

Uses and Abuses of Infectious Disease Models

Jason Andrews

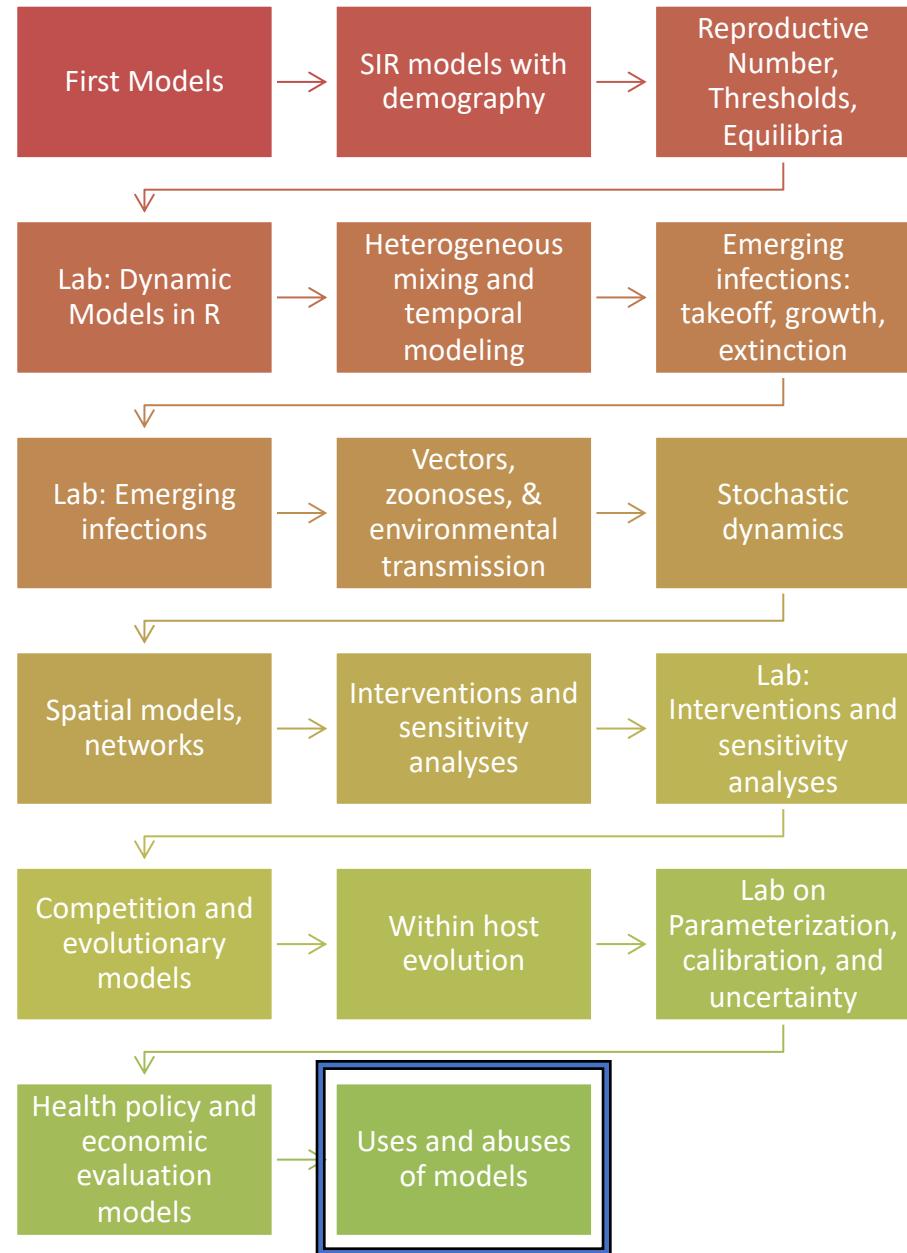
Jeremy Goldhaber-Fiebert

June 2, 2020

Announcements

- Review session / Q&A on Thursday 10:30-11:50
- Second midterm distributed this evening, due end of day June 11

Course Roadmap



Learning Objectives

- Become familiar with historical challenges in forecasting epidemics, and some of the drivers of forecasting errors
- Understand the importance of accurately capturing and conveying uncertainty
- Identify tradeoffs associated with model complexity
- Distinguish between mechanistic models and statistical / machine learning (ML) models for disease forecasting
- Understand how infectious disease models can be useful, and what properties can enhance their utility and reliability

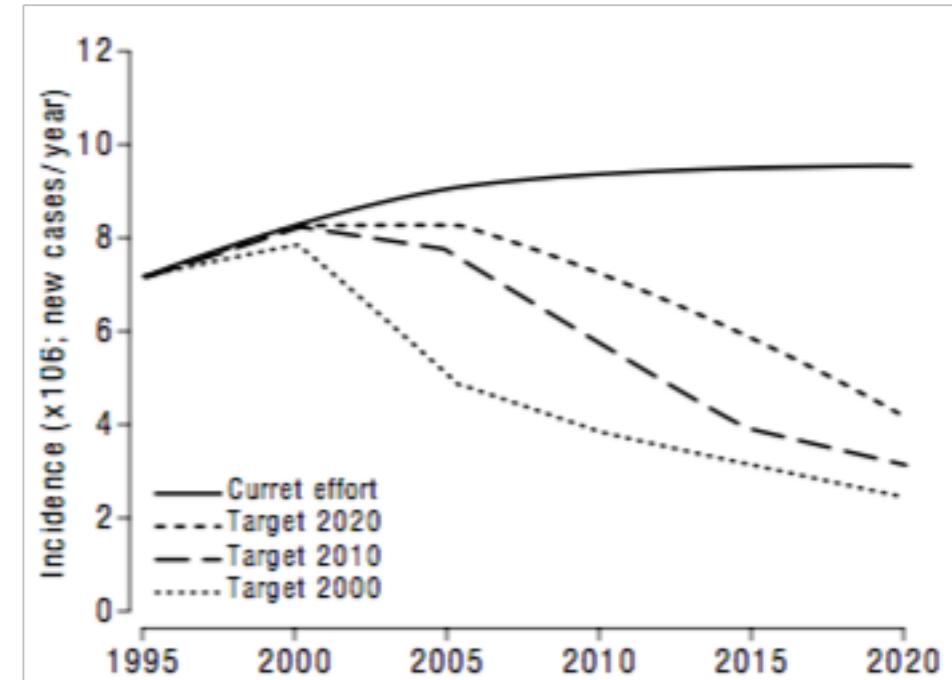
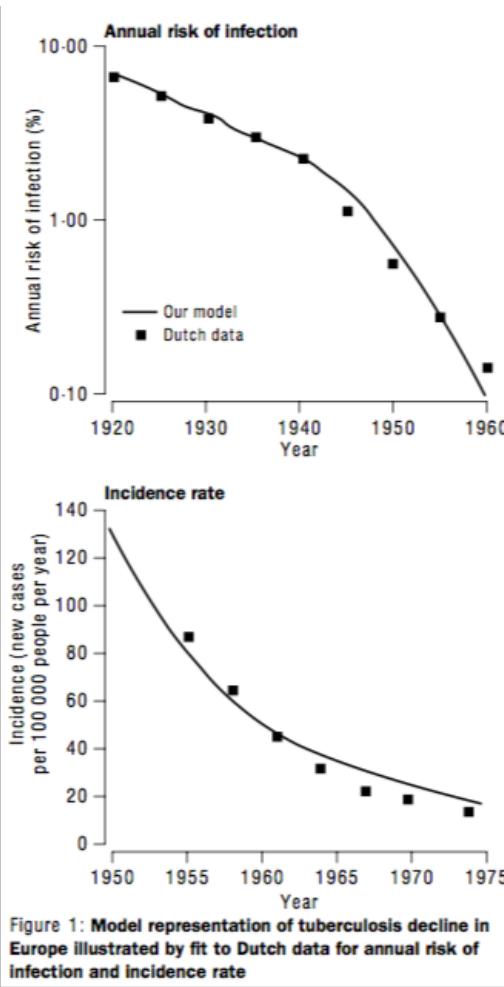
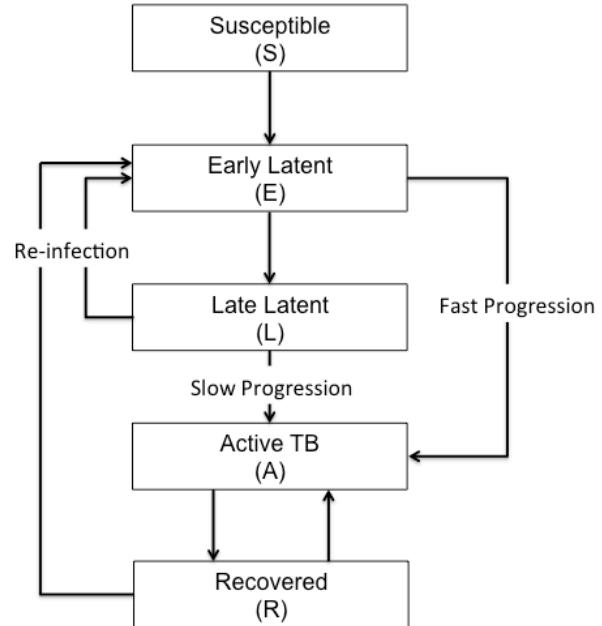
Uses of infectious disease models

- Explain observations of disease dynamics
 - Invasion thresholds, herd immunity, periodicity, age patterns of infection, traveling waves, heterogeneity and thresholds, evolution and fitness
- Estimate unknown parameters relevant to disease natural history and transmission of pathogens
 - Inferring serial and transmission intervals, effective contact rates in emerging epidemics, immunity duration
- Predict spread and future burden of disease
- Project the impact of potential interventions

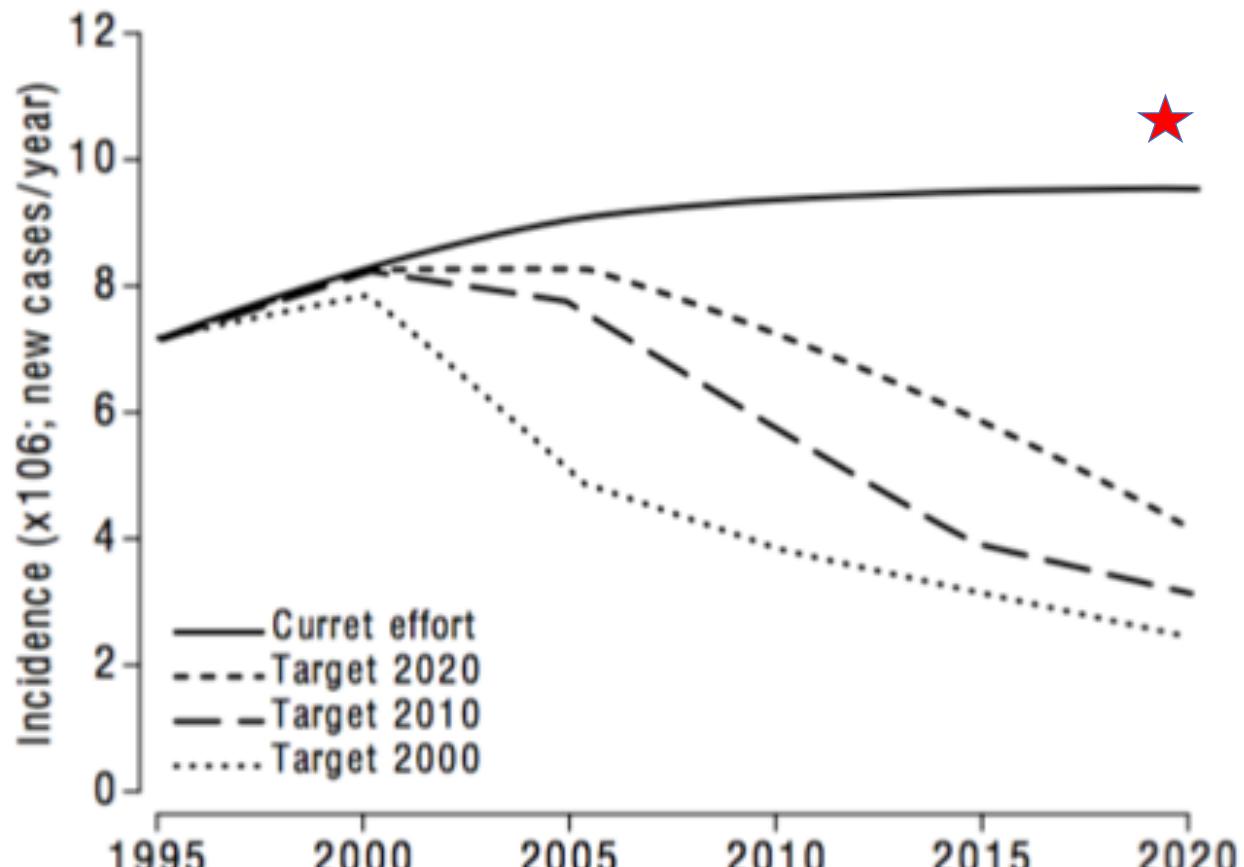
Models for Prediction / Forecasting

Prospects for worldwide tuberculosis control under the WHO DOTS strategy

Christopher Dye, Geoffrey P Garnett, Karen Sleeman, Brian G Williams



Projected Impact of TB Control Efforts



Centers for Disease Control and Prevention



Morbidity and Mortality Weekly Report

Supplement / Vol. 63 / No. 3

September 26, 2014

**Estimating the Future Number of Cases
in the Ebola Epidemic —
Liberia and Sierra Leone, 2014–2015**

“EbolaResponse”: CDC Ebola Model

TABLE 1. Calculated risk for onward transmission of Ebola, by patient category — EbolaResponse modeling tool, West Africa, 2014

Patient category	Values used to fit to data for Liberia and Sierra Leone*	Daily risk for onward transmission		No. infected per infectious person (95% CI) [§]	
		Values from literature†		From literature	Model estimates
		DRC (95% CI)	Uganda (95% CI)		
Hospitalized	0.02	0.1134 (0.00001–0.5842)	0.0017 (0.0–0.918)	0.4 (0–2.2) 0.01 (0–3.5)	0.12
Home or in a community setting such that there is a reduced risk for disease transmission (including safe burial when needed)	0.03	0.084 (0.06–0.313)	0.5045 (0.0576–0.5391)	0.5 (0.4–1.9) 2.6 (0.3–2.8)	0.18
Home with no effective isolation	0.3	1.0932 (0.00001–1.4281)	0.066 (0.0–3.0367)	1.8 (0–2.3) 0.1 (0–3.2)	1.8

Compartmental SIIR Model

FIGURE 4. Distribution of Ebola virus incubation period, by days of incubation

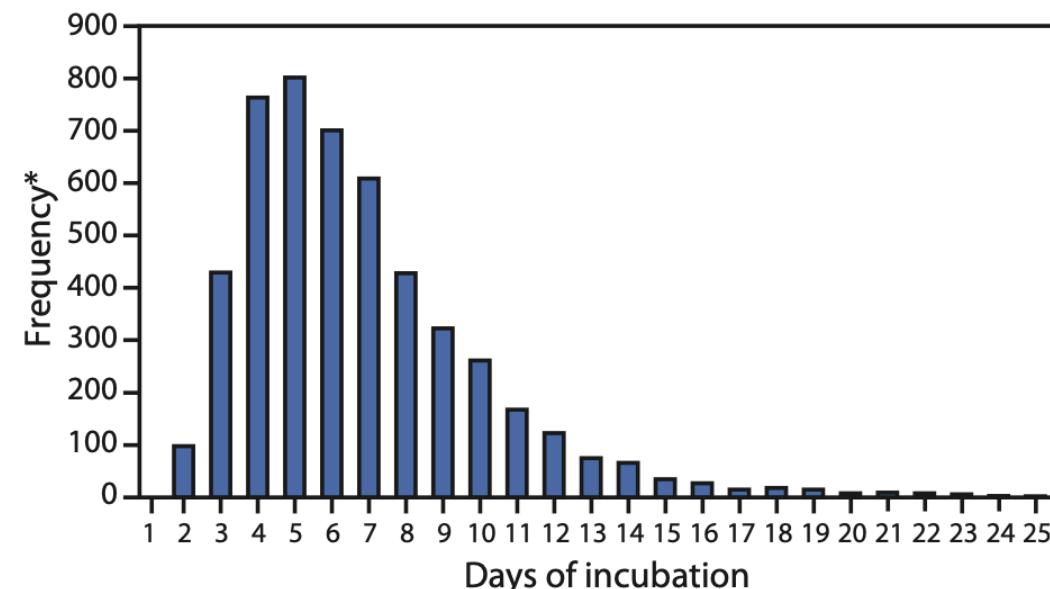
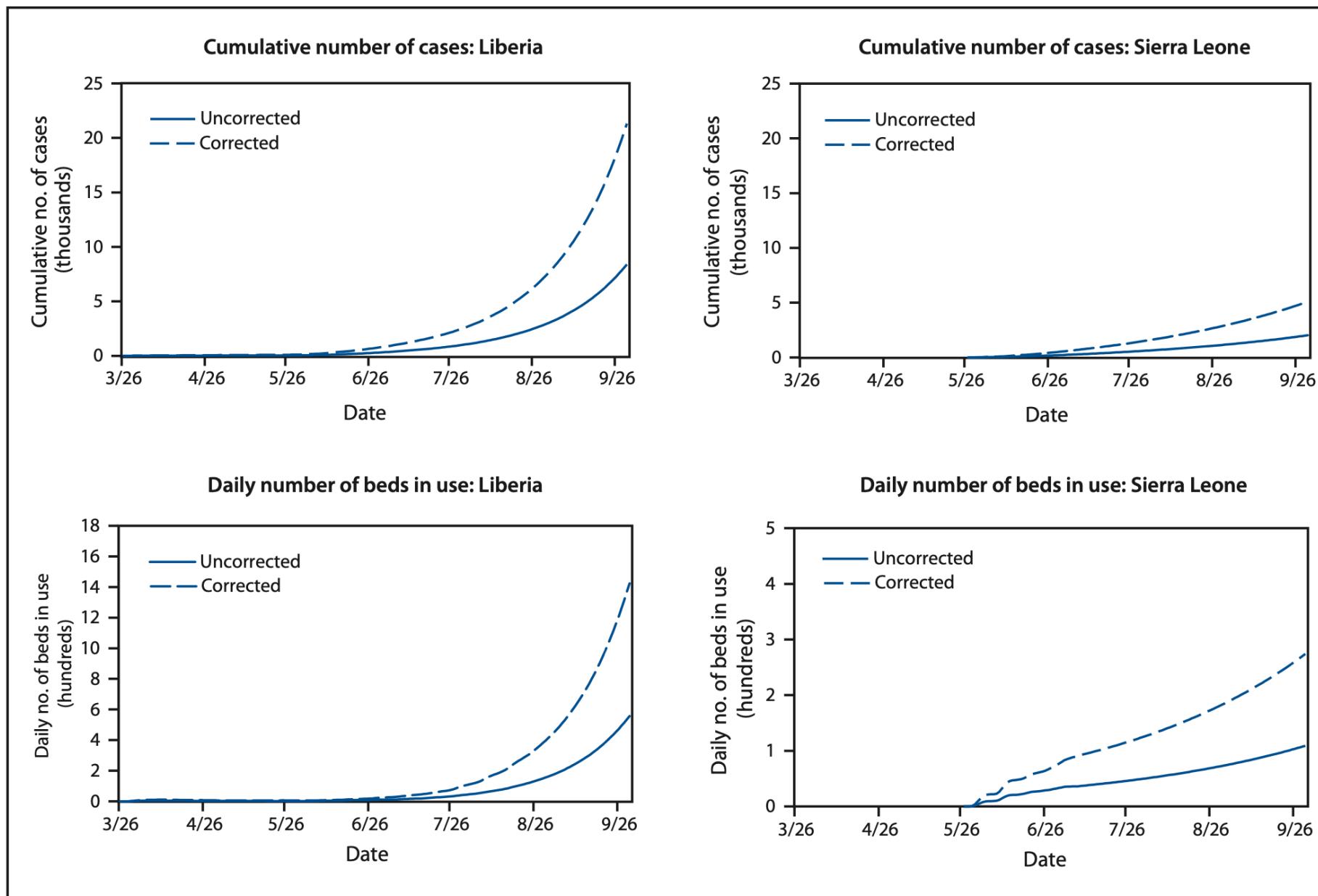


FIGURE 1. Estimated number of Ebola cases and daily number of beds in use,* with and without correction for underreporting,† through September 30
— EbolaResponse modeling tool, Liberia and Sierra Leone, 2014

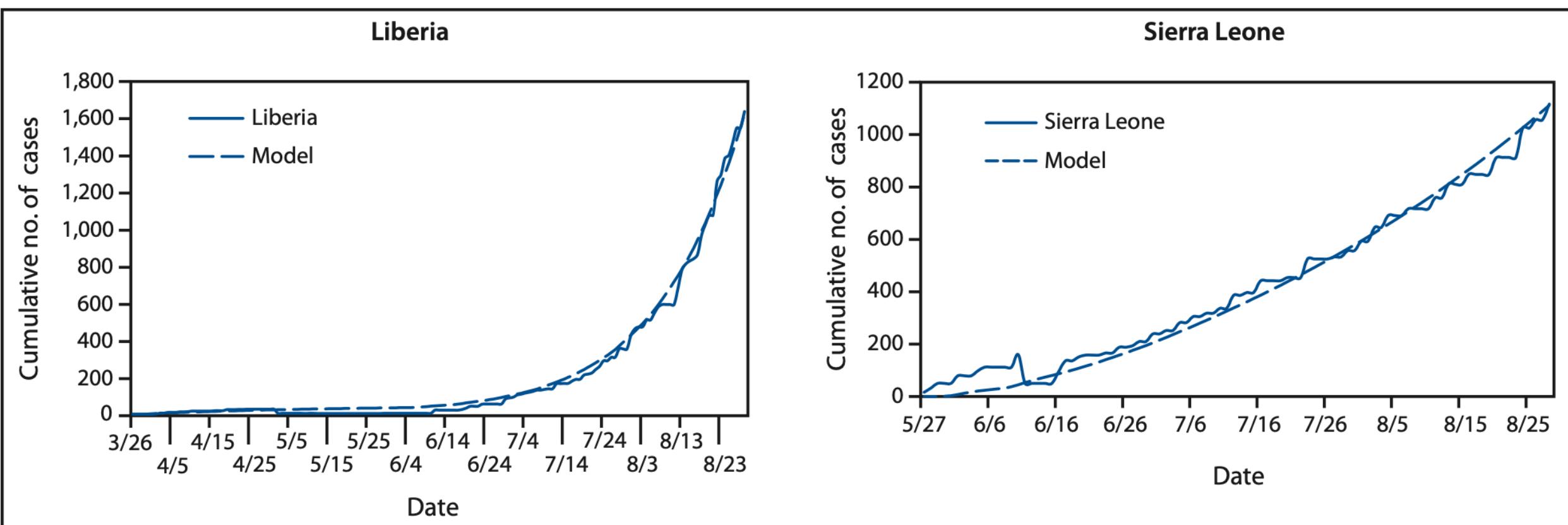


* Estimates of daily number of beds in use are calculated using estimates of likelihood of going to an Ebola treatment unit (ETU) and days in the ETU (Table 3).

† Corrected for potential underreporting by multiplying reported cases by a factor of 2.5 (Table 4).

CDC Model Calibration

FIGURE 6. Goodness of fit: comparison of cumulative reported and model-predicted numbers of Ebola cases* — EbolaResponse modeling tool, Liberia and Sierra Leone, 2014



Cumulative number of cases

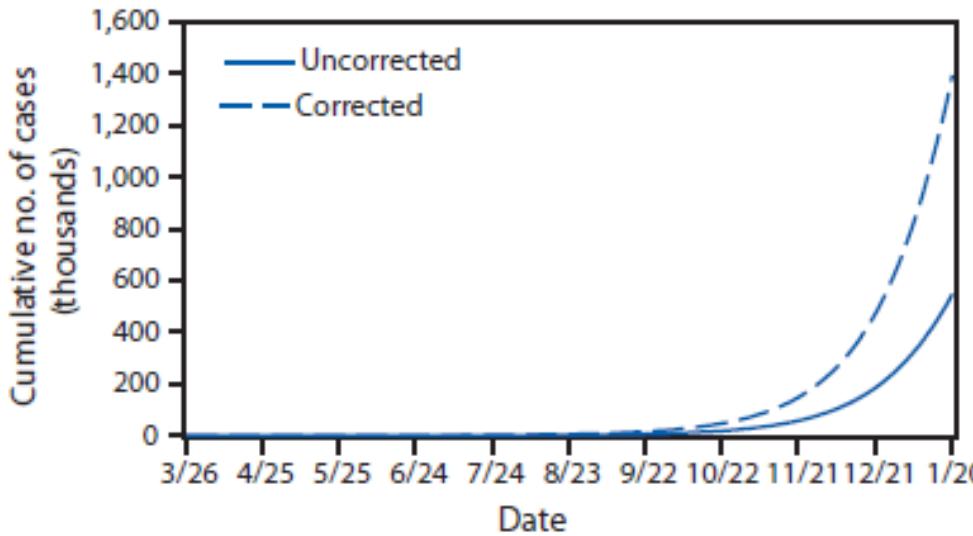
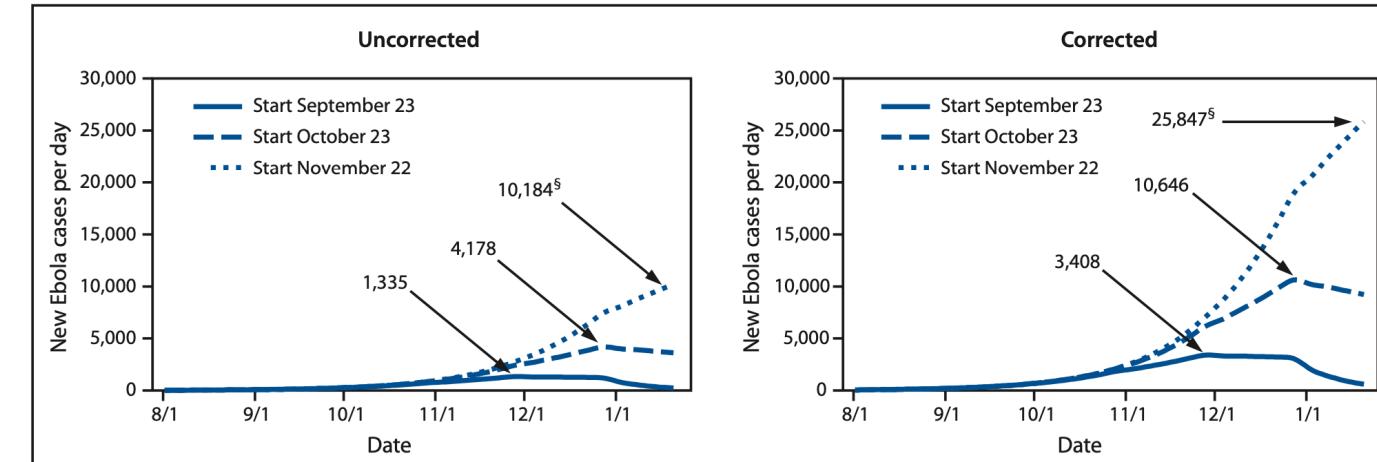
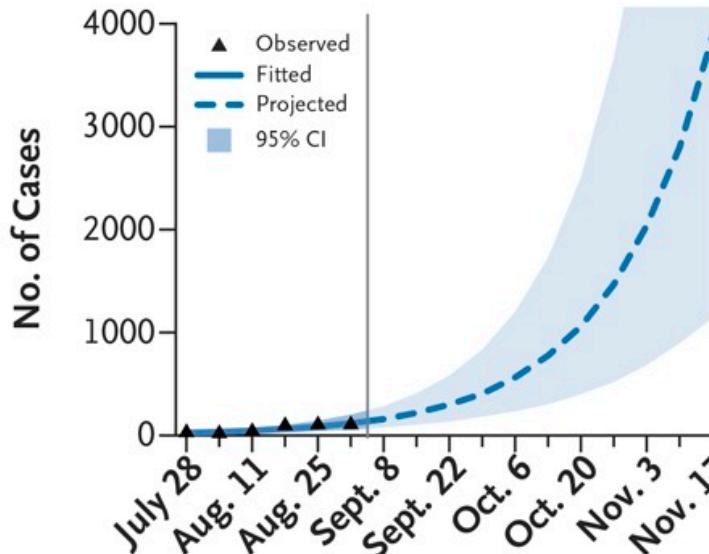


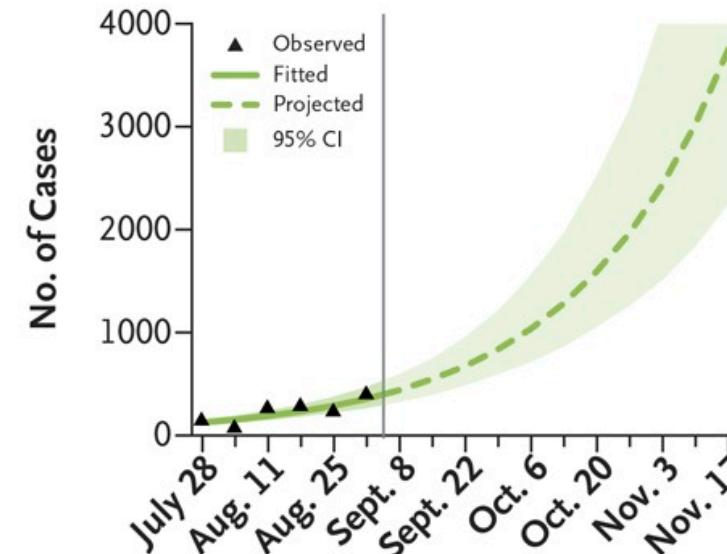
FIGURE 10. Estimated impact of delaying intervention* on daily number of Ebola cases, with and without correction for underreporting† — EbolaResponse modeling tool, Liberia, 2014–2015



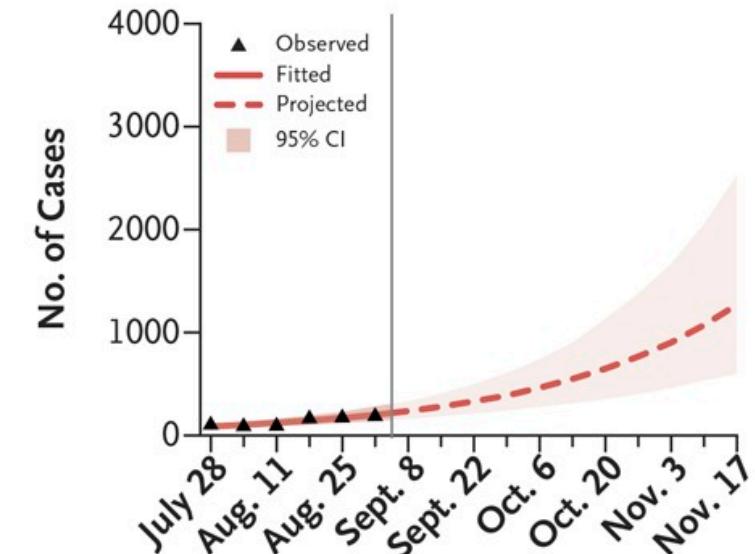
A Guinea



B Liberia



C Sierra Leone



Ebola Cases Could Reach 1.4 Million Within Four Months, C.D.C. Estimates

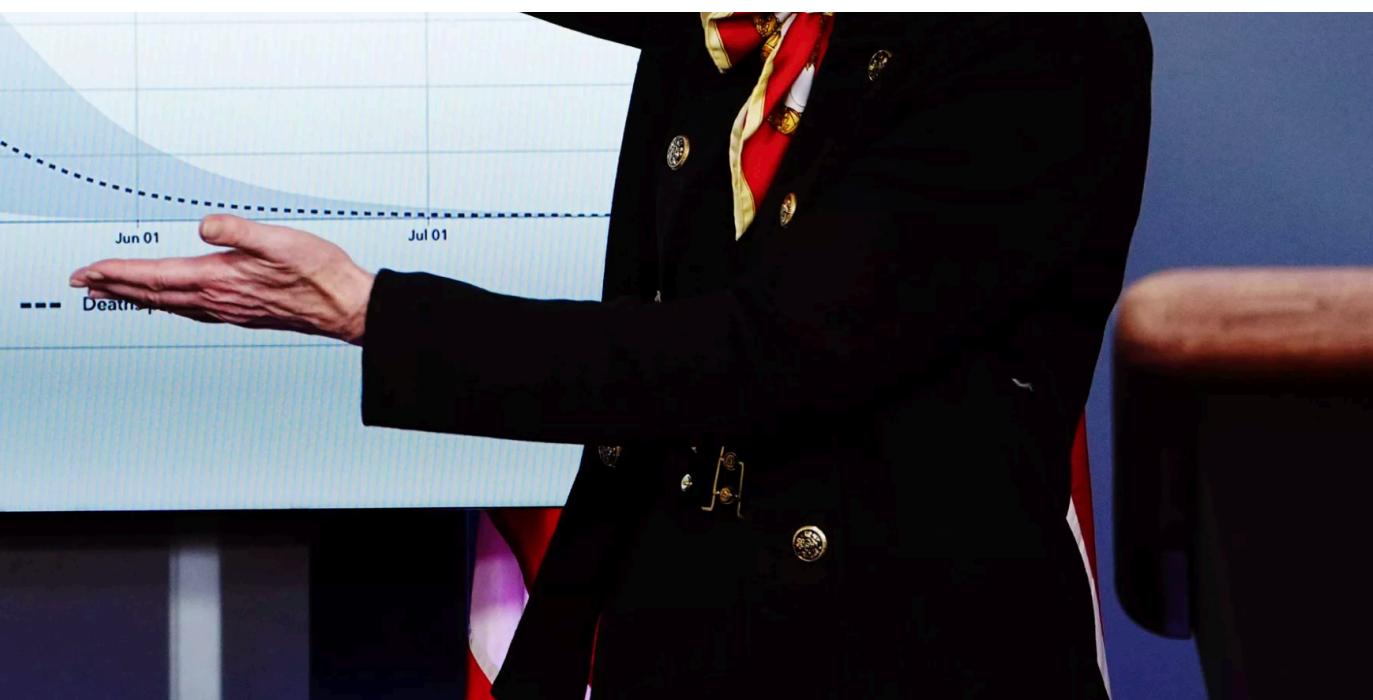
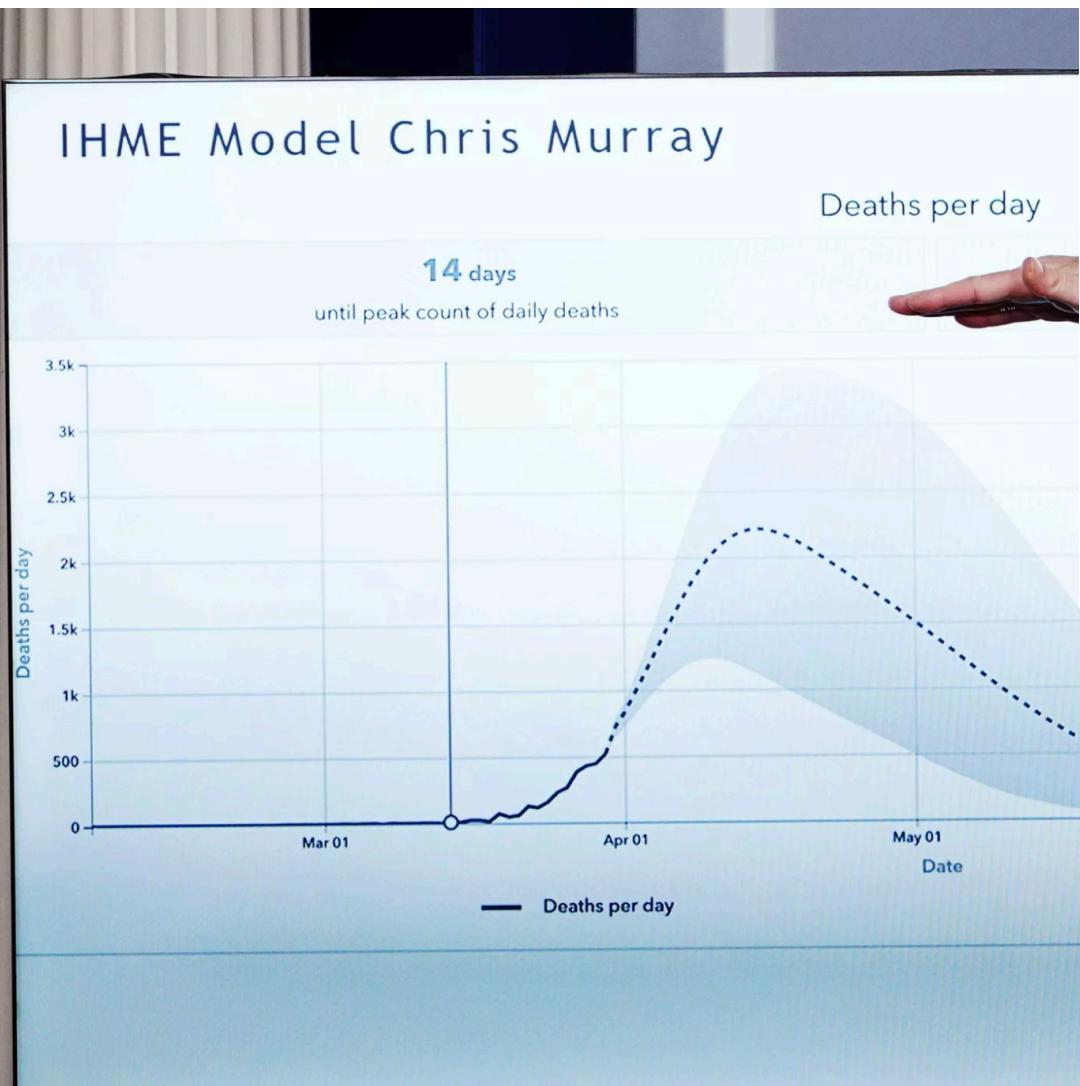
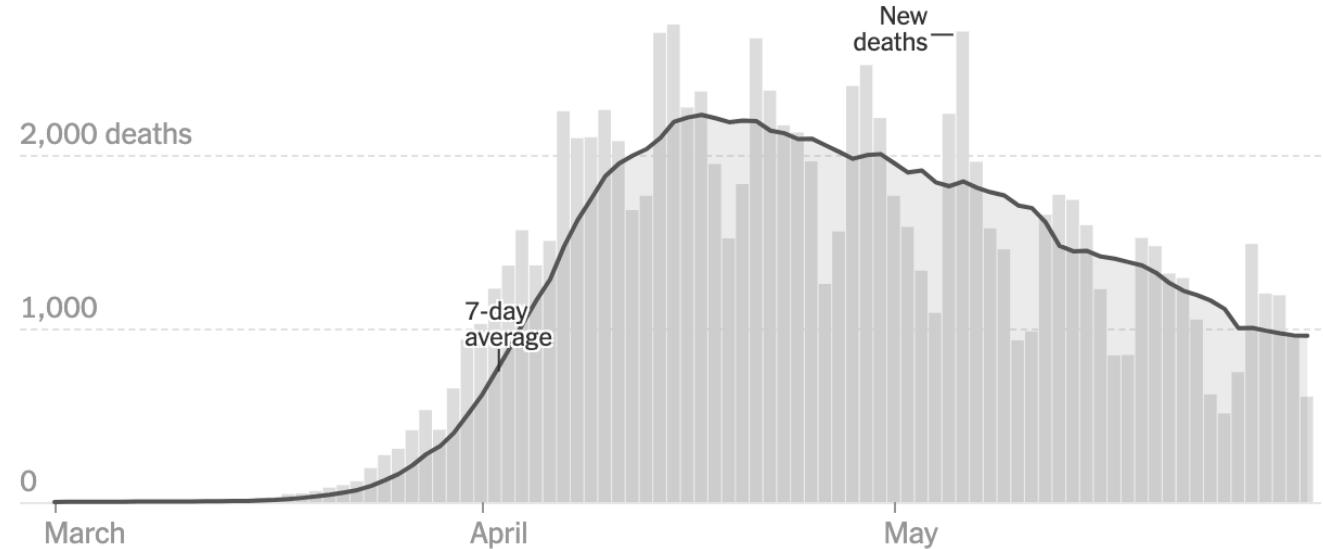
By DENISE GRADY SEPT. 23, 2014



Total Cases Reported: 28,616

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. Daniel Berehulak for The New York Times

New reported deaths by day in the United States



Influential Covid-19 model uses flawed methods and shouldn't guide U.S. policies, critics say

By SHARON BEGLE @sxbegle / APRIL 17, 2020



ADBE

Annals of Internal Medicine

Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic

Nicholas P. Jewell, PhD; Joseph A. Lewnard, PhD; and Britta L. Jewell, PhD

IDEAS AND OPINIONS

“It’s not a model that most of us in the infectious disease epidemiology field think is well suited” to projecting Covid-19 deaths, epidemiologist Marc Lipsitch of the Harvard T.H. Chan School of Public Health told reporters this week, referring to projections by the Institute for Health Metrics and Evaluation at the University of Washington.

Other experts, including some colleagues of the model-makers, are even harsher. “That the IHME model keeps changing is evidence of its lack of reliability as a predictive tool,” said epidemiologist Ruth Etzioni of the Fred Hutchinson Cancer Center, who has served on a search committee for IHME. “That it is being used for policy decisions and its results interpreted wrongly is a travesty unfolding before our eyes.”

This coronavirus model keeps being wrong. Why are we still listening to it?

A model that the White House has relied on has come under fire for its flawed projections.

By Kelsey Piper | May 2, 2020, 8:00am EDT

nature

Subscribe

NEWS FEATURE · 02 APRIL 2020 · CORRECTION [03 APRIL 2020](#)

Special report: The simulations driving the world's response to COVID-19

How epidemiologists rushed to model the coronavirus pandemic.

Neil Ferguson's Imperial model could be the most devastating software mistake of all time

The boss of a top software firm asks why the Government failed to get a second opinion from a computer scientist

DAVID RICHARDS AND KONSTANTIN BOUDNIK

16 May 2020 • 1:22pm

Computer code for Prof Lockdown's model which predicted 500,000 would die from Covid-19 and inspired Britain's 'Stay Home' plan is a 'mess which would get you fired in private industry' say data experts

- Professor Neil Ferguson's Imperial College London coding branded 'unreliable'
- University of Edinburgh scientists ran the same model and had different results
- Model was criticised early on by University of Oxford and public health expert
- Prof Ferguson left the government's Sage group after breaking lockdown rules
- [Here's how to help people impacted by Covid-19](#)

By [VANESSA CHALMERS](#) HEALTH REPORTER FOR MAILONLINE and [LUKE MAY](#)

PUBLISHED: 04:24 EDT, 17 May 2020 | UPDATED: 11:49 EDT, 17 May 2020

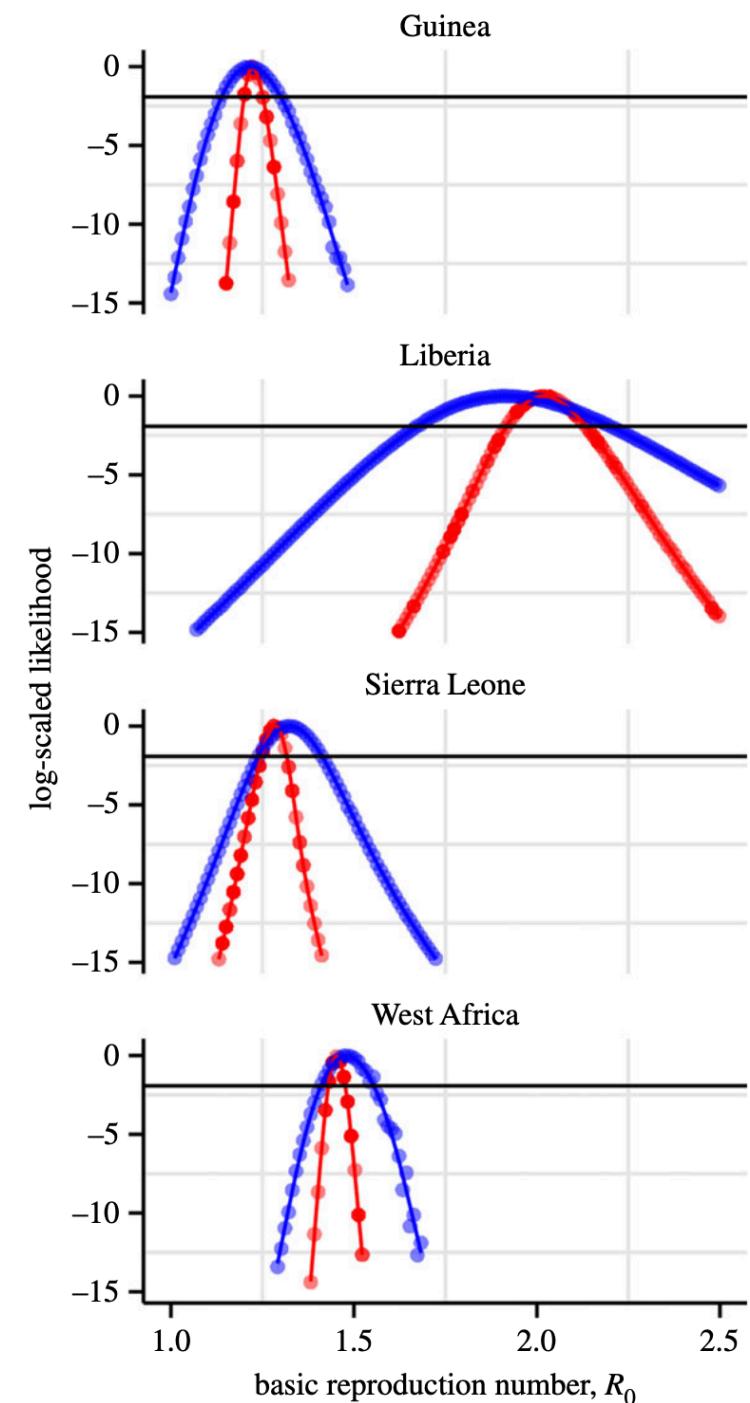
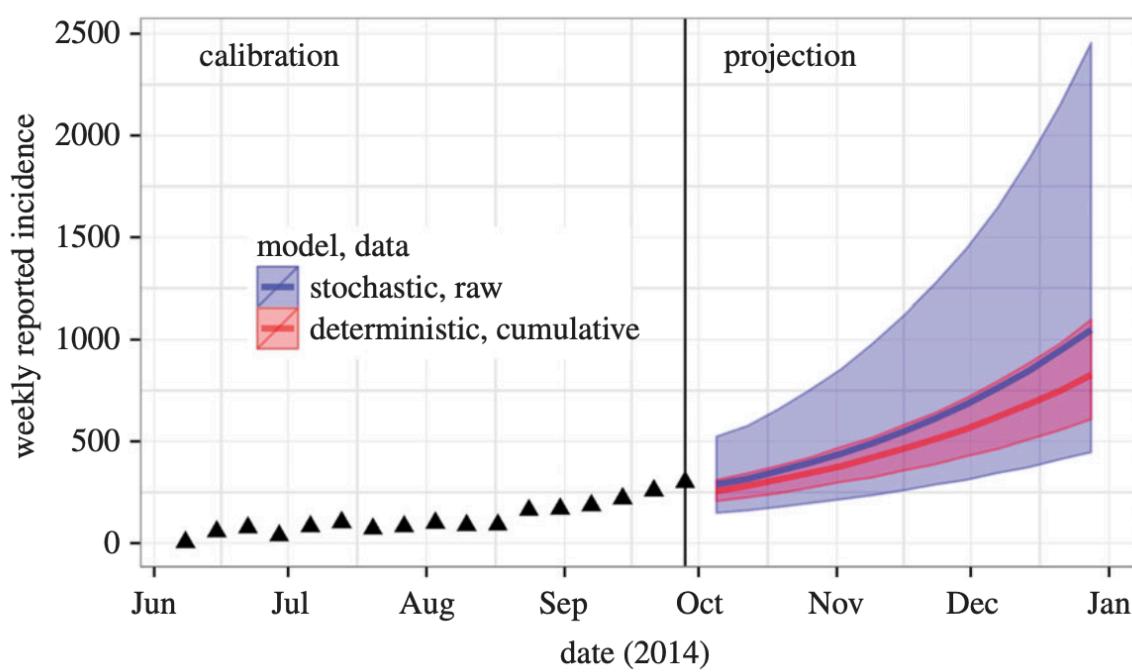
What makes a model good?

What properties enhance the utility
and reliability of models

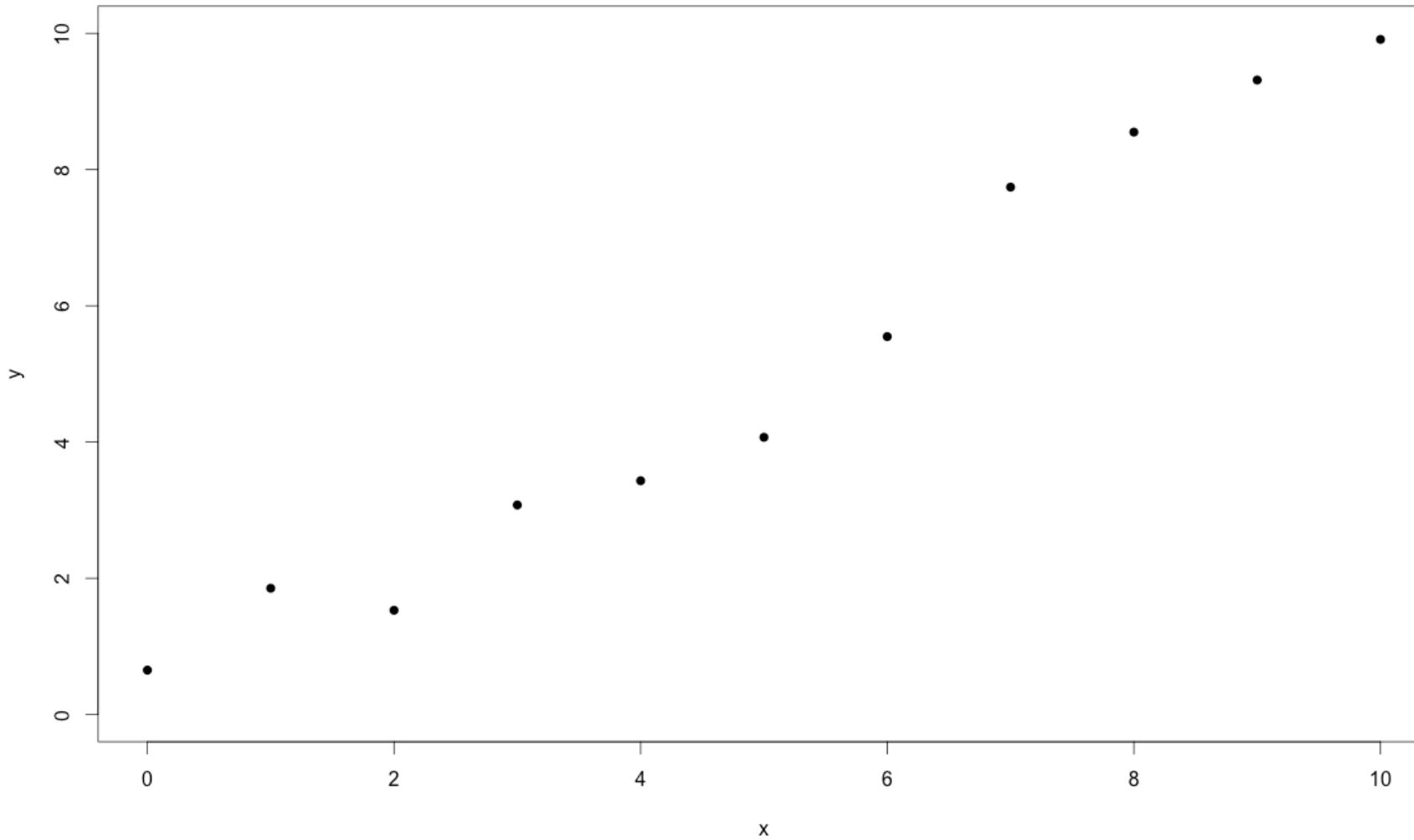
Properly Account for Uncertainty

Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola

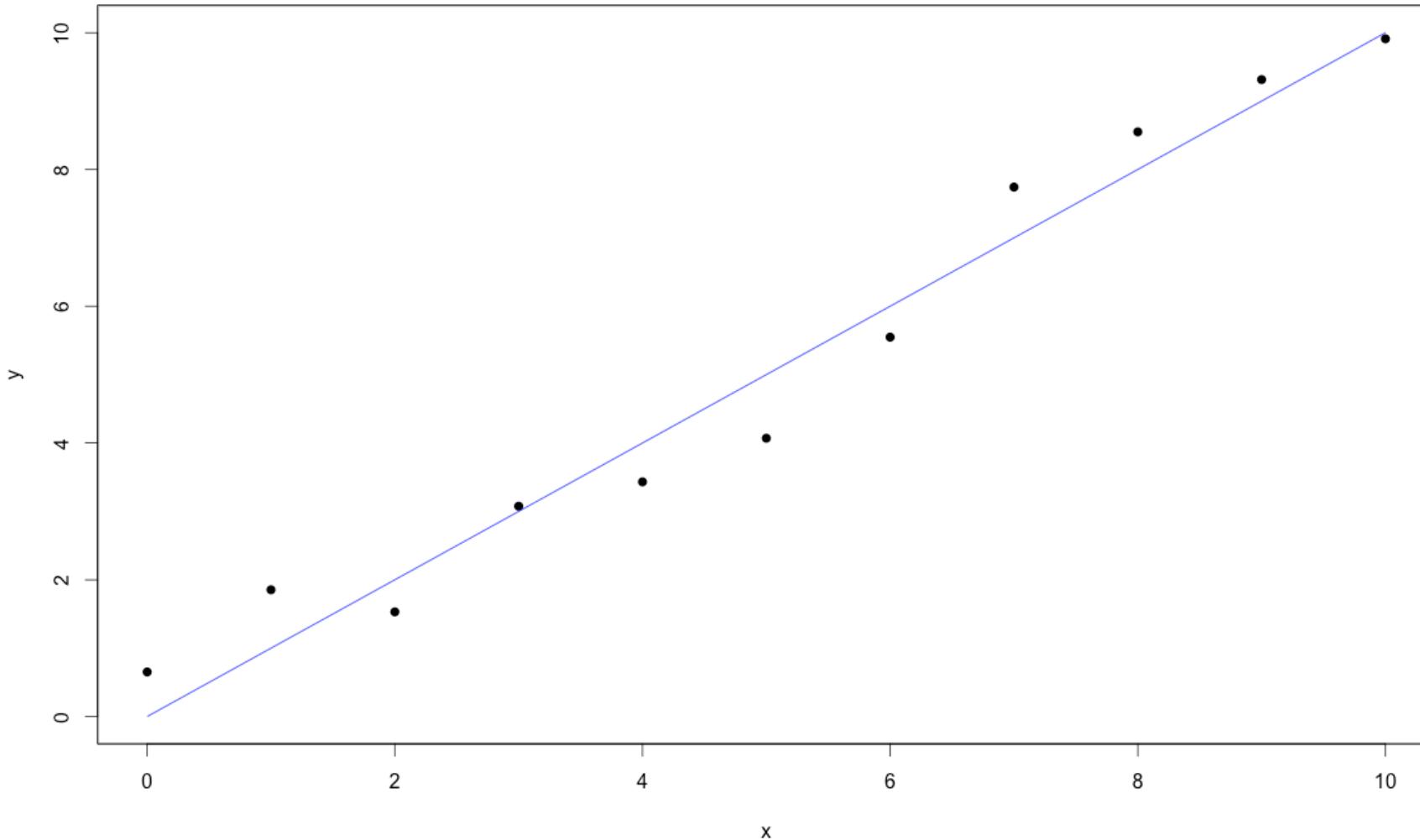
Aaron A. King^{1,2,3,4}, Matthieu Domenech de Cellès¹, Felicia M. G. Magpantay¹
and Pejman Rohani^{1,2,4}



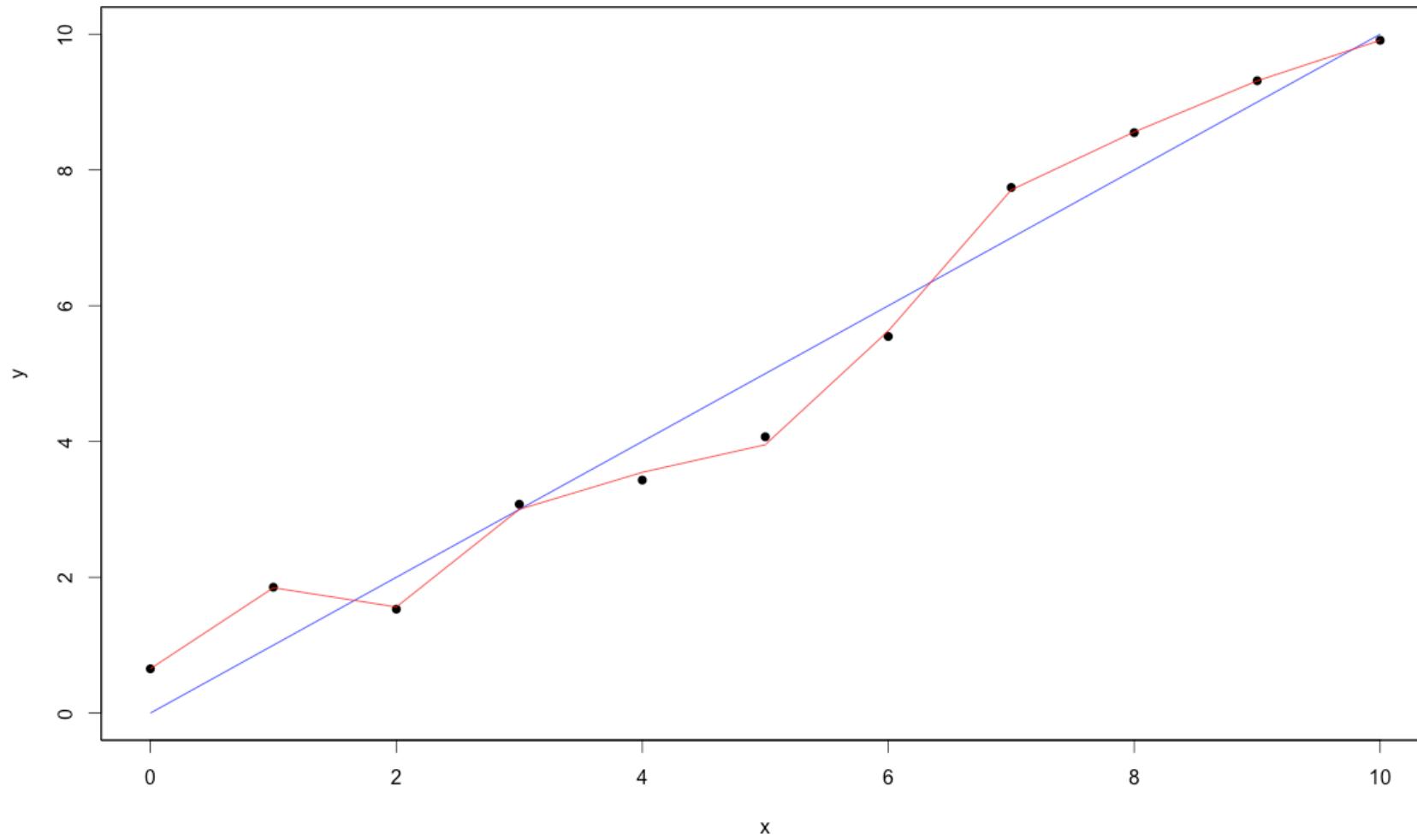
Model Fit?



Model Fit?



Model Fit?



Good Fit \neq Good Model

Model Complexity

- Complex models can (almost) always fit data better
- However, they require estimating more parameters, introduce more uncertainty, non-linearities and risk for misspecification
- “Over-fitting”
- Complex models are not necessarily more accurate or reliable simply because they fit better

OPEN  ACCESS Freely available online



Guidelines and Guidance

Complexity in Mathematical Models of Public Health Policies: A Guide for Consumers of Models

Sanjay Basu^{1,2,3,4*}, Jason Andrews⁵

Summary – Model Complexity

- There are always tradeoffs between creating models that are “realistic” versus those that are grounded in well characterized data about disease processes
- Complex models can suffer parameter estimation problems (too few data, too many unknowns)
- Models should capture essential components of disease processes
- When in doubt comparing simpler and more complex models can be valuable, penalizing goodness of fit (e.g. AIC)
- There are formal methods for assessing structural and parameter identifiability

“Everything should be made as simple as possible, but not simpler”
-(possibly) Einstein

Farr's Law Applied to AIDS Projections

Dennis J. Bregman, PhD, Alexander D. Langmuir, MD, MPH

Farr's Law of Epidemics, first promulgated in 1840 and resurrected by Brownlee in the early 1900s, states that epidemics tend to rise and fall in a roughly symmetrical pattern that can be approximated by a normal bell-shaped curve. We applied this simple law to the reported annual incidence of cases of acquired immunodeficiency syndrome in the United States from 1982 through 1987. The 6 years of incidence data closely fit a normal distribution that crests in late 1988 and then declines to a low point by the mid-1990s. The projected size of the epidemic falls in the range of 200 000 cases. A continuing incidence of endemic cases can be expected to emerge, but we believe it will occur at a low level.

(JAMA. 1990;263:1522-1525)

nomena controlled by forces that could be divined by scientific inquiry and expressed in mathematical terms. Surely, he was one of the earliest modern epidemic theorists—if not the first.

At the beginning of the 20th century, during the years in which Sir Ronald Ross^{1,2} (1908 and 1915) and William Hamer³ (1906) developed the basis of today's classic epidemic theories, John Brownlee, an inveterate Scottish public



Mechanistic Plausibility

HIV Model for “Test-and-Treat”

Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model

Reuben M Granich, Charles F Gilks, Christopher Dye, Kevin M De Cock, Brian G Williams

Lancet 2009; 373: 48–57
Published Online
November 26, 2008
DOI:10.1016/S0140-6736(08)61697-9

See Comment pages 7 and 9

Department of HIV/AIDS
(R M Granich MD,
Prof C Gilks DPhil,
Prof K M De Cock MD) and Stop
TB Department
(Prof C Dye DPhil,
B G Williams PhD), WHO,
Geneva, Switzerland

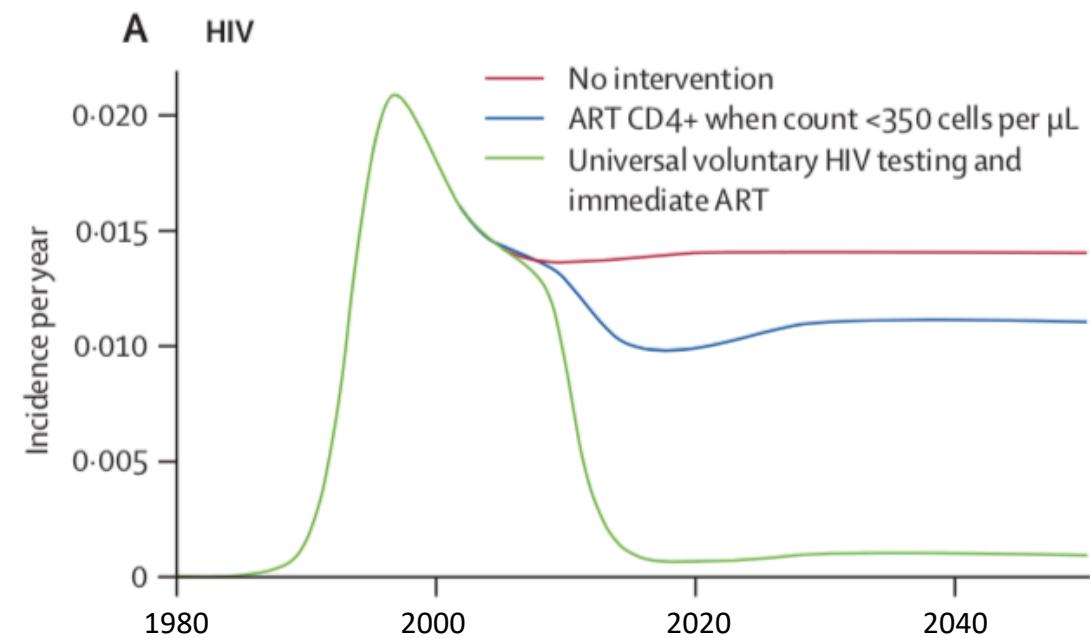
Correspondence to:
Reuben Granich, Antiretroviral
Treatment and HIV Care Unit,
Department of HIV/AIDS, WHO,
Avenue Appia 20, CH-1211,
Geneva 27, Switzerland
granichr@who.int

Background Roughly 3 million people worldwide were receiving antiretroviral therapy (ART) at the end of 2007, but an estimated 6·7 million were still in need of treatment and a further 2·7 million became infected with HIV in 2007. Prevention efforts might reduce HIV incidence but are unlikely to eliminate this disease. We investigated a theoretical strategy of universal voluntary HIV testing and immediate treatment with ART, and examined the conditions under which the HIV epidemic could be driven towards elimination.

Methods We used mathematical models to explore the effect on the case reproduction number (stochastic model) and long-term dynamics of the HIV epidemic (deterministic transmission model) of testing all people in our test-case community (aged 15 years and older) for HIV every year and starting people on ART immediately after they are diagnosed HIV positive. We used data from South Africa as the test case for a generalised epidemic, and assumed that all HIV transmission was heterosexual.

Findings The studied strategy could greatly accelerate the transition from the present endemic phase, in which most adults living with HIV are not receiving ART, to an elimination phase, in which most are on ART, within 5 years. It could reduce HIV incidence and mortality to less than one case per 1000 people per year by 2016, or within 10 years of full implementation of the strategy, and reduce the prevalence of HIV to less than 1% within 50 years. We estimate that in 2032, the yearly cost of the present strategy and the theoretical strategy would both be US\$1·7 billion; however, after this time, the cost of the present strategy would continue to increase whereas that of the theoretical strategy would decrease.

Interpretation Universal voluntary HIV testing and immediate ART, combined with present prevention approaches, could have a major effect on severe generalised HIV/AIDS epidemics. This approach merits further mathematical modelling, research, and broad consultation.

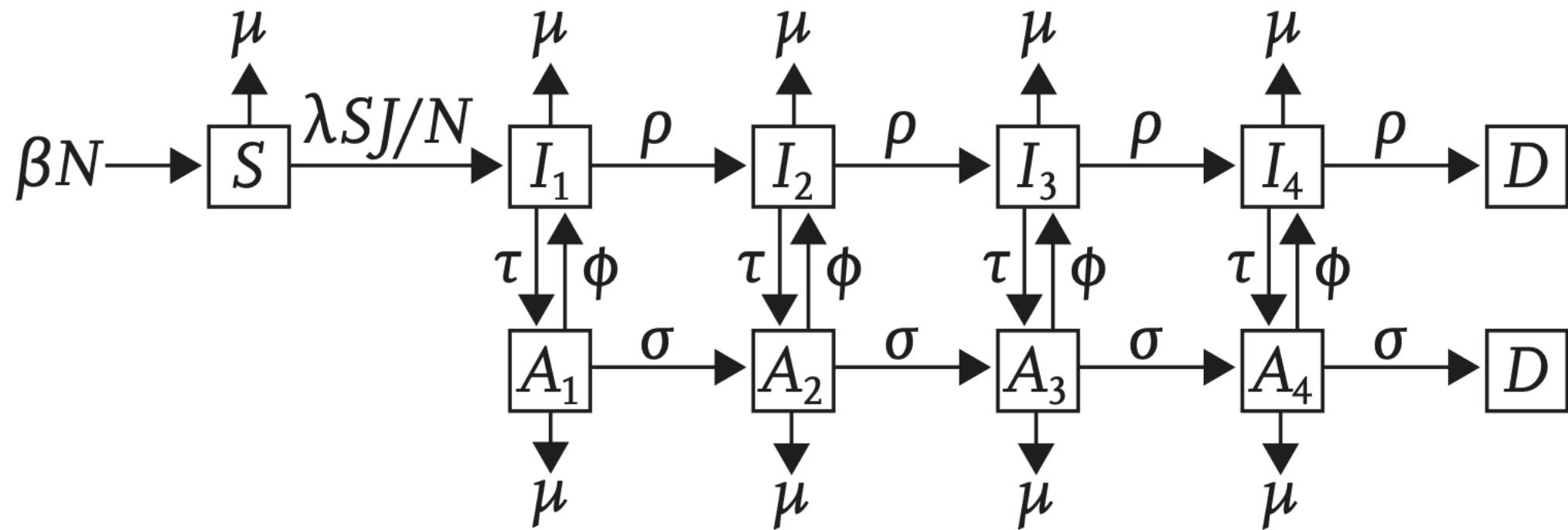


$$\lambda = \lambda_0 e^{-\alpha P n}$$

$$P = \frac{I}{N}$$

$$I = \sum_i (I_i + A_i)$$

$$J = \sum_i (I_i + \varepsilon A_i)$$



Projecting the Benefits of Antiretroviral Therapy for HIV Prevention: The Impact of Population Mobility and Linkage to Care

Jason R. Andrews,¹ Robin Wood,⁴ Linda-Gail Bekker,⁴ Keren Middelkoop,⁴ and Rochelle P. Walensky^{1,2,3}

¹Division of Infectious Diseases, Massachusetts General Hospital, ²Division of General Medicine, Department of Medicine, Massachusetts General Hospital, Boston, and ³Harvard University Center for AIDS Research, Harvard Institute for Global Health Cambridge, Massachusetts; ⁴Desmond Tutu HIV Centre, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Republic of South Africa

Background. Recent mathematical models suggested that frequent human immunodeficiency virus (HIV) testing with immediate initiation of antiretroviral therapy (ART) to individuals with a positive test result could profoundly curb transmission. The debate about ART as prevention has focused largely on parameter values. We aimed to evaluate structural assumptions regarding linkage to care and population mobility, which have received less attention.

Methods. We modified the linkage structure of published models of ART as prevention, such that individuals who decline initial testing or treatment do not link to care until late-stage HIV infection. We then added population mobility to the models. We populated the models with demographic, clinical, immigration, emigration, and linkage data from a South African township.

Results. In the refined linkage model, elimination of HIV transmission (defined as an incidence of <0.1%) did not occur by 30 years, even with optimistic assumptions about the linkage rate. Across a wide range of estimates, models were more sensitive to structural assumptions about linkage than to parameter values. Incorporating population mobility further attenuated the reduction in incidence conferred by ART as prevention.

Conclusions. Linkage to care and population mobility are critical features of ART-as-prevention models. Clinical trials should incorporate relevant data on linkage to care and migration to evaluate the impact of this strategy.



Community health workers with the Population Effects of Antiretroviral Therapy to Reduce HIV Transmission study did door-to-door HIV testing of 1 million people annually for 3 years. KIM CLOETE

Largest ever HIV prevention study delivers sobering message

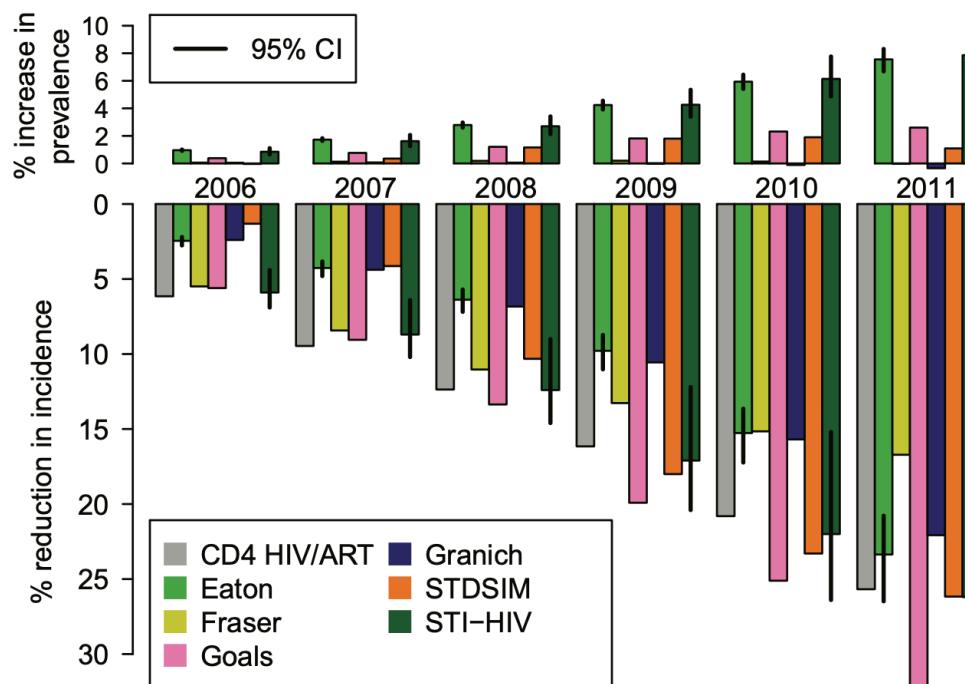
By [Jon Cohen](#) | Mar. 11, 2019 , 3:55 PM

Model Comparisons

HIV Treatment as Prevention: Systematic Comparison of Mathematical Models of the Potential Impact of Antiretroviral Therapy on HIV Incidence in South Africa

Jeffrey W. Eaton^{1*}, Leigh F. Johnson², Joshua A. Salomon³, Till Bärnighausen^{3,4}, Eran Bendavid⁵, Anna Bershteyn⁶, David E. Bloom³, Valentina Cambiano⁷, Christophe Fraser⁸, Jan A. C. Hontelez^{4,9,10}, Salal Humair^{3,11}, Daniel J. Klein⁶, Elisa F. Long¹², Andrew N. Phillips⁷, Carel Pretorius¹³, John Stover¹³, Edward A. Wenger⁶, Brian G. Williams¹⁴, Timothy B. Hallett¹

1 Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **2** Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa, **3** Harvard School of Public Health, Boston, Massachusetts, United States of America, **4** Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Mtubatuba, South Africa, **5** Department of Medicine, Stanford University, Stanford, California, United States of America, **6** Intellectual Ventures Laboratory, Bellevue, Washington, United States of America, **7** Research Department of Infection & Population Health, University College London, London, United Kingdom, **8** Medical Research Council Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom, **9** Erasmus University, Rotterdam, Netherlands, **10** Department of Primary and Community Care, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands, **11** School of Science and Engineering, Lahore University of Management Sciences, Lahore, Pakistan, **12** Yale University, New Haven, Connecticut, United States of America, **13** Futures Institute, Glastonbury, Connecticut, United States of America, **14** South African Centre for Epidemiological Modelling and Analysis, Stellenbosch, South Africa



Model Comparisons with Truth Sets

An open challenge to advance probabilistic forecasting for dengue epidemics

Michael A. Johansson^{a,b,1}, Karyn M. Apfeldorf^c, Scott Dobson^c, Jason Devita^c, Anna L. Buczak^d, Benjamin Baugher^d, Linda J. Moniz^d, Thomas Bagley^d, Steven M. Babin^d, Erhan Guven^d, Teresa K. Yamana^e, Jeffrey Shaman^e, Terry Moschou^f, Nick Lothian^f, Aaron Lane^f, Grant Osborne^f, Gao Jiang^g, Logan C. Brooks^h, David C. Farrow^h, Sangwon Hyunⁱ, Ryan J. Tibshirani^{h,i}, Roni Rosenfeld^h, Justin Lessler^j, Nicholas G. Reich^k, Derek A. T. Cummings^{l,m}, Stephen A. Lauer^k, Sean M. Moore^{n,o}, Hannah E. Clapham^p, Rachel Lowe^{q,r}, Trevor C. Bailey^s, Markel García-Díez^t, Marilia Sá Carvalho^u, Xavier Rodó^{r,v}, Tridip Sardar^w, Richard Paul^{x,y}, Evan L. Ray^z, Krzysztof Sakrejda^k, Alexandria C. Brown^k, Xi Meng^k, Osonde Osoba^{aa}, Raffaele Vardavas^{aa}, David Manheim^{bb}, Melinda Moore^{aa}, Dhananjai M. Rao^{cc}, Travis C. Porco^{dd}, Sarah Ackley^{dd}, Fengchen Liu^{dd}, Lee Worden^{dd}, Matteo Convertino^{ee}, Yang Liu^{ff}, Abraham Reddy^{ff}, Eloy Ortiz^{gg}, Jorge Rivero^{gg}, Humberto Brito^{gg,hh}, Alicia Juarrero^{gg,ii}, Leah R. Johnson^{jj}, Robert B. Gramacy^{kk}, Jeremy M. Cohen^{kk}, Erin A. Mordecai^{ll}, Courtney C. Murdock^{mm,nn}, Jason R. Rohr^{n,o}, Sadie J. Ryan^{m,oo,pp}, Anna M. Stewart-Ibarra^{qq}, Daniel P. Weikel^{rr}, Antarpreet Jutla^{ss}, Rakibul Khan^{ss}, Marissa Poultney^{ss}, Rita R. Colwell^{tt}, Brenda Rivera-García^{uu}, Christopher M. Barker^{vv}, Jesse E. Bell^{ww}, Matthew Biggerstaff^{xx}, David Swerdlow^{xx}, Luis Mier-y-Teran-Romero^{a,j}, Brett M. Forshey^{yy}, Juli Trtanj^{zz}, Jason Asher^{aaa}, Matt Clay^{aaa}, Harold S. Margolis^a, Andrew M. Hebbeler^{bbb,ccc}, Dylan George^{ccc,ddd}, and Jean-Paul Chretien^{ccc,eee}

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved September 30, 2019 (received for review June 18, 2019)

Dengue Forecasting Project

- Launched in 2015 by the White House Office of Science and Technology Policy
- Teams were provided 4 seasons of Dengue data in two locations
- They used these data, along with climate data (if they desired) to train their models, and submitted their models and performance on training sets
- They were then provided data for two additional seasons and asked to make out-of-sample predictions at 4-week intervals for two subsequent test seasons
- Asked to predict:
 - 1) Peak incidence for season
 - 2) Week in which peak incidence occurred
 - 3) Total cases occurring over the season

Types of Models (Mechanistic vs Statistical)

- *Team A.* The model is a dynamical two-strain susceptible-exposed-infectious-recovered-susceptible (SEIRS) compartmental model with a multi-life stage model for vector populations. Parameters were derived from the literature and data included dengue case data, precipitation, and minimum and maximum temperatures.
- *Team B.* Forecasts for each target-location were generated from an ensemble of three types of statistical models: Holt-Winters exponential smoothing (time series smoothing of historical dengue at local, seasonal, and long-term scales), multidimensional analogues (on historical dengue data and historical precipitation data), and historical average models (the seasonal distributions of historical cases). Ensemble components were assigned individual weights for each target-location pair based on mean absolute error in predictions over the previous 4 years
- *Team D.* The model used the K-spectral centroid clustering algorithm to generate normalized clusters of incidence patterns from similar seasons to a sliding window of normalized cases in the current season. The most similar curve was selected and scaled to project mean incidence for the rest of the season. Probabilities for each target were specified based on previous season error for the same targets.
- *Team P.* The model was a Bayesian statistical, time-series regression model including smoothed, lagged dengue case data, lagged climate variables (precipitation, minimum temperature, and relative humidity), and an indicator for serotype switches within the past two years. Cases were modeled as a negative binomial process with an offset for estimated population size.

Evaluating Model Performance

- Each model had to assign a probability to every outcome, p_i
- For example, suppose a model predicted that the most likely peak week of epidemic was week 23, and that it was very unlikely to be week 8, $p_{23} = 0.40$ and $p_8 = 0.01$.
- The score for each model, over n predictions:

$$S_n = \frac{1}{n} \sum_{i=1}^n \log(p_i)$$

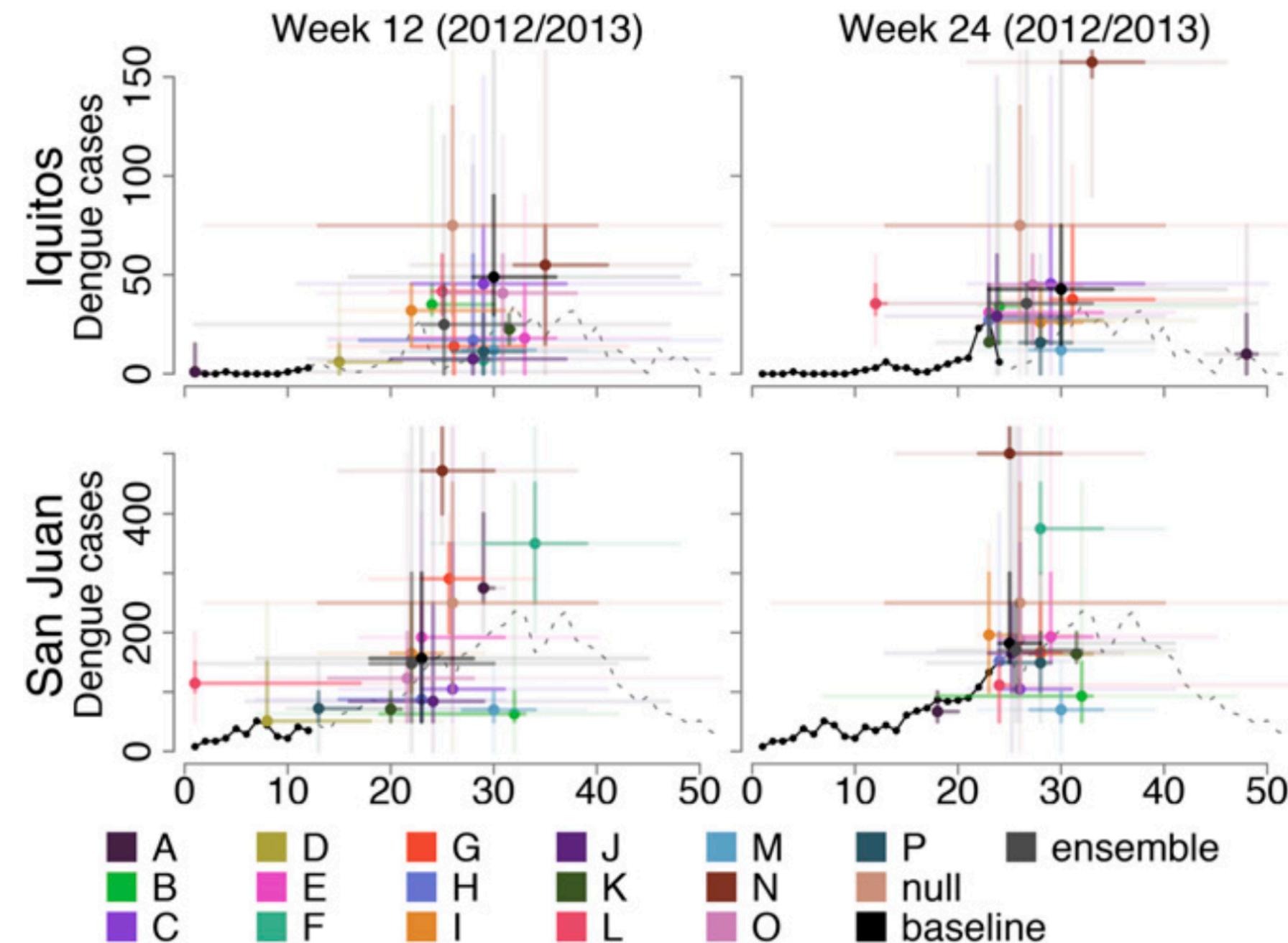


Table S1. Model characteristics and forecasting scores for Weeks 0-24 in the testing seasons (2009/2010 to 2012/2013). The highest score for each target is indicated in bold, with both the top team and the baseline indicated if the baseline outperformed all teams.

	Model Characteristics				San Juan (logarithmic scores)			Iquitos (logarithmic scores)		
Team	Mech. [‡]	Ensemble	Climate	Serotype	Peak Incidence	Peak Week	Season Incidence	Peak Incidence	Peak Week	Season Incidence
A	Yes	No	Yes	No	-4.62 [†]	-6.02 [†]	-4.79 [†]	-3.08 [†]	-6.07 [†]	-3.03 [†]
B	No	Yes	Yes	No	-2.57	-4.24	-2.04*	-1.85*	-6.38	-2.03*
C	Yes	Yes	No	No	-2.12*	-4.14	-2.13*	-2.08*	-3.36*	-2.47
D	No	No	No	No	-2.62	-6.45 [†]	-5.07	-4.76	-5.83	-5.43
E	No	Yes	Yes	No	-1.43*	-3.70*	-2.81	-2.54	-3.29*	-3.14
F	No	Yes	Yes	No	-2.99 [†]	-3.98 [†]	-3.98 [†]	-5.45 [†]	-4.66 [†]	-6.71 [†]
G	Yes	Yes	Yes	No	-1.23*	-4.88	-1.992*	-2.18*	-3.36*	-2.49
H	No	Yes	No	No	-2.60	-5.37	-2.64	-3.07	-3.45*	-2.54
I	Yes	No	Yes	No	-2.95 [†]	-6.03 [†]	-3.32 [†]	-4.18 [†]	-3.91* [†]	-4.02 [†]
J	No	Yes	Yes	No	-1.74*	-4.20	-1.986*	-2.27*	-3.61*	-2.66
K	No	No	Yes	No	-4.49 [†]	-6.17 [†]	-4.38 [†]	-6.68 [†]	-6.91 [†]	-6.91 [†]
L	Yes	Yes	Yes	No	-3.30 [†]	-5.45 [†]	-4.19 [†]	-5.29 [†]	-4.05 [†]	-4.61 [†]
M	No	No	No	No	-5.17	-5.75 [†]	-7.66	-5.78	-2.98*	-3.50
N	No	No	No	No	-4.98 [†]	-4.06	-2.18*	-3.88 [†]	-2.96*	-2.28*
O	No	No	No	No	-2.66	-3.90*	-2.84	-5.24	-4.35	-3.61
P	No	No	Yes	Yes	-2.63 [†]	-5.36 [†]	-4.55 [†]	-2.89 [†]	-3.08*	-4.81 [†]
null	NA	No	No	No	-2.40	-3.95	-2.40	-2.40	-3.95	-2.40
baseline	No	No	Yes	No	-1.43*	-3.47*	-2.15*	-2.88	-2.55*	-3.62 [†]
ensemble	Yes	Yes	Yes	Yes	-1.68*	-3.60*	-2.13*	-2.14*	-3.10*	-2.09*

[‡]Yes if the model included any mechanistic component.

[†]Forecasts with zero probability assigned to at least one observed outcome. Those individual forecast probabilities were changed to 0.001 to calculate the average.

*Forecasts with scores higher than the null model.

Mechanistic models had lower log scores (-0.65, 95% CI: -0.8 to -0.49) than purely statistical

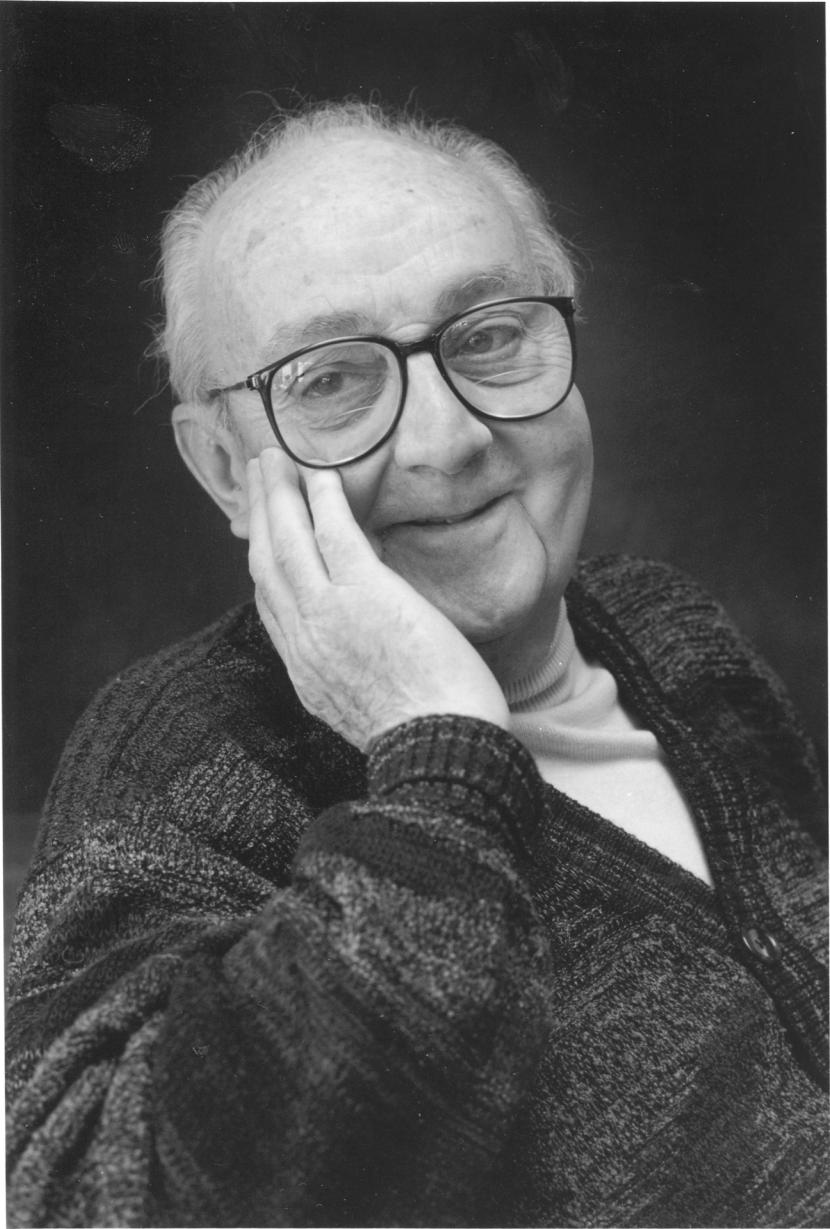
Ensemble models had higher scores (+1.02)

Models using climate had lower log scores (-0.14, 95% CI: -0.19 to -0.09)

Summary (to here)

- Models frequently err substantially
- Historical predictions often had uncertainty intervals not containing the subsequent observations
- Data often insufficient to identify changes in fundamental parameters or observation processes (e.g. case detection rates)
- Mechanistic models often underperform statistical or ML models in prediction, but are typically needed to project intervention impact

Should we even attempt to predict the impact of interventions on disease outcomes?



“All models are wrong,
but some are useful.”

-George Box

Decisions must be made

Models as Decision Guides

- Models as tools for making assumptions explicit
 - Transparency critical in methods and assumptions
- Start with simplicity, add complexity when clearly motivated
 - Model comparisons of simpler and more complex models
- Perform extensive sensitivity analyses to understand and convey what affects outcomes
- Ensure that uncertainty analyses fully capture uncertainty
- Interpretation and communication – emphasize robust results, sometimes qualitative insights over quantitative predictions
 - E.g. rank order of intervention impact may be far more robust than predictions of epidemic trajectories



Perspective

Wrong but Useful — What Covid-19 Epidemiologic Models Can and Cannot Tell Us

Inga Holmdahl, S.M., and Caroline Buckee, D.Phil.

Amid enormous uncertainty about the future of the Covid-19 pandemic, epidemiologic models are critical planning tools for policy-makers, clinicians, and public health practitioners.

Some models with apparently conflicting conclusions have received substantial press coverage, giving the impression that mathematical models are in general unreliable or inherently flawed. But infectious disease modeling is an expansive field with a long history, encompassing a range of methods and assumptions that are not necessarily directly comparable, or even designed for the same purpose (see box).

Covid-19 modeling studies generally follow one of two general approaches that we will refer to here as forecasting models and mechanistic models. Although there are hybrid approaches, these two model types tend to address different questions on different time scales, and they deal differently with uncertainty.

Forecasting models are often statistical in nature, fitting a line or curve to data and extrapolating from there — like seeing a pattern in a sequence of numbers and guessing the next number, without incorporating the process that produces the pattern. Well-constructed statistical frameworks can be used for short-term forecasts, using machine learning or regression, for example, to crunch epidemiologic data from the past or a different location and project

Five Questions to Ask about Model Results.

1. What is the purpose and time frame of this model? For example, is it a purely statistical model intended to provide short-term forecasts or a mechanistic model investigating future scenarios? These two types of models have different limitations.
2. What are the basic model assumptions? What is being assumed about immunity and asymptomatic transmission, for example? How are contact parameters included?
3. How is uncertainty being displayed? For statistical models, how are confidence intervals calculated and displayed? Uncertainty should increase as we move into the future. For mechanistic models, what parameters are being varied? Reliable modeling descriptions will usually include a table of parameter ranges — check to see whether those ranges make sense.
4. If the model is fitted to data, which data are used? Models fitted to confirmed Covid-19 cases are unlikely to be reliable. Models fitted to hospitalization or death data may be more reliable, but their reliability will depend on the setting.
5. Is the model general, or does it reflect a particular context? If the latter, is the spatial scale — national, regional, or local — appropriate for the modeling questions being asked and are the assumptions relevant for the setting? Population density will play an important role in determining model appropriateness, for example, and contact-rate parameters are likely to be context-specific.

“Unlike other scientific efforts, in which researchers continuously refine methods and collectively attempt to approach a truth about the world, epidemiologic models are often designed to help us systematically examine the implications of various assumptions about a highly nonlinear process that is hard to predict using only intuition. Models are constrained by what we know and what we assume, but used appropriately and with an understanding of these limitations, they can and should help guide us through this pandemic.”

Conclusions

- Historically, models have frequently erred substantially in predictions, and uncertainty intervals have often been too narrow
- Some of these issues are technical (e.g. fitting to data with correlated errors) or failing to account for changes in the epidemic (e.g. effective contact rate) or its observation (e.g. case detection rates), which can be very challenging
- When models are used for forecasting, the model fit to training data and mechanistic/biological plausibility may be insufficient properties for selecting models with high predictive accuracy
- Statistical models and ML often outperform mechanistic models in predictive accuracy, but typically cannot be used for predicting intervention impact
- In predicting impact of interventions, models are best viewed as tools for making assumptions explicit and generating insights beyond what intuition might afford us, but important to fully convey uncertainty and limitations

Thank you