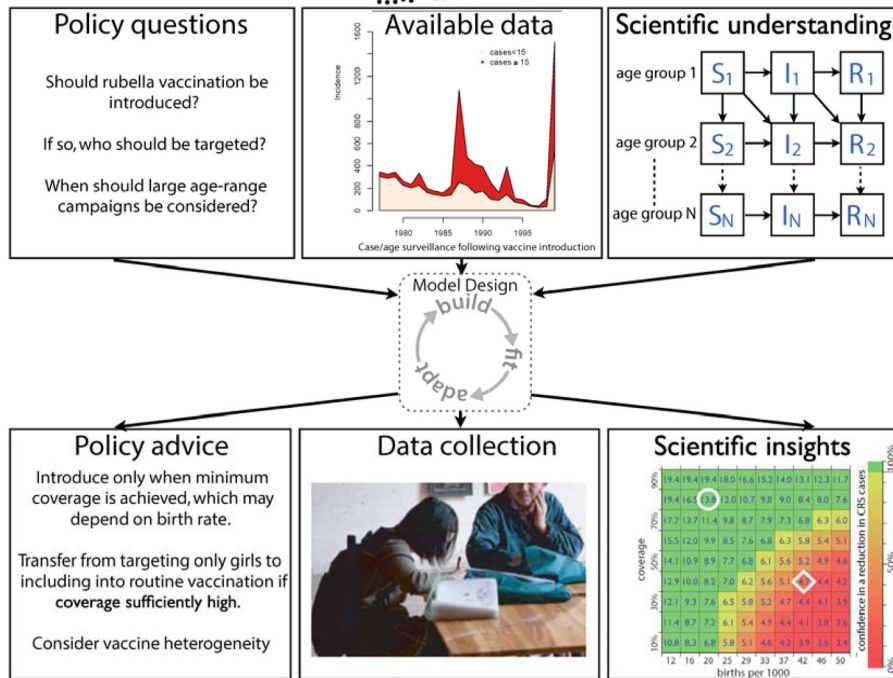
31. Koel B, et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science*. 2013; 342:976–979. [PubMed: 24264991]
32. Vrancken B, et al. The Genealogical Population Dynamics of HIV-1 in a Large Transmission Chain: Bridging within and among Host Evolutionary Rates. *PLoS Comp Biol*. 2014; 10:e1003505.
33. Ayres D, et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Systematic Biology*. 2012; 61:170–173. [PubMed: 21963610]
34. Telford S, et al. Species interaction in a parasite community drive infection risk in a wildlife population. *Science*. 2010; 330:243–246. [PubMed: 20929776]
35. Cattadori I, et al. Infections do not predict shedding in co-infections with two helminths from a natural system. *Ecology*. 2013; 95:1684–1692. [PubMed: 25039232]
36. Anderson RM, Truscott J, Hollingsworth TD. The coverage and frequency of mass drug administration required to eliminate persistent transmission of soil-transmitted helminths. *Phil Trans R Soc B*. 2014; 369:20130435. [PubMed: 24821921]
37. Lessler J, et al. Evidence for antigenic seniority in influenza A (H3N2) antibody responses in Southern China. *PLoS Pathogens*. 2012; 8:e1002802. [PubMed: 22829765]
38. Miller MA, Viboud C, Balinska M, Simonsen L. The signature features of influenza pandemics - Implications for policy. *N Engl J Med*. 2009; 360:2595–2598. [PubMed: 19423872]
39. Lavine JS, King AA, Bjornstad ON. Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure. *Proc Natl Acad Sci USA*. 2011; 108:7259–7264. [PubMed: 21422281]
40. Cobey S, Lipsitch M. Niche and Neutral Effects of Acquired Immunity Permit Coexistence of Pneumococcal Serotypes. *Science*. 2012; 335:1376–1380. [PubMed: 22383809]
41. Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS*. 1989; 3:807–817. [PubMed: 2517202]
42. Danon L, et al. Networks and the epidemiology of infectious disease. *Interdiscip Perspect Infect Dis*. 2011:284909. [PubMed: 21437001]
43. Bansal S, Grenfell BT, Meyers LA. When individual behaviour matters: homogeneous and network models in epidemiology. *J R Soc Interface*. 2007; 4:879–891. [PubMed: 17640863]
44. Bansal S, Read J, Pourbohloul B, Meyers LA. The dynamic nature of contact networks in infectious disease epidemiology. *J Biol Dyn*. 2010; 4:478–89. [PubMed: 22877143]
45. Funk S, Gilad E, Watkins C, Jansen VAA. The spread of awareness and its impact on epidemic outbreaks. *Proc Natl Acad Sci USA*. 2009; 106:6872–6877. [PubMed: 19332788]
46. Oraby T, Thampi V, Bauch CT. The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases. *Proc R Soc B*. 2014; 280:20133172.
47. Centola D. The spread of behavior in an online social network experiment. *Science*. 2010; 329:1194–1197. [PubMed: 20813952]
48. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comp Biol*. 2011; 7:e1002199.
49. Ferrari MJ, et al. The dynamics of measles in sub-saharan Africa. *Nature*. 2008; 451:679–684. [PubMed: 18256664]
50. Hollingsworth TD, Ferguson NM, Anderson RM. Will travel restrictions control the international spread of pandemic influenza? *Nature Med*. 2006; 12:497. [PubMed: 16675989]
51. Perkins TA, Scott TW, Le Menach A, Smith DL. Heterogeneity, Mixing, and the Spatial Scales of Mosquito-Borne Pathogen Transmission. *PLoS Comp Biol*. 2013; 9:e1003327.
52. Wesolowski A, et al. Quantifying the impact of human mobility in malaria. *Science*. 2012; 338:267–270. [PubMed: 23066082]
53. Ross, R. *The Prevention of Malaria*. 2. John Murray; London: 1911.
54. Bartlett MS. Measles periodicity and community size. *J Roy Stat Soc*. 1957; 120:48–70.
55. Anderson RM, May RM. Directly transmitted infectious diseases: control by vaccination. *Science*. 1982; 215:1053–1060. [PubMed: 7063839]
56. McLean AR, Anderson RM. Measles in Developing Countries. Part II. The Predicted Impact of Mass Vaccination. *Epid Infect*. 1988; 100:419–442.

57. Saint-Victor DS, Omer SB. Vaccine refusal and the endgame: walking the last mile first. *Phil Trans R Soc B*. 2013; 368:20120148. [PubMed: 23798696]
58. Klepac P, Laxminarayan R, Grenfell BT. Synthesizing epidemiological and economic optima for control of immunizing infections. *Proc Natl Acad Sci USA*. 2011; 108:14366–14370. [PubMed: 21825129]
59. Salathé M, et al. Digital epidemiology. *PLoS Comp Biol*. 2012; 8:e1002616.
60. Salathé M, Freifeld CC, Mearu SR, Tomasulo AF, Brownstein JS. Influenza (H7N9) and the importance of digital epidemiology. *N Engl J Med*. 2013; 369:401–404. [PubMed: 23822655]
61. Read AF, Lynch PA, Thomas MB. How to Make Evolution-Proof Insecticides for Malaria Control. *PLoS Biol*. 2009; 7:e1000058. [PubMed: 19355786]
62. Arinaminpathy N, Lavine JS, Grenfell BT. Self-boosting vaccines and their implications for herd immunity. *Proc Natl Acad Sci USA*. 2012; 109:20154–20159. [PubMed: 23169630]
63. Diekmann, O.; Heesterbeek, JAP.; Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton Univ. Press; Princeton: 2013.
64. Jewell CP, Roberts GO. Enhancing Bayesian risk prediction for epidemics using contact tracing. *Biostatistics*. 2012; 13:567–579. [PubMed: 22674466]
65. Knock ES, O’Neill PD. Bayesian model choice for epidemic models with two levels of mixing. *Biostatistics*. 2014; 15:46–59. [PubMed: 23887980]
66. Dureau J, Kalogeropoulos K, Baguelin M. Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. *Biostatistics*. 2013; 14:541–555. [PubMed: 23292757]
67. King AA, Ionides EL, Pascual M, Bouma MJ. Inapparent infections and cholera dynamics. *Nature*. 2008; 454:877–880. [PubMed: 18704085]
68. Ratmann O, Donker G, Meijer A, Fraser C, Koelle K. Phylodynamic Inference and Model Assessment with Approximate Bayesian Computation: Influenza as a Case Study. *PLoS Comput Biol*. 2012; 8:e1002835. [PubMed: 23300420]
69. Farah M, Birrell PJ, Conti S, De Angelis D. Bayesian Emulation and Calibration of a Dynamic Epidemic Model for A/H1N1 Influenza. *J Amer Statist Assoc*. 2014; 109:1080-1090. doi:10.1080/01621459.2014.934453
70. Baguelin M, et al. Assessing optimal target populations for influenza vaccination programmes: an evidence synthesis and modeling study. *PloS Med*. 2013; 10:e1001527. [PubMed: 24115913]
71. De Angelis D, Presanis AM, Conti S, Ades AE. Estimation of HIV Burden through Bayesian Evidence Synthesis. *Statistical Science*. 2014; 29:9–17.
72. Ginsberg J, et al. Detecting influenza epidemics using search engine query data. *Nature*. 2009; 457:1012–1014. [PubMed: 19020500]
73. Bharti N, et al. Explaining seasonal fluctuations of measles in Niger using nighttime lights imagery. *Science*. 2011; 334:1424–1427. [PubMed: 22158822]
74. Messina JP, et al. Global spread of dengue virus types: mapping the 70-year history. *Trends in Microbiology*. 2014; 22:138–146. [PubMed: 24468533]
75. Dye C, et al. WHO and the future of disease control programmes. *Lancet*. 2013; 381:413–418. [PubMed: 23374479]
76. Dye C, O’Garra A. After 2015: infectious diseases in a new era of health and development. *Phil Trans R Soc B*. 2014; 369(1645)
77. Woolhouse MEJ, Farrar J. An intergovernmental panel on antimicrobial resistance. *Nature*. 2014; 509:555–557. [PubMed: 24877180]
78. Terry RF, Salm JF, Nannei C, Dye C. Creating a global observatory for health R&D. *Science*. 2014; 345:1302–1304. [PubMed: 25214621]
79. Rohani P, King A. Never mind the length, feel the quality: the impact of long-term epidemiological data sets on theory, application and policy. *Trends Ecol Evol*. 2010; 25:611–618. [PubMed: 20800928]
80. Halloran, ME.; Longini, IM.; Struchiner, CJ. *Design and Analysis of Vaccine studies*. Springer; Berlin: 2010.

81. Ndeffo ML, et al. Cost-effectiveness of a community-based intervention for reducing the transmission of *Schistosoma haematobium* and HIV in Africa. *Proc Natl Acad Sci USA*. 2013; 110:7952–7957. [PubMed: 23589884]
82. Pitman RJ, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM modeling good research practices task force-5. *Value Health*. 2012; 15:828–834. [PubMed: 22999132]
83. Woolhouse, MEJ., et al. Guide to good practice for quantitative veterinary epidemiology. 2011. <http://www.qve-goodpracticeguide.org.uk/>
84. Hollingsworth TD, Klinkenberg D, Heesterbeek H, Anderson RM. Mitigation strategies for pandemic influenza A: balancing conflicting policy objectives. *PLoS Comp Biol*. 2011; 7:e1001076.
85. Wu JT, Cowling BJ. The use of mathematical models to inform influenza pandemic preparedness and response. *Exp Biol Med*. 2011; 236:955–961.
86. Lee BY, Wiringa A. The 2009 H1N1 influenza pandemic: a case study of how modeling can assist all stages of vaccine decision-making. *Hum Vaccin*. 2011; 7:115–119. [PubMed: 21263227]
87. Lipsitch M, et al. Improving the evidence base for decision making during a pandemic: the example of 2009 influenzaA/H1N1. *Biosecur Bioterror*. 2011; 9:89–115. [PubMed: 21612363]
88. House T, et al. Modeling the impact of local reactive school closures on critical care provision during an influenza pandemic. *Proc R Soc B*. 2011; 278:2753–2760.
89. Van Kerkhove MD, Ferguson NM. Epidemic and intervention modelling - a scientific rationale for policy decision? Lessons from the 2009 influenza pandemic. *Bull World Health Organ*. 2012; 90:306–310. [PubMed: 22511828]
90. Luksza M, Lassig M. A predictive fitness model for influenza. *Nature*. 2014; 507:57–61. [PubMed: 24572367]
91. UK Government. 2012. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/224775/JCVI-statement-on-the-annual-influenza-vaccination-programme-25-July-2012.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224775/JCVI-statement-on-the-annual-influenza-vaccination-programme-25-July-2012.pdf)
92. Rose MA, et al. The epidemiological impact of childhood influenza vaccination using live-attenuated influenza vaccine (LAIV) in Germany: predictions of a simulation study. *BMC Inf Dis*. 2014; 14:40.
93. HIV Modeling Consortium. HIV treatment as prevention: models, data, and questions—towards evidence-based decision-making. *PloS Med*. 2012; 9:e1001259. [PubMed: 22802739]
94. Perelson AS, et al. HIV-1 dynamics in vivo: virion clearance rate, infected cell lifespan, and viral generation time. *Science*. 1996; 271:1582–1586. [PubMed: 8599114]
95. Faria NR, et al. The early spread and epidemic ignition of HIV-1 in human populations. *Science*. 2014; 346:56–61. [PubMed: 25278604]
96. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. *J Infect Dis*. 2008; 198:687–693. [PubMed: 18662132]
97. Bezemer D, et al. A resurgent HIV-1 epidemic among men who have sex with men in the era of potent antiretroviral therapy. *AIDS*. 2008; 22:1071–1077. [PubMed: 18520351]
98. Montaner JS, et al. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. *Lancet*. 2006; 368:531–536. [PubMed: 16890841]
99. 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*. 2009; 373:48–57. [PubMed: 19038438]
100. Cohen MS, et al. Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med*. 2011; 365:493–505. [PubMed: 21767103]
101. Bellan SE, et al. Extra-couple HIV transmission in sub-Saharan Africa: a mathematical modeling study of survey data. *Lancet*. 2013; 381:1561–1569. [PubMed: 23391466]
102. Tanser F, et al. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*. 2013; 339:966–971. [PubMed: 23430656]
103. Vermund SH, Fidler SJ, Ayles H, Beyers N, Hayes RJ. Can combination prevention strategies reduce HIV transmission in generalized epidemic settings in Africa? The HPTN 071 (PopART) study plan in South Africa and Zambia. *J Acquired Immune Defic Syndr*. 2013; 63(Suppl 2):S221–227. [PubMed: 23764639]

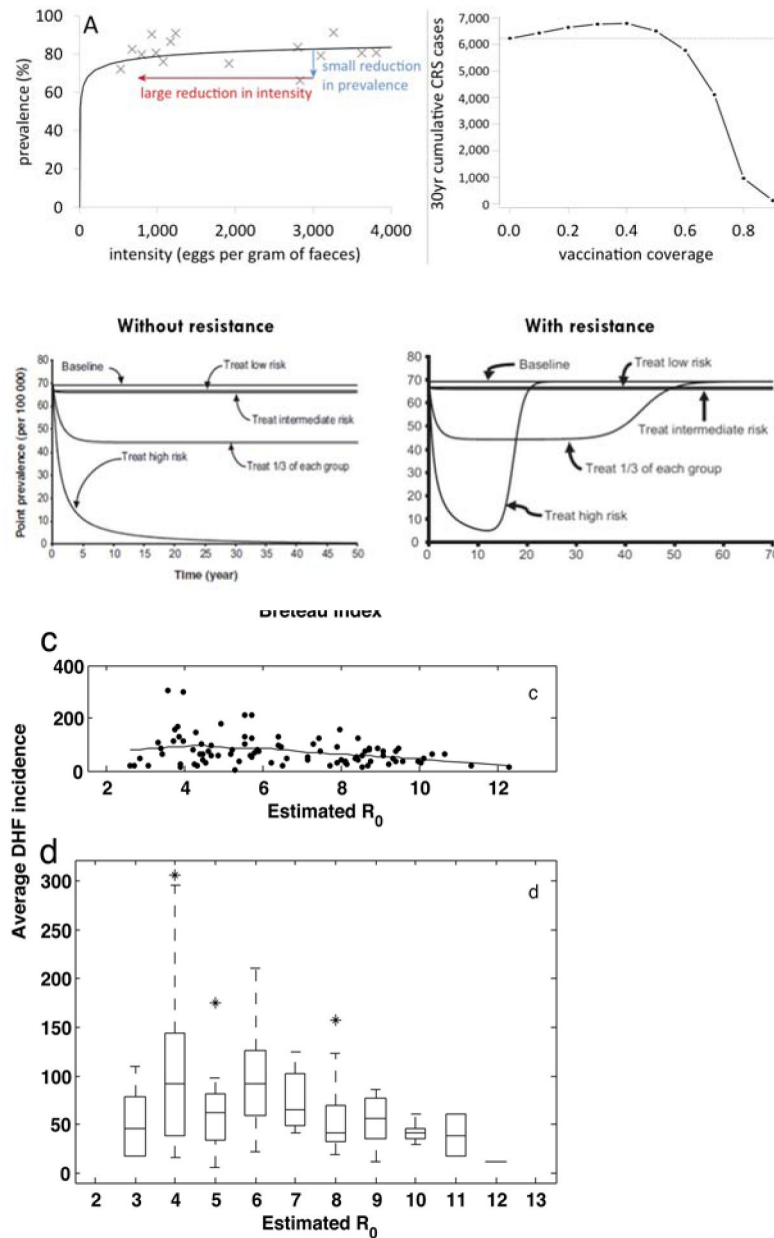
104. Smith MK, Powers KA, Muessig KE, Miller WC, Cohen MS. HIV treatment as prevention: the utility and limitations of ecological observation. *PLoS Med.* 2012; 9:e1001260. [PubMed: 22802740]
105. Cohen MS, et al. Antiretroviral treatment of HIV-1 prevents transmission of HIV-1: where do we go from here? *Lancet.* 2013; 382:1515–1524. [PubMed: 24152938]
106. Macdonald G. The analysis of equilibrium in malaria. *Trop Dis Bull.* 1952; 49:813–828. [PubMed: 12995455]
107. Dietz K, Molinieux L, Thomas A. A malaria model tested in the African Savannah. *Bull Wld Hlth Org.* 1974; 50:347–357.
108. Anderson, RM.; May, RM. *Infectious Diseases of Humans: Dynamics and Control.* Oxford Univ. Press; Oxford: 1992.
109. Perelson AS, Ribeiro RM. Modeling the within-host dynamics of HIV infection. *BMC Biol.* 2013; 11:96. [PubMed: 24020860]
110. Donnelly, CA.; Ferguson, NM. *Statistical Aspects of BSE and vCJD: Models for Epidemics.* Vol. 2000. Chapman & Hall/CRC; Boca Raton: 2000.
111. Keeling MJ. Models of foot-and-mouth disease. *Proc R Soc B.* 2005; 272:1195–1202.
112. Anderson RM, et al. Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic. *Phil Trans R Soc B.* 2004; 359:1091–1105. [PubMed: 15306395]
113. Blumberg S, Lloyd-Smith J. Inference of  $R_0$  and Transmission Heterogeneity from the Size Distribution of Stuttering Chains. *PLoS Comp Biol.* 2013; 9:e1002993.
114. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature.* 2005; 438:355–359. [PubMed: 16292310]
115. Matthews L, et al. Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control. *Proc Natl Acad Sci USA.* 2006; 103:547–552. [PubMed: 16407143]
116. Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. *Nature.* 2003; 422:428–433. [PubMed: 12660783]
117. Koelle K, Cobey S, Grenfell B, Pascual M. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science.* 2006; 314:1898–1903. [PubMed: 17185596]
118. Recker M, Pybus O, Nee S, Gupta S. The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc Natl Acad Sci USA.* 2007; 104:7711–7716. [PubMed: 17460037]
119. Funk S, Cecchi F, Nishiura H, Heesterbeek JAP, Edmunds J. Identifying transmission cycles at the human-animal interface: the role of animal reservoirs in maintaining gambiense Human African Trypanosomiasis. *PLoS Comp Biol.* 2013; 9:e1002855.
120. Reiner RC Jr, et al. A systematic review of mathematical models of mosquito-borne pathogen transmission 1970–2010. *J R Soc Interface.* 2013; 10:20120921. [PubMed: 23407571]
121. Stuckey EM, Smith TA, Chitnis N. Estimating malaria transmission through mathematical models. *Trends Parasitol.* 2013; 29:477–482. [PubMed: 24001452]
122. Reich NG, et al. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. *J R Soc Interface.* 2013; 10:20130414. [PubMed: 23825116]
123. Cooper BS, et al. Methicillin-resistant *Staphylococcus aureus* in hospitals and the community: Stealth dynamics and control catastrophes. *Proc Natl Acad Sci USA.* 2004; 101:10223–10228. [PubMed: 15220470]
124. Dye C, Williams BG. The population dynamics and control of tuberculosis. *Science.* 2010; 328:856–861. [PubMed: 20466923]
125. Read AF, Day T, Huijben S. The evolution of drug resistance and the curious orthodoxy of aggressive chemotherapy. *Proc Natl Acad Sci USA.* 2011; 108:10871–10877. [PubMed: 21690376]
126. Cobey S, Lipsitch M. Pathogen Diversity and Hidden Regimes of Apparent Competition. *Am Nat.* 2013; 181:12–24. [PubMed: 23234842]
127. Metcalf CJE, et al. Structured models of infectious disease: inference with discrete data. *Theor Pop Biol.* 2012; 82:275–282. [PubMed: 22178687]

128. Lessler J, Metcalf CJE. Balancing evidence and uncertainty when considering rubella vaccine introduction. *PloS One*. 2013; 8:e67639. [PubMed: 23861777]
129. Anderson RM, Hollingsworth D, Truscott J, Brooker S. Optimisation of mass chemotherapy to control soil-transmitted helminth infection. *Lancet*. 2012; 379:289–290. [PubMed: 22284644]
130. Metcalf CJE, et al. Implications of spatially heterogeneous vaccination coverage for the risk of congenital rubella syndrome in South Africa. *J R Soc Interface*. 2013; 10:20120756. [PubMed: 23152104]
131. Panagiotopoulos T, Antoniadou I, Valassi-Adam E. Increase in congenital rubella occurrence after immunisation in Greece: retrospective survey and systematic review. *Br Med J*. 1999; 319:1462–1467. [PubMed: 10582926]
132. Chan CH, McCabe CJ, Fisman DN. Core groups, antimicrobial resistance and rebound in gonorrhoea in North America. *Sex Transm Infect*. 2012; 88:200–204. [PubMed: 22169277]
133. Nagao Y, Koelle KK. Decreases in dengue transmission may act to increase the incidence of dengue hemorrhagic fever. *Proc Natl Acad Sci USA*. 2008; 105:2238–2243. [PubMed: 18250338]
134. Isaac Newton Institute IDD Collaboration: In addition to the authors already mentioned, this collaboration includes: Nimalan Arinaminpathy<sup>1</sup>, Frank Ball<sup>2</sup>, Tiffany Bogich<sup>3</sup>, Julia Gog<sup>4</sup>, Bryan Grenfell<sup>3</sup>, Alun Lloyd<sup>5</sup>, Angela Mclean<sup>6</sup>, Philip O’Neill<sup>2</sup>, Carl Pearson<sup>11</sup>, Steven Riley<sup>1</sup>, Gianpaolo Scalia Tomba<sup>12</sup>, Pieter Trapman<sup>13</sup>, James Wood<sup>7</sup>. Affiliations: <sup>1</sup>Imperial College, London, UK. <sup>2</sup>University of Nottingham, Nottingham, UK. <sup>3</sup>Princeton University, Princeton, USA. <sup>4</sup>University of Cambridge, Cambridge, UK. <sup>5</sup>North Carolina State University, Raleigh, USA. <sup>6</sup>University of Oxford, Oxford, UK. <sup>11</sup>University of Florida, Gainesville, USA. <sup>12</sup>University of Rome “La Sapienza”, Rome, Italy. <sup>13</sup>University of Stockholm, Stockholm, Sweden.



**Figure 1.**

Example of the process of modeling for public health, based on rubella. Policy questions are formulated; available data are brought to bear on the question, here illustrated by incidence of rubella following the introduction of vaccination in individuals aged less than 15 years, or 15 or more years in Costa Rica (127). Scientific understanding and subsequent policy advice is derived from model-based analyses of data, in this case using a non-linear age-structured SIR model (see Box 4), and can lead to collection of key missing data for model improvement. A plot illustrates insight where each square depicts a combination of birth rate and infant vaccine coverage (reflecting the situation in different countries, e.g. Somalia, diamond, Nepal, circle) for routine vaccination only. The color of the square indicates the confidence in the true value of  $R_0$  being higher (red) or lower (green) than a critical value that depends on birth rate and coverage. This translates into confidence that the public health burden caused by the rubella-complication of Congenital Rubella Syndrome (CRS) in newborns is likely to be reduced (green) compared to the situation without vaccination (128), suggesting that introduction of routine measles-rubella (MR) vaccine in Nepal is likely to succeed in bringing down CRS, while in Somalia routine MR vaccine would increase CRS burden without substantive improvements in the fraction that is vaccinated.



**Figure 2.** Examples of counter-intuitive effects of non-linear infection dynamics. *Upper left graph:* Non-linear interaction between prevalence of a helminth infection and infection pressure (as measured by the mean intensity of existing infections) means that control measures must have a disproportionately large impact on intensity before prevalence is reduced. This effect is predicted by a mathematical model (solid line) and corroborated by field data (crosses) (129). *Upper right graph* (adapted from 130): Non-linear relation between total number of cases of congenital rubella syndrome (CRS) and rubella vaccine coverage, showing that sub-optimal levels of vaccine coverage cause worse health outcomes than no vaccination. The line shows model predictions; similar effects have been documented for real rubella control situations (131). *Middle graphs* (from 132): Modeling results of rebound of gonorrhoea

transmission with different treatment strategies without (left panel) and with (right panel) antimicrobial resistance developing. In the presence of resistance, focusing treatment on the high-risk core group leads to an increase to un-treated baseline prevalence, after initially strong decline for more than a decade. *Bottom graphs* (from 133): Field data and box plot of a non-linear relation between  $R_0$  for dengue transmission and average dengue hemorrhagic fever incidence across Thailand, showing that starting control that brings down transmission from a situation with high  $R_0$  may paradoxically increase cases of DHF.



**Table 1**

As different infections have become the focus of public health attention, the modeling community has responded by developing improved concepts and methods. The table concentrates on the period since 1950. The first column lists the (classes) of infection, the second column lists factors whose importance to infection dynamics became particularly clear in relation to those infections; the third and fourth columns highlight concepts and methods that were developed in response. For each row, only a few typical references are given. Many factors, concepts and methods are relevant, in current use, and in continual development for much larger classes of infectious agents.

<b>Motivating studies</b>	<b>Important factors</b>	<b>Concepts</b>	<b>Methods</b>
Malaria (1910s, 1950s onwards, 53,106,107)	Transmission via insect vectors; nonlinear dependence of transmission on mosquito biting rate; influence of environmental and climatic variables.	Threshold for control, basic reproduction number.	Models with two host species (host-vector models); using models to support and guide field campaigns; relating models to field data.
Childhood infectious diseases, e.g. measles. (1950s onwards, 54,108)	Immunizing infections; spatial and temporal heterogeneity; demography; age structure; household structure.	Critical community size and herd immunity; periodic outbreaks; fade out; vaccine efficacy.	<i>SIR</i> models; age-structured models; models with periodic forcing; spatial and stochastic models; metapopulation models; time-series models.
Macroparasites. (1970s onwards, 108)	Clumped infections, multi-strain and multi-species infections, cross immunity, concurrent infections.	Consequences of overdispersed distribution of parasite load (Figure 1)	Stochastic models, approximations including hybrid models and moment closure.
Sexually transmitted infections, e.g. HIV. (1980s onwards, 108,109)	High/low risk groups; non-random contact structure; partnerships; within-host strain diversity and evolution; time scale.	Incubation and infectious period distribution; core group; next-generation matrix and operator; partnership dynamics.	Statistical methods (e.g. back calculation); Models with (dis)assortative mixing; pair formation models; within-host dynamic models.
Veterinary outbreaks, e.g. BSE and FMD. (1990s onwards, 110,111)	Fixed spatial locations with changing contact networks.	Local versus long-range transmission; spatial intervention (ring vaccination/culling); conflict of priorities at different scales.	Individual based models and spatial simulations (FMD); data-driven real-time modeling; inference of transmission trees.
Novel emerging infections, e.g. SARS, Nipah virus, MERS. (2000 onwards, 6,112–115)	Behavior change; global interconnectedness and international cooperation in control; responses in absence of biomedical measures; animal reservoirs.	Zoonotic spillover; stuttering chains; importance of index case; super-spreaders; unobserved dynamics in an animal reservoir; super-shedding	Contact tracing; modeling international spread and control; quarantine and case isolation; individual heterogeneity in infectiousness, incubation and latency period.
Influenza including avian influenza. (Present, 27,28,116–118)	Distribution of prior immunity; within-population and species strain differences, virus evolution and interaction; role of wildlife and farm animals.	Pandemics; spillover between wild birds, and farmed birds; phylodynamics.	Interaction between immunological and epidemiological dynamics; integrating phylogenetic and epidemic methods and models.
Vector-borne diseases, e.g. dengue, malaria. (Present, 119–122)	The influence of climate and environment on vector and pathogen development; animal reservoirs; interaction between strains within-host and between-host.	Dilution effect and role of biodiversity in infectious disease dynamics; re-emerging infections.	Evolutionary impact of vaccines/ other interventions; synthesis of data from ecology and epidemiology; elimination modeling; statistical modeling of environmental vector suitability.
Bacterial infections, e.g. pneumococcal disease, MRSA and tuberculosis. (Present, 39,123–126)	Antibiotic/drug resistance; adaptive dynamics.	Vaccine effectiveness; interacting natural immunity and vaccine boosting.	Modeling interacting and emerging strains; stochastic models in small populations.