When will computational epidemiologists be replaced by AI?
Planes, trains and autodidacts
HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time

Alan S. Perelson, Avidan U. Neumann, Martin Markowitz, John M. Leonard, David D. Ho*

A new mathematical model was used to analyze a detailed set of human immunodeficiency virus–type 1 (HIV-1) viral load data collected from five infected individuals after the administration of a potent inhibitor of HIV-1 protease. Productively infected cells were estimated to have, on average, a life-span of 2.2 days (half-life $t_{1/2} = 1.6$ days), and plasma virions were estimated to have a mean life-span of 0.3 days ($t_{1/2} = 0.24$ days). The estimated average total HIV-1 production was $10.3 \times 10^9$ virions per day, which is substantially greater than previous minimum estimates. The results also suggest that the
Natural History of HIV Infection

- **Primary infection**
  - Possible acute HIV syndrome.
  - Wide dissemination of virus
  - Seeding of lymphoid organs

- **Clinical latency**

- **Opportunistic diseases**

- **Constitutional symptoms**

- **Death**

**Graph details**

- Y-axis: CD4+ T Cells/mm³
- X-axis: Weeks, Years
- Plasma Viremia Titre

Try math epidemiology!

Barrier to entry low (no PhD, unlike econ)

Potential for major impact

Reusable physics skills!
From enthusiasm to emergency
Lord Robert May & Sir Roy Anderson
Epidemiological patterns

Roy M. Anderson
"Epidemiological patterns..."
Host population
The simple SIR epidemic model

S
Susceptible

I
Infected

R
Recovered
The simple SIR epidemic model

\[ S \xrightarrow{\lambda} I \xrightarrow{\frac{1}{d}} R \]

\(\lambda = \text{force of infection}\)

\(d = \text{duration of infectiousness}\)
2009: H1N1 influenza pandemic
### Asian Lineage Avian Influenza A (H7N9) Virus

#### Background

Human infections with an Asian lineage avian influenza A (H7N9) virus ("Asian H7N9") were first reported in China in March 2013. Annual epidemics of sporadic human infections with Asian H7N9 viruses in China have been reported since that time. China is currently experiencing its 5th epidemic of Asian H7N9 human infections. This is the largest annual epidemic to date. As of September 13, 2017, the World Health Organization (WHO) has reported 764 human infections with Asian H7N9 virus during the 5th epidemic, making the largest epidemic to date. This brings the total cumulative number of human infections with Asian lineage H7N9 reported by WHO to 1562.

Additional infections have been reported, but not yet publically announced by WHO. During epidemics one through four, about 40 percent of people confirmed with Asian H7N9 virus infection died.

#### Asian H7N9 Outbreak Characterization

- Asian H7N9 virus infections in poultry in China
- Sporadic infections in people; most with poultry exposure
Middle East Respiratory Syndrome (MERS) is viral respiratory illness that was recently recognized in humans. It was first reported in Saudi Arabia in 2012 and has since spread to several other countries, including the United States. Most people identified as infected with MERS-CoV developed severe acute respiratory illness, including fever, cough, and shortness of breath. Many of them have died.

ABOUT MERS
Information about MERS including symptoms and complications, how it spreads, prevention and treatment...

PEOPLE WHO MAY BE AT INCREASED RISK FOR MERS
Information for travelers from the Arabian Peninsula, contacts of ill travelers from this area, contacts of a confirmed case of MERS, healthcare personnel not using infection-control precautions, and people with exposure to camels...

Countries with Lab-Confirmed MERS Cases
Countries in or near the Arabian Peninsula with MERS cases: Bahrain, Iran, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, United Arab Emirates (UAE), and Yemen.
Infectious Disease Modeling Methods as Tools for Informing Response to Novel Influenza Viruses of Unknown Pandemic Potential

Manoj Gambhir,1,2,9 Catherine Bozzo,5,6 Justin J. O'Hagan,2,3,9 Amir Uzicanin,9 Lucinda E. Johnson,6 Matthew Biggerstaff,8 and David L. Swerdlow7

1Epidemiological Modelling Unit, Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia; 2Modeling Unit, National Center for Immunization and Respiratory Diseases (NCIRD), Centers for Disease Control and Prevention (CDC), 3InRc Inc, 4Graduate Program in Epidemiology and Molecules to Mankind, Laney Graduate School, Emory University, 5Division of Global Migration and Quarantine, National Center for Emerging and Zoonotic Infectious Diseases, 6Influence Division, and 7Modeling Unit and Office of the Director, NCIRD, CDC, Atlanta, Georgia

The rising importance of infectious disease modeling makes this an appropriate time for a guide for public health practitioners tasked with preparing for, and responding to, an influenza pandemic. We list several questions that public health practitioners commonly ask about pandemic influenza and match these with analytical methods, giving details on when during a pandemic the methods can be used, how long it might take to implement them, and what data are required. Although software to perform these tasks is available, care needs to be taken to understand: (1) the type of data needed, (2) the implementation of the methods, and (3) the interpretation of results in terms of model uncertainty and sensitivity. Public health leaders can use this article to evaluate the modeling literature, determine which methods can provide appropriate evidence for decision-making, and to help them request modeling work from in-house teams or academic groups.

The 2009 influenza A (H1N1) pandemic was one of the most closely tracked and studied epidemics in history. Traditional epidemiological methods, such as outbreak investigations and laboratory-based surveillance, were rapidly used to inform policy decisions [1-4]. These methods were enhanced by newer computational techniques such as bioinformatics and digital surveillance methods [5]. Simultaneously, substantial contributions during an outbreak of influenza with pandemic potential, public health leaders ask a range of questions to inform situational awareness, help assess severity [11] and guide decisions that aim to control the spread and impact of disease. Critical questions include:

- What is the case-fatality ratio?
- What is the case-hospitalization ratio?
<table>
<thead>
<tr>
<th>Questions</th>
<th>Pandemic Stage</th>
<th>Analysis Time Commitment</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the basic reproduction number ($R_0$) and the current value, or</td>
<td>P</td>
<td></td>
<td>Growth rate of case incidence curve</td>
</tr>
<tr>
<td>the time course, of the effective reproduction number ($R_{eff}$)?</td>
<td>E</td>
<td></td>
<td>Infection tree reconstruction</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Richards population growth model</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Chain binomial model</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Case renewal process</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Influenza genetic sequence analysis</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Age-structured SEIR model</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Maximum likelihood estimation</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Coalescent analysis</td>
</tr>
<tr>
<td>X</td>
<td>L</td>
<td>✔</td>
<td>Next generation matrix</td>
</tr>
<tr>
<td>What is the predicted peak number of cases and time? What is the</td>
<td>X</td>
<td></td>
<td>Age-structured SEIR model</td>
</tr>
<tr>
<td>predicted cumulative number of cases over the epidemic (ie, final</td>
<td></td>
<td></td>
<td>Digital surveillance methods</td>
</tr>
<tr>
<td>attack rate)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the possible spatiotemporal patterns of spread of the</td>
<td>X</td>
<td></td>
<td>Individual-based model</td>
</tr>
<tr>
<td>infection?</td>
<td></td>
<td></td>
<td>Metapopulation model</td>
</tr>
<tr>
<td>What was the likely sequence of spatiotemporal spread of infection</td>
<td>X</td>
<td></td>
<td>Infection tree reconstruction &amp; travel</td>
</tr>
<tr>
<td>since the outbreak began?</td>
<td></td>
<td></td>
<td>pattern modeling</td>
</tr>
<tr>
<td>What is the severity of the virus(es) (ie, case-hospitalization/death-</td>
<td>X</td>
<td></td>
<td>Individual-based model</td>
</tr>
<tr>
<td>rate) accounting for ascertainment biases (eg, more likely to detect</td>
<td></td>
<td></td>
<td>Bayesian evidence synthesis</td>
</tr>
<tr>
<td>severe cases)?</td>
<td></td>
<td></td>
<td>Incidence curve backcalculation</td>
</tr>
<tr>
<td>What is the transmission probability of the virus(es)?</td>
<td>X</td>
<td></td>
<td>Contact rate matrices &amp; SEIR model</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Data Needed</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Age-structured SEIR model</td>
<td>A compartmental model in which hosts are grouped into population compartments composed of their age-group and their infection status, e.g., an SEIR model. These models can be deterministic or stochastic, and the transitions between infection states are governed by contact and recovery rates [10, 12].</td>
<td>Case incidence stratified by age, contact matrix by age, cross-sectional serosurveys, physician visit/hospitalization rates to calculate symptomatic proportion/disease reporting rate/proportion immune, severity of infection across risk groups, initial number of infected individuals (or date on which the first infected individual was introduced into the population).</td>
<td></td>
</tr>
<tr>
<td>Antigenic cartography</td>
<td>A method for quantifying and visualizing the antigenic evolution of the influenza virus according to antigenic distances [13].</td>
<td>Influenza virus genetic sequences, antigenic distances between subtypes (using e.g., hemagglutination inhibition assay).</td>
<td></td>
</tr>
<tr>
<td>Antigenic distance</td>
<td>Antigenic distances of proposed vaccine strains from predicted dominant circulating strain(s) are correlated with prior years’ vaccine effectiveness estimates [14].</td>
<td>Hemagglutination inhibition assay distances of potential circulating strain(s) and record of vaccine effectiveness from prior years with amino acid sequences of past vaccine strains and dominant seasonal strains.</td>
<td></td>
</tr>
<tr>
<td>Bayesian evidence synthesis</td>
<td>Prior knowledge and distinct surveillance data sources are combined to estimate epidemiologic quantities (e.g., number infected, case-hospitalization rate) [9].</td>
<td>Repeated cross-sectional serosurveys, numbers and dates of onset of confirmed cases, symptomatic cases, hospitalizations, intensive care admissions, dates of severe outcomes.</td>
<td></td>
</tr>
<tr>
<td>Branching process analysis</td>
<td>Branching process theory is used to estimate the number of offspring of primary cases [15]. The generation time distribution between households and incidence of infection of households [16] is estimated.</td>
<td>Contact tracing data, surveillance datasets, $R_0$ population distribution (i.e., the probability associated with an individual in the population generating $R_0$ secondary cases at the start of the epidemic).</td>
<td></td>
</tr>
<tr>
<td>Case renewal process</td>
<td>Initial cases are modeled as a renewal process, which is a generalization of the Poisson process in which the time between cases is random and arbitrary, but independent and identically distributed [8, 17].</td>
<td>Case incidence time series (infection/hospitalization/death).</td>
<td></td>
</tr>
<tr>
<td>Chain binomial model</td>
<td>Initial cases are modeled as a discrete time chain of infections from one individual to another with probability of infection, or escape from infection, calculated using the binomial probability distribution [18].</td>
<td>Case incidence time series (infection/hospitalization/death).</td>
<td></td>
</tr>
<tr>
<td>Coalescent analysis</td>
<td>A Bayesian phylogenetic “coalescent” model is fitted to genetic sequence data obtained from isolates sampled from the infected population [19]. Growth rates of the epidemic are inferred.</td>
<td>Influenza genetic sequences and sampling times.</td>
<td></td>
</tr>
</tbody>
</table>
Key data for outbreak evaluation: building on the Ebola experience

Anne Cori, Christl A. Donnelly, Ilaria Dorigatti, Neil M. Ferguson, Christophe Fraser, Tini Garske, Thibaut Jombart, Gemma Nedjati-Gilani, Pierre Nouvellet, Steven Riley, Maria D. Van Kerkhove, Harriet L. Mills, and Isobel M. Blake

Medical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London W2 1PG, UK
Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK
Centre for Global Health, Institut Pasteur, 25-28 Rue du Dr Roux, 75015 Paris, France
MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK
School of Veterinary Sciences, University of Bristol, Bristol BS40 SDU, UK

Following the detection of an infectious disease outbreak, rapid epidemiological assessment is critical for guiding an effective public health response. To understand the transmission dynamics and potential impact of an outbreak, several types of data are necessary. Here we build on experience gained in the West African Ebola epidemic and prior emerging infectious disease outbreaks to set out a checklist of data needed to: (1) quantify severity and transmissibility; (2) characterize heterogeneities in transmission and their determinants; and (3)
Schematic illustrating the data needed to answer questions at different stages of the epidemic to inform the response.
Four key challenges in infectious disease modelling using data from multiple sources
Mathematical models: A key tool for outbreak response

Information. These models can clarify how the disease is spreading and provide timely guidance to policymakers. However, the use of models in public health often meets resistance (1), from doubts in peer review about the utility of such analyses to public skepticism that models can contribute when the means to control an epidemic are already known (2). Even when they are discussed in a positive light, models are often portrayed as arcane and largely inaccessible thought experiments (3). However, the role of models is crucial: they can be used to quantify the effect of mitigation efforts, provide guidance...
WARNING:
WE ARE NOT READY FOR
THE NEXT PANDEMIC

SCIENCE KNOWS
HOW TO FIGHT
AN OUTBREAK—
BUT POLICY STILL
GETS IN THE WAY
BY BRYAN WALSH

HOW TO KEEP THE
WORLD SAFE
BY BILL GATES
However, this isn’t working

At least not on a reasonable timescale

Math epi has been around for 5 decades but it’s barely used in public health agencies, unless...
The sky is falling down

Ebola
CDC Emergency Response
Activation Levels

Level 1
The highest level of response reserved for critical emergencies. CDC assigns the largest number of staff possible to work 24/7 on the response. To date, there have been three Level 1 responses: Ebola outbreak (2014), H1N1 influenza outbreak (2009) and Hurricane Katrina (2005).

Level 2
The CDC experts in the particular disease lead the response with a large number of other staff from the program area. A large number of staff from CDC’s Emergency Operations Center may assist with the response.

Level 3
The CDC experts in the particular disease lead the response with some of their own staff. Some staff from CDC’s Emergency Operations Center may assist in the response. CDC decides when a different level of response is needed.
Level 1

The highest level of response reserved for critical emergencies. CDC assigns the largest number of staff possible to work 24/7 on the response. To date, there have been three Level 1 responses: Ebola outbreak (2014), H1N1 influenza outbreak (2009) and Hurricane Katrina (2005).

Level 2
CDC leaders integral to the Ebola response, including epidemiologists, laboratorians, logistics, and more, assemble in agency’s command center to discuss next steps in directing the response at CDC Emergency Operations Center in Atlanta, August 8.

Spencer Lowell for TIME magazine
Questions from leadership

How many cases might there be?

When will the epidemic end?

What will it take to end the epidemic?
Ebola estimate

Without intervention, the total number of Ebola cases in the West African countries of Liberia and Sierra Leone could top 1 million by January.

CUMULATIVE CASES, IN MILLIONS (Liberia, Sierra Leone)

- Without intervention
- With proper intervention (70% of cases confined to treatment centers)

June 24, July 24, Aug. 23, Sept. 22, Oct. 21, Nov. 21, Dec. 20, Jan. 20

SOURCE: Centers for Disease Control and Prevention

PATRICK CLARK
THE WASHINGTON POST
Ebola Cases Could Reach 1.4 Million Within Four Months, C.D.C. Estimates

Worst-Case Scenario Can Still Be Avoided

by Andrew E. Kramer

A Red Cross team removed the body of a woman believed to have died of Ebola in Monrovia, Liberia, last week. Officials urge caution in handling victims' bodies.
Job creation/destruction
So, why is it so hard to get traction?

Policymakers don’t trust the model(s), they trust the person presenting the model.

They don’t trust single models, they need ensembles.

They’re comfortable with statistics but not mechanistic modelling.
Multi-model ensembles
IPCC report (AR4)
Neglected Tropical Disease Modelling Consortium

- **9 universities**: Warwick, Yale, Erasmus, Notre Dame, Imperial College London, Case Western Reserve, Monash, London and Liverpool Schools of Hygiene

- **9 diseases** incl: schistosomiasis, lymphatic filariasis, trachoma, soil transmitted helminths
2 questions from BMGF

Are we on target for the 2020 goals with current strategies?
If not, what other strategies will be required, and where?
Probabilistic forecasts of trachoma transmission at the district level: A statistical model comparison

Amy Pinsent a,⁎, Fengchen Liu b, Michael Deiner b,⁎, Paul Emerson c,d, Ana Bhaktiari c, Travis C. Porco b,e,f, Thomas Lietman b,e,f,g, Manoj Gambhir a

a Department of Public Health and Preventative Medicine, Monash University, Melbourne, Australia
b F.J. Proctor Foundation, University of California San Francisco, San Francisco, CA, USA
c International Trachoma Initiative, Atlanta, GA, USA
d School of Public Health, Emory University, Atlanta, GA, USA
e Department of Ophthalmology, University of California San Francisco, San Francisco, CA, USA
f Department of Epidemiology and Bio-statistics, University of California San Francisco, San Francisco, CA, USA
g Global Health Sciences, University of California San Francisco, San Francisco, CA, USA

ABSTRACT

The World Health Organization and its partners are aiming to eliminate trachoma as a public health problem by 2020. In this study, we compare forecasts of TF prevalence in 2011 for 7 different statistical and mechanistic models across 9 de-identified trachoma endemic districts, representing 4 unique trachoma endemic countries. We forecast TF prevalence between 1–6 years ahead in time and compare the 7 different models to the observed 2011 data using a log-likelihood score. An SIS model, including a district-specific random effect for the district-specific transmission coefficient, had the highest log-likelihood score across all 9 districts and was therefore the best performing model. While overall the deterministic transmission model was the least well-performing model, although it did comparably well to the other.
COORDINATING RESEARCH ACTIVITIES IN MATHEMATICAL MODELLING

HIV Modelling Consortium

The HIV Modelling Consortium aims to help improve scientific support for decision making by co-coordinating a wide range of research activities in mathematically modelling the HIV epidemic.
HIV Treatment as Prevention: Systematic Comparison of Mathematical Models of the Potential Impact of Antiretroviral Therapy

Abstract

Background: Many mathematical models have been developed to inform policy on new HIV infections. Comparing results from these models can highlight the differences in the assumptions and projections and hence provide insights into the reliability of the models. We compared a systematic review of mathematical models simulating the same epidemiological scenarios.

Methods and Findings: Twelve independent mathematical models were developed. Two different scenarios were evaluated: a) a fixed treatment threshold for treatment eligibility, access to treatment for all individuals with CD4 count below 350 cells per μL, and b) expanding the treatment threshold to 500 cells per μL. The models project that HIV incidence in South Africa will decrease by 39% and 55% for the former and latter scenarios, respectively, by 2020. The models estimated that treatment would prevent 21 million future infections by 2020.

Health benefits, costs, and cost-effectiveness of earlier eligibility for adult antiretroviral therapy and expanded treatment coverage: a combined analysis of 12 mathematical models

Summary

Background: New WHO guidelines recommend initiation of antiretroviral therapy for HIV-positive adults with CD4 counts of 500 cells per μL or less, a threshold that was previously recommended. Country decision makers have to decide whether to further expand eligibility for antiretroviral therapy accordingly. We aimed to assess the potential health benefits, costs, and cost-effectiveness of various criteria for adult antiretroviral therapy and expanded treatment coverage.

Methods: We used seven independent mathematical models in four settings—South Africa (generalized epidemic, moderate antiretroviral therapy coverage), Zambia (generalized epidemic, high antiretroviral therapy coverage), India (concentrated epidemic, moderate antiretroviral therapy coverage), and Vietnam (concentrated epidemic, low antiretroviral therapy coverage)—to assess the potential health benefits, costs, and cost-effectiveness of various eligibility criteria for adult antiretroviral therapy under scenarios of existing and expanded treatment coverage, with results projected over 20 years. Analyses assessed the extension of eligibility to include individuals with CD4 counts of 500 cells per μL or less, all HIV-positive adults compared with the previous (WHO) recommendation of initiation with CD4 counts of 350 cells per μL or less. We assessed costs from a health-system perspective, and the incremental cost of adding eligibility and the incremental cost-effectiveness ratio (ICER) were calculated. The ICER was calculated as the incremental cost (in US$) per disability-adjusted life-year (DALY) averted compared with the previous (WHO) eligibility criteria.

Findings: In South Africa, the cost per DALY averted of extending eligibility for antiretroviral therapy to adult patients with CD4 counts of 500 cells per μL or less ranged from $237 to $369 per DALY averted compared with 2001 guidelines. In Zambia, expansion of eligibility to adults with a CD4 count threshold of 500 cells per μL ranged from improving health outcomes while reducing costs (i.e., eliminating the previous guideline) to $120 per DALY averted. In Latin American and Caribbean countries, results were similar for expansion of eligibility to all HIV-positive adults, and when substantially expanded treatment coverage was assessed, expansion of treatment coverage to the antiretroviral model was cost-effective. In South Africa, the cost per DALY averted was less than three times the per capita GDP.
TB diagnosis

The practical process of preparing sputum samples for TB diagnosis.

Mission Statement

TB MAC aims to increase the effectiveness and efficiency of TB control policy and practice at global and country level.

We will do this by:
- building stronger and more effective links between decision makers, modelers, and researchers.
Modelling to guide the effective use of diagnostics in global health strategies
Improved training to data
The transmission dynamics of human immunodeficiency virus (HIV)

By R. M. May, F.R.S., and R. M. Anderson, F.R.S.

1 Department of Biology, Princeton University, Princeton, N.J. 08544, U.S.A.

2 Department of Pure and Applied Biology, Imperial College,
London University, London SW7 2BB, U.K.
Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic

Roy M. Anderson¹*, Christophe Fraser¹, Azra C. Ghani¹, Christl A. Donnelly¹, Steven Riley¹, Neil M. Ferguson¹, Gabriel M. Leung², T. H. Lam² and Anthony J. Hedley²

¹Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, St Mary’s Campus, Norfolk Place, London W2 1PG, UK
²21 Sassoon Road, Faculty of Medicine Building, University of Hong Kong, Pokfulam, Hong Kong, China

Figure 12. The SARS epidemic in Hong Kong and the fit of a multi-compartment meta-population stochastic model (from Riley et al. 2003). The dots are reported SARS cases and the solid line is the best fit model. The vertical grey bars denote 95% prediction intervals.
Ebola Virus Disease in West Africa — The First 9 Months of the Epidemic and Forward Projections

WHO Ebola Response Team

ABSTRACT

BACKGROUND
On March 23, 2014, the World Health Organization (WHO) was notified of an outbreak of Ebola virus disease (EVD) in Guinea. On August 8, the WHO declared the epidemic to be a “public health emergency of international concern.”

METHODS
By September 14, 2014, a total of 4507 probable and confirmed cases, including 2296 deaths from EVD (Zaire species) had been reported from five countries in West Africa — Guinea, Liberia, Nigeria, Senegal, and Sierra Leone. We analyzed a detailed subset of data on 3343 confirmed and 667 probable Ebola cases collected in Guinea, Liberia, Nigeria, and Sierra Leone as of September 14.
Figure 4. Observed and Projected Case Incidence.

Observed and projected weekly case incidence in Guinea (Panel A), Liberia (Panel B), and Sierra Leone (Panel C) are shown on linear (upper panels) and logarithmic (lower panels) scales.
Figure 1. Case Reproduction Numbers and Weekly Incidence in Guinea, Liberia, and Sierra Leone.

Shown are the estimated case reproduction number ($R_t$) over time (upper panels) and the observed and projected weekly incidence (lower panels) of confirmed and probable cases of Ebola virus disease (EVD), according to the date of symptom onset, from the week beginning June 30, 2014, until the week beginning January 12, 2015, on the basis of data reported through December 7 for Guinea and November 30 for Liberia and Sierra Leone. The projections shown in the lower panels were generated from $R_t$ estimates derived from data on case incidence (daily situation reports) for the 7 weeks through December 7 for Guinea and November 30 for Liberia and Sierra Leone (the time period delineated by the vertical dotted lines).
Epidemic Prediction Initiative **BETA**
Moving forecasting from research to decisions.

EPI aims to improve the science and usability of epidemic forecasts by facilitating open forecasting projects with specific public health objectives. Links to current and past projects can be found below. Learn more about EPI [here](#).

**CURRENT PROJECTS**

**State FluSight 2017-18**
Seasonal Influenza Forecasting at the US State Level

**FluSight 2017-18**
Seasonal Influenza Forecasting

**Influenza Hospitalizations 2017-18**
Forecasting laboratory confirmed influenza hospitalizations
Prof Nicholas Reich: http://flusightnetwork.io
Automation
So, what does a mathematical epidemiologist do?

✓ Devises (and performs) data collection
✓ Cleans the data
✓ Selects appropriate mathematical models
✓ Trains those models on the data
✓ Forecasts/Nowcasts/Scenario Analyses
✓ Communicates results to leadership
Which of these can be automated?

- Devises (and performs) data collection
- Cleans the data
- Selects appropriate mathematical models
- Trains those models on the data
- Forecasts/Nowcasts/Scenario Analyses
- Communicates results to leadership
However, things are changing

✓ New data types
✓ New mathematical models
✓ New training methods
✓ New visualisation of data/results

So, the AI epidemiologist would need to be upgraded frequently
In addition...

- Open sourcing code and data (when possible): reproducibility
- Breakthrough in model training needed
- ML methods are flexible to adding in new data types
- ML models can be reusable: ‘transfer learning’
Let's get started,
Select a project or create a new one

- Dengue Fever
- Zika Virus
- Measles
- Dragon Pox
- Malaria
Calibrate Model

\[ C_T = \frac{\sum_{t=1}^{T} (K_T + b_t - F(T + b_t))^2}{\sum_{t=1}^{T} (K_T + b_t - T)^2} \]

**Change**
- Click variable to change
- to
- Enter new value

**Activity Log**

- Reset
- Change Dataset
- Change Model
- Generate Visuals
Thank you
Cases of whooping cough in United States highest since 1959

December 12, 2012 | By Don Sapatkin, Inquirer Staff Writer

With pertussis at its highest level nationally in a half-century, the Philadelphia region has been weathering a spike that in some places is more than triple the previous record set two years ago.

“We’re sort of way off the scale this year,” said Stephen Ostroff, Pennsylvania’s acting physician general. “It really started picking up in the summer, and once kids got back to school, the [pertussis] was already there.”

Cases of pertussis, also known as whooping cough, often decline in late fall into early winter. In Philadelphia, which recorded 50 cases for August - more typical of an entire year - infections plummeted last month. But there has been no major decrease statewide, Ostroff said.
How worried should we be about the whooping cough epidemic?

MARY-ROSE MACCOLL, The Australian  April 26, 2012  12:03AM

Newborn babies are most at risk of death from the disease.
Reported pertussis cases – 1922-2010

Number of cases

Year

SOURCE: CDC, National Notifiable Diseases Surveillance System and Supplemental Pertussis Surveillance System and 1922-1949, passive reports to the Public Health Service
Questions from leadership

Is the effectiveness and duration of protection of the new vaccine different to the old?
Table 1. Descriptions of the nested models that were fitted to the NNDSS incidence data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Protection duration of whole cell vaccine same as natural infection; acellular vaccine same as whole-cell</td>
<td>-9720</td>
</tr>
<tr>
<td>2</td>
<td>Protection duration of whole cell vaccine same as natural infection; different efficacy for acellular vaccine</td>
<td>-9570</td>
</tr>
<tr>
<td>3</td>
<td>Protection duration of whole cell vaccine same as natural infection; different protection duration for acellular vaccine</td>
<td>-9250</td>
</tr>
<tr>
<td>4</td>
<td>Protection duration of whole cell vaccine different from natural infection; acellular vaccine same as whole-cell</td>
<td>-9800</td>
</tr>
<tr>
<td>5</td>
<td>Protection duration of whole cell vaccine same as natural infection; protection duration and efficacy different for acellular vaccine</td>
<td>-8422</td>
</tr>
<tr>
<td>6</td>
<td>Whole cell vaccine protection duration different from natural infection; different efficacy for acellular vaccine</td>
<td>-9183</td>
</tr>
<tr>
<td>7</td>
<td>Whole cell vaccine protection duration different from natural infection; different protection duration for acellular vaccine</td>
<td>-9230</td>
</tr>
<tr>
<td>8</td>
<td>Whole cell vaccine protection duration different from natural infection; protection duration and efficacy different for acellular vaccine</td>
<td>-8417</td>
</tr>
</tbody>
</table>

The mean posterior values of the Deviance Information Criterion (DIC) of the models are given in the rightmost column.

doi:10.1371/journal.pcbi.1004138
http://journals.plos.org/ploscompbiol/article?id=info:doi/10.1371/journal.pcbi.1004138
Total incidence since vaccination began: model vs. data
Projecting forward in time
$VE^*coverage = 10\%$
40%
90%
RESEARCH ARTICLE

A Change in Vaccine Efficacy and Duration of Protection Explains Recent Rises in Pertussis Incidence in the United States

Manoj Gambhir\textsuperscript{1,2,3,*}, Thomas A. Clark\textsuperscript{4}, Simon Cauchemez\textsuperscript{5,6}, Sara Y. Tartof\textsuperscript{7}, David L. Swerdlow\textsuperscript{2,8}, Neil M. Ferguson\textsuperscript{8}

1 Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia, 2 Modeling Unit, National Center for Immunization and Respiratory Diseases (NCIRD), Centers for Disease Control and Prevention (CDC), Atlanta, Georgia, United States of America, 3 IHRC, Inc., Atlanta, Georgia, United States of America, 4 Meningitis and Vaccine Preventable Diseases Branch, Division of Bacterial Diseases, NCIRD, CDC, Atlanta, Georgia, United States of America, 5 Medical Research Council Centre for Outbreak Analysis and Modelling, Imperial College London, London, United Kingdom, 6 Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France, 7 Kaiser Permanente Southern California, Kaiser Permanente Research, Department of Research & Evaluation, Pasadena, California, United States of America, 8 Office of Science and Integrative Programs, NCIRD, CDC, Atlanta, Georgia, United States of America.

* manoj.gambhir@monash.edu
http://journals.plos.org/ploscompbiol/article?id=info:doi/10.1371/journal.pcbi.1004138
Projects throughout CDC

**Pertussis**  Explaining the recent upsurge in cases in 7-10 yos and rise in overall cases

**Ebola**  2014-2015 West African epidemic
Lessons
Modelling’s major contribution comes very early (when sit. awareness is poor)

Embed within a public health agency

Academic publication often isn’t useful during an emergency (but is afterward)
Thank you for your time!

Special thanks to:

David Swerdlow
Lyn Finelli
Carrie Reed
Matt Biggerstaff
Cristina Carias
Martin Meltzer
Rebekah Borse
Isaac Fung
Neil Ferguson
Simon Cauchemez
Christl Donnelly
Tom Clark
Ben Lopman
Amy Pinsent

+ many others
Figure: Temporal trends on Twitter and Google about Ebola and influenza (flu) before, during, and after Ebola cases in the USA, September to November, 2014
*Numbers are relative to the highest number of searches done on Google (for Ebola on Oct 16).
Table 2. Parameter estimates for the best-fitting model, Model 8 (models outlined in Table 1).

<table>
<thead>
<tr>
<th>Parameter description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vaccine efficacies &amp; waning</td>
<td></td>
</tr>
<tr>
<td>Whole-cell</td>
<td></td>
</tr>
<tr>
<td>Vaccine efficacy of 1st 3 doses/4th/5th dose</td>
<td>90% [87%, 94%]</td>
</tr>
<tr>
<td>Rate of loss of whole-cell vaccine immunity</td>
<td>$3 \times 10^{-5} \text{yr}^{-1} [2 \times 10^{-6}, 2 \times 10^{-5}]$, i.e., essentially lifelong</td>
</tr>
<tr>
<td>Acellular</td>
<td></td>
</tr>
<tr>
<td>Vaccine efficacy of 1st 3 doses/4th/5th dose</td>
<td>80% [78%, 82%]</td>
</tr>
<tr>
<td>Rate of loss of acellular vaccine immunity</td>
<td>$0.018 \text{yr}^{-1} [0.015, 0.020]$, i.e., average of approx. 50 yrs protection</td>
</tr>
<tr>
<td>Tdap</td>
<td></td>
</tr>
<tr>
<td>Vaccine efficacy</td>
<td>As acellular</td>
</tr>
<tr>
<td>Epidemiological Parameters</td>
<td></td>
</tr>
<tr>
<td>Basic reproduction number, $R_0$</td>
<td>11.0 [9.9, 11.5]</td>
</tr>
<tr>
<td>Rate of loss of natural immunity</td>
<td>$3 \times 10^{-5} \text{yr}^{-1} [2 \times 10^{-6}, 2 \times 10^{-5}]$, i.e., essentially lifelong (as for whole-cell)</td>
</tr>
<tr>
<td>Relative susceptibility of individuals to subsequent infection (with reference to naive individuals)</td>
<td>32% [29%, 35%]</td>
</tr>
<tr>
<td>Relative infectiousness of individuals with subsequent infection (with reference to primary-infected individuals)</td>
<td>17% [14%, 23%]</td>
</tr>
<tr>
<td>Year of reporting rate change</td>
<td>None</td>
</tr>
<tr>
<td>Mean reporting rate prior to change</td>
<td>6.0% [0.1%, 22%]</td>
</tr>
<tr>
<td>Mean reporting rate after change</td>
<td>n/a</td>
</tr>
</tbody>
</table>

http://journals.plos.org/ploscompbiol/article?id=info:doi/10.1371/journal.pcbi.1004138
Fig 3. Cross-sectional incidence of disease over age of population.

http://journals.plos.org/ploscompbiol/article?id=info:doi/10.1371/journal.pcbi.1004138
Disease modelers use math to try to provide a more precise picture of a certain situation or to predict how the situation will change, and have become critical in the world of infectious diseases. But the accuracy — or inaccuracy — of such models is increasingly a talking point.
Model equations

\[
\frac{dS}{dt} \quad \text{Susceptible}
\]

\[
\frac{dI}{dt} \quad \text{Infected}
\]

\[
\frac{dR}{dt} \quad \text{Recovered}
\]
Inflow & outflow

\[
\frac{dS}{dt} = -\beta S I
\]

\[
\frac{dI}{dt} = \beta S I - \frac{1}{d} I
\]

\[
\frac{dR}{dt} = \frac{1}{d} I
\]

- Susceptible
- Infected
- Recovered
As infecteds increase, *rate* increases
Assessing the International Spreading Risk Associated with the 2014 West African Ebola Outbreak

SEPTEMBER 2, 2014 - RESEARCH

This article is either a revised version or has previous revisions
Edition 1 - September 2, 2014

AUTHORS
Marcelo F. C. Gomes  Ana Pastore y Piontti  Luca Rossi  Dennis Chao  Ira Longini  M. Elizabeth Halloran  Alessandro Vespignani

ABSTRACT
Background: The 2014 West African Ebola Outbreak is so far the largest and deadliest recorded in history. The affected countries, Sierra Leone, Guinea, Liberia, and Nigeria, have been struggling to contain and mitigate the outbreak. The ongoing rise in confirmed and suspected cases, 2615 as of 20 August 2014, is considered to increase the risk of international dissemination, especially because the epidemic is now affecting cities with major commercial airports.

Method: We use the Global Epidemic and Mobility Model to generate stochastic, individual-based simulations...
CENTERS FOR DISEASE CONTROL & PREVENTION (CDC)

Sierra Leone EbolaResponse (ER)
Modeling the spread of disease impact & intervention
Version 3.0

Contributors: Michael Washington, Charisma Atkins, Martin Meltzer

Division of Preparedness & Emerging Infections
Health Economics & Modeling Unit (HEMU)
December 4, 2014
Questions from leadership

What’s a viable vaccine trial design during the outbreak?
Stepped wedge study design - 18 week phase-in of vaccination

Bolded square highlights follow-up time usable for efficacy analyses, excluding 21 day sero-conversion time. Usable follow-up weeks contain both unvaccinated and vaccinated cohort time.

<table>
<thead>
<tr>
<th>type of person-time</th>
<th>proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>unvaccinated</td>
<td>0.50</td>
</tr>
<tr>
<td>vaccinated, seroconverting</td>
<td>n/a</td>
</tr>
<tr>
<td>vaccinated, seroconverted</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Example Vaccination Groups: (1) facility HCW such as doctors, nurses, phlebotomists (2) facility support such as cooking and food delivery, housekeeping, sanitation (3) ambulance teams (4) burial teams. Each of 3 shifts is treated as a different Vaccination Group. Vaccination Groups and shifts are distributed evenly across Vaccination Weeks, with a vaccination weeks assigned at random.
Specific questions

Will an e.g. Cox Proportional Hazards approach be able to account for:

- Declining background disease risk
- Clustering of disease risk
- Healthy vaccinee effect
Example simulation (single model run):

Unvaccinated: 43 cases
0.81 cases/person-month

Vaccinated: 27 cases
0.51 cases/person-month

Hazard ratio: 0.55 (0.32 – 0.96)
Vaccine Efficacy: 45% (4% - 68%)

No cases included until first vaccinee reaches end of seroconversion period
Longer lag to accrue vaccinated cases
No bias: Predicted VE
1000 runs at each VE input (range 50% to 90%)
Statistical power and validity of Ebola vaccine trials in Sierra Leone: a simulation study of trial design and analysis

Steven E Bellan, Juliet R C Pulliam, Carl A B Pearson, David Chomparens, Spencer J Fox, Laura Skrip, Alison P Galvani, Manoj Gambhir, Ben A Lopman, Travis C Porco, Lauren Ancel Meyers, Jonathan Dushoff

Summary
Background Safe and effective vaccines could help to end the ongoing Ebola virus disease epidemic in parts of west Africa, and mitigate future outbreaks of the virus. We assess the statistical validity and power of randomised controlled trial (RCT) and stepped-wedge cluster trial (SWCT) designs in Sierra Leone, where the incidence of Ebola virus disease is spatiotemporally heterogeneous, and is decreasing rapidly.

Methods We projected district-level Ebola virus disease incidence for the next 6 months, using a stochastic model fitted to data from Sierra Leone. We then simulated RCT and SWCT designs in trial populations comprising geographically distinct clusters at high risk, taking into account realistic logistical constraints, and both individual- and cluster-level variations in risk. We assessed false-positive rates and power for parametric and non-parametric analyses of simulated trial data, across a range of vaccine efficacies and trial start dates.

Findings For an SWCT, regional variation in Ebola virus disease incidence trends produced increased false-positive rates (up to 0.15 at α=0.05) under standard statistical models, but not when analysed by a permutation test, whereas analyses of RCTs remained statistically valid under all models. With the assumption of a 6-month trial starting on Feb 18, 2015, we estimate the power to detect a 90% effective vaccine to be between 49% and 89% for an RCT, and between 6% and 26% for an SWCT, depending on the Ebola virus disease incidence within the trial population. We estimate that a 1-month delay in trial initiation will reduce the power of the RCT by 20% and that of the SWCT by 49%.

Interpretation Spatiotemporal variation in infection risk undermines the statistical power of the SWCT. This variation also undercuts the SWCT’s expected ethical advantages over the RCT, because an RCT, but not an SWCT, can prioritise vaccination of high-risk clusters.
Questions from leadership

Where should ETUs be constructed next?

Which neighboring countries are at the highest risk?
Can we learn from the business/start-up world too?

Research: do it once

Development: can it be done many times?

Product/Service: do it many, many times