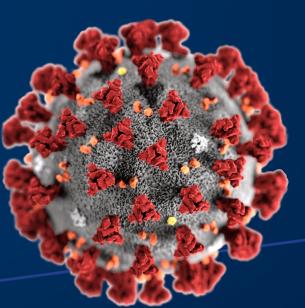
Statistics and Epidemics: COVID-19's lessons on myths and public health mobilization

Wednesday, December 9, 2020

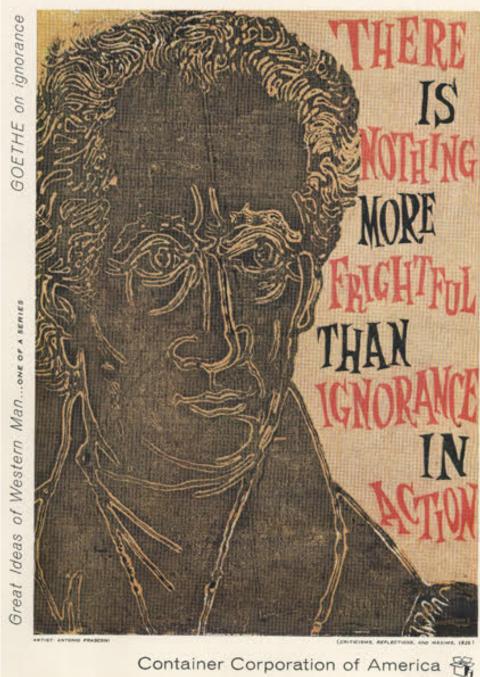
Nicholas P. Jewell Chair of Biostatistics & Epidemiology London School of Hygiene & Tropical Medicine, London, UK & UC Berkeley







Begin with humility . . .







2020...

- Surprise
- Shock
- Panic
- Disruption
- Sorrow
- Grief
- Fatigue
- Threats
- Feeling bad about not doing more COVID-19 work
- Feeling bad about not doing more non-COVID-19 work
- Empathy for
 - Teachers and students
 - HCWs
 - Essential workers
 - Farm workers
 - Parents
 - Children





Global Novel Coronavirus Cases and Deaths (as of Dec 8, 2020)

Worldwide Cases: ~68M (~15M in the US; ~1.8M in the UK) Worldwide Deaths: ~1.6M (~285K in the US; ~62K in the UK) Attack Rate:

Attack rate for 1918-19 influenza: about 33% (50M+ deaths worldwide; 675K in USA; 228K in UK) Attack rate for 2009 H1N1: about 10-20% (284K deaths worldwide; 12K in USA; 200 in UK)

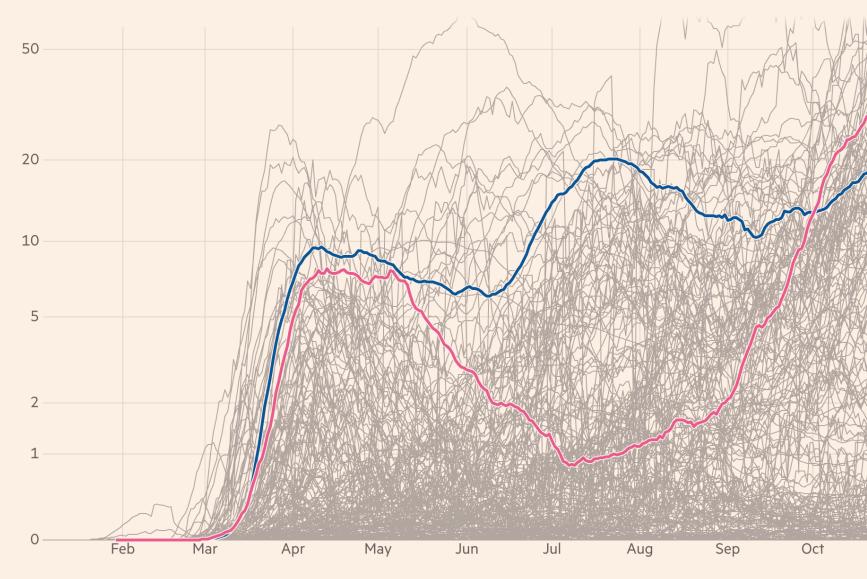




Cases Dec 8, 2020

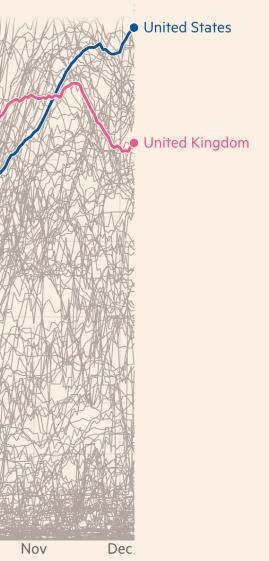
New confirmed cases of Covid-19 in United States and United Kingdom

Seven-day rolling average of new cases (per 100k)





Source: Financial Times analysis of data from the European Centre for Disease Prevention and Control, the Covid Tracking Project, the UK Government coronavirus dashboard, the Spanish Ministry of Health and the Swedish Public Health Agency. Data updated December 8 2020 1.37pm GMT. Interactive version: ft.com/covid19

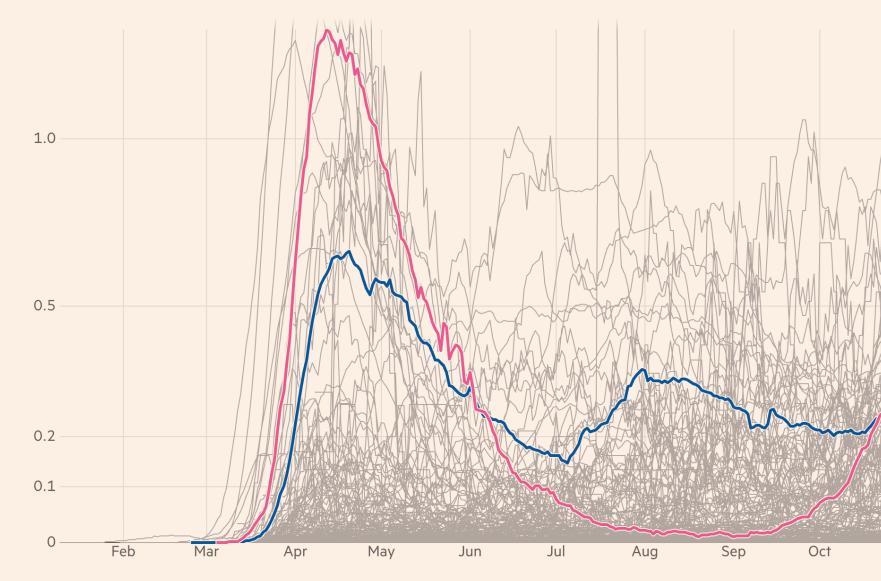


FINANCIAL TIMES

Deaths Dec 1, 2020

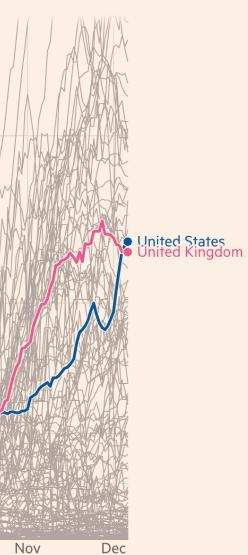
New deaths attributed to Covid-19 in United States and United Kingdom

Seven-day rolling average of new deaths (per 100k)





Source: Financial Times analysis of data from the European Centre for Disease Prevention and Control, the Covid Tracking Project, the UK Government coronavirus dashboard, the Spanish Ministry of Health and the Swedish Public Health Agency. Data updated December 8 2020 1.37pm GMT. Interactive version: ft.com/covid19



FINANCIAL TIMES

Myths about COVID-19

- US China travel ban in late January saved many lives it didn't
- The US/UK acted quickly we didn't
- Covid-19 is just like the seasonal 'flu it isn't
- We have tested adequately we haven't
- When will it be time to fully open up we need to thread a needle/vaccines will help now
- We have been learning as quickly as we can we haven't
- Statistical and modeling work will guide us we haven't always been that successful







Huge Progress . . .

- Enormous scientific advances in understanding SARS-CoV-2
- Impressive if belated response in March 2020
- Vaccine Development

... and Many Disappointments

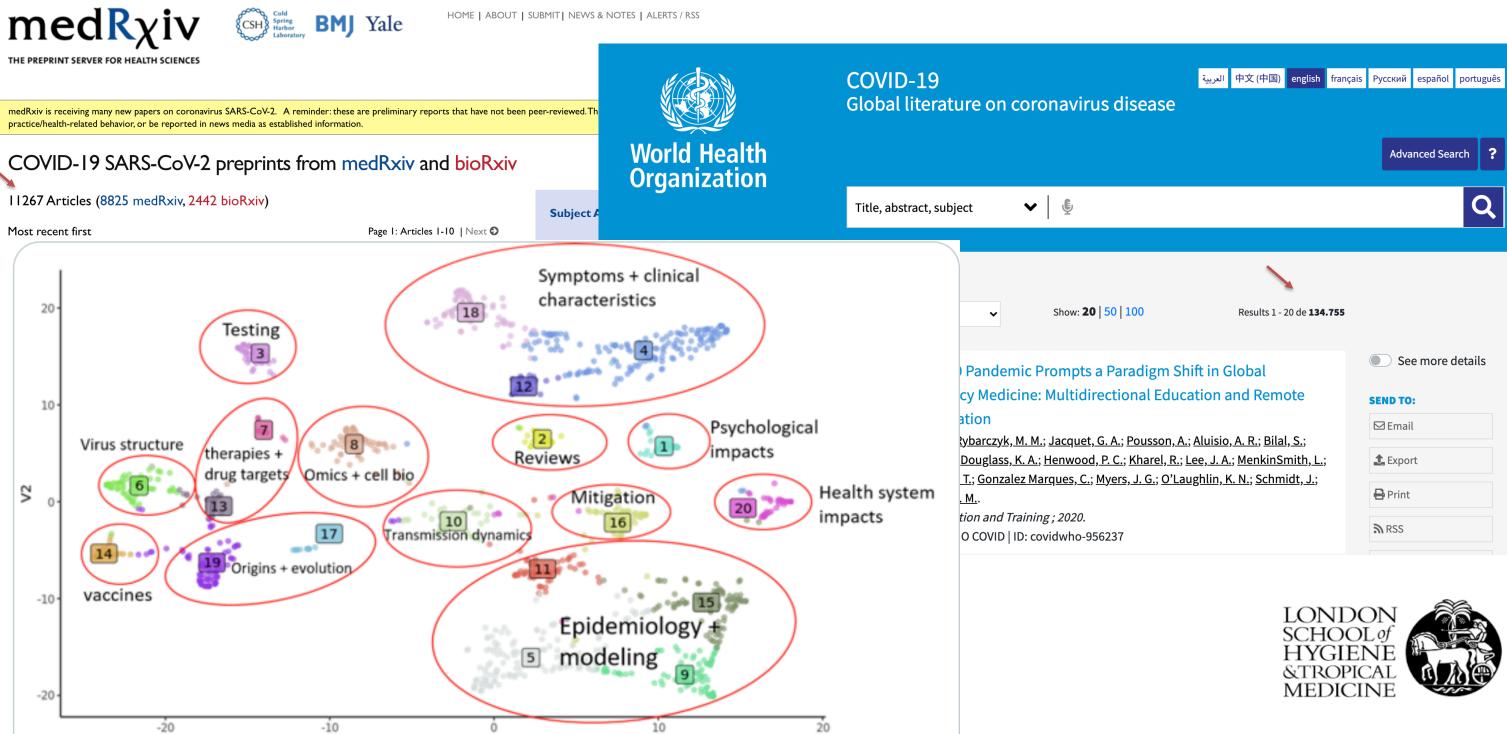
- Testing and Contact Tracing have not functioned well
- Uncoordinated data systems persist
- Slow ramp-up in data analysis
- Pandemic has cast light on underlying issues



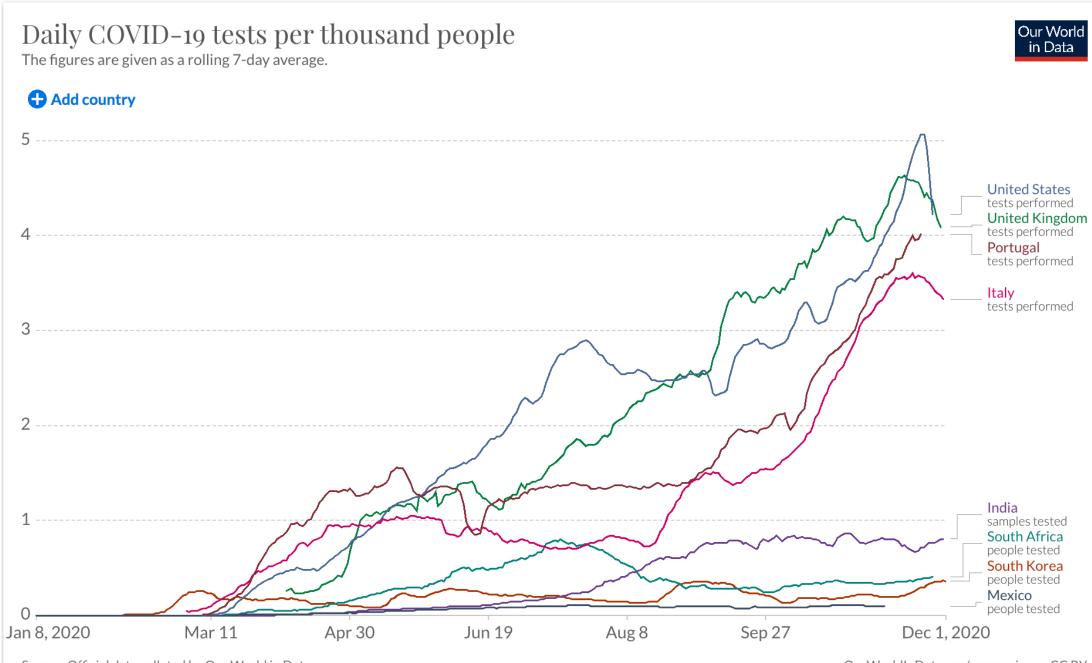


Publishing research

V1



Testing



Source: Official data collated by Our World in Data

OurWorldInData.org/coronavirus • CC BY Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Daily data is interpolated for countries not reporting testing data on a daily basis. Details can be found at our Testing Dataset page.



Testing

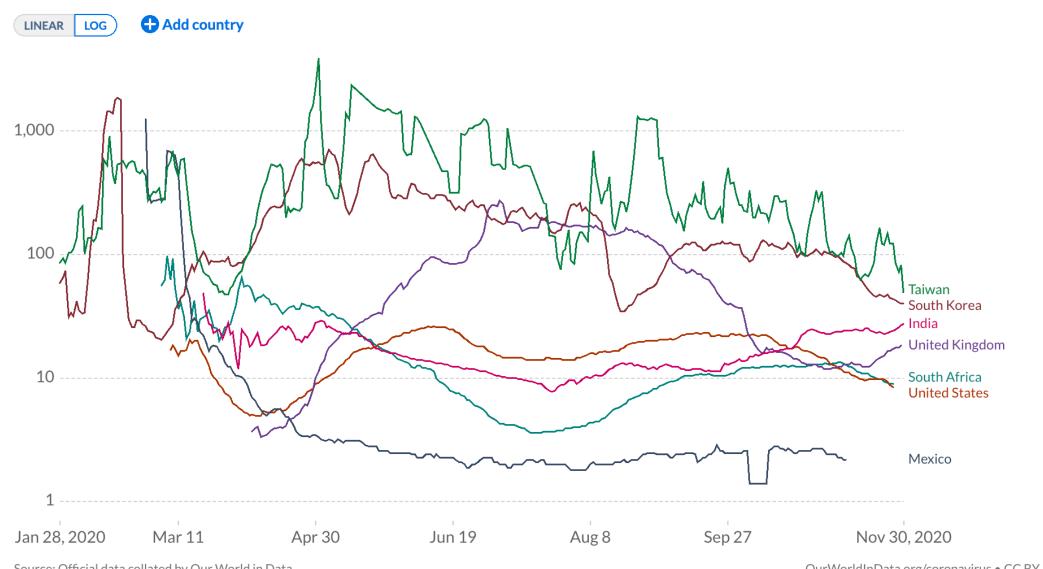
Test positive rate is the inverse of what is shown here

US is not reaching out far enough into many communities



Tests conducted per new confirmed case of COVID-19

Shown is the daily number of tests for each new confirmed case. This is a rolling 7-day average.



Source: Official data collated by Our World in Data

OurWorldInData.org/coronavirus • CC BY Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Daily data is interpolated for countries not reporting testing data on a daily basis. Details can be found at our Testing Dataset page.



SEIR Mathematical Models

"The tendency of some modelers to present them as scientific predictions of the future rather than models does not help. Models are widely used in government, and some models have arguably too much influence. They are generally most useful when they identify impacts of policy decisions which are not predictable by commonsense; the key is usually not that they are 'right', but that they provide an unexpected insight." (Chris Whitty, CMO England, 2015)

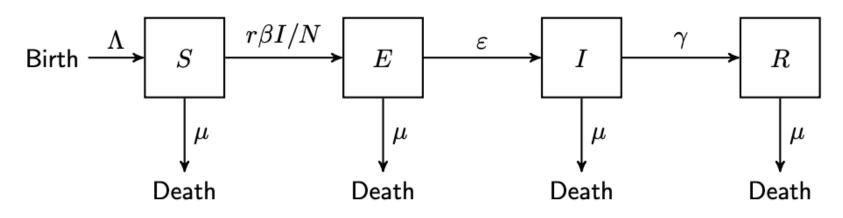
Models:

- Provide descriptions of the natural history of infections at a population/individual level
- Provide projections of the likely future (hard to predict peak) •
- Provide insight into the impact of possible interventions ('thought experiments' in hypothetical 'worlds')



SEIR Mathematical Models

Susceptible-Exposed-Infectious-Recovered Model: applicable to measles, mumps, rubella.



- Exposed (latent) humans E:
- Per-capita rate of progression to infectious state ε :

Also Agent-Based Modeling





 $\frac{\mathrm{d}S}{\mathrm{d}t} = \Lambda - r\beta S \frac{I}{N} - \mu S$ $\frac{\mathrm{d}E}{\mathrm{d}t} = r\beta S\frac{I}{N} - \varepsilon E$ $\frac{\mathrm{d}I}{\mathrm{d}t} = \varepsilon E - \gamma I - \mu I$ $\frac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - \mu R$

N = S + E + I + R.

SEIR Mathematical Models

$$R_0 = \begin{pmatrix} \mathsf{Number of} \\ \mathsf{contacts} \\ \mathsf{per unit time} \end{pmatrix} \begin{pmatrix} \mathsf{Probability of} \\ \mathsf{transmission} \\ \mathsf{per contact} \end{pmatrix} \begin{pmatrix} \mathsf{Duration of} \\ \mathsf{infection} \end{pmatrix} \\ \times \begin{pmatrix} \mathsf{Probabililty of} \\ \mathsf{surviving} \\ \mathsf{exposed stage} \end{pmatrix}$$

$$R_0 = r \times \beta \times \frac{1}{\gamma + \mu} \times \frac{\varepsilon}{\varepsilon + \mu}$$
$$= \frac{r\beta\varepsilon}{(\gamma + \mu)(\varepsilon + \mu)}$$

- If $R_0 < 1$, the disease-free equilibrium point is globally asymptotically stable and there is no endemic equilibrium point (the disease dies out).
- If R₀ > 1, the disease-free equilibrium point is unstable and a globally asymptotically stable endemic equilibrium point exists.

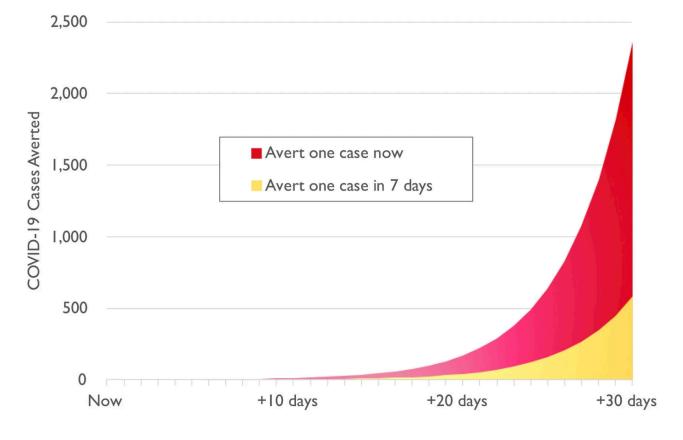


Maxim: Timing of Mitigation Measures is Crucial

The Exponential Power of Now

The explosive spread of coronavirus can be turned to our advantage, two infectious disease experts argue: "But only if we intervene early. That means now."

4 times as many infections averted in a month if we start today rather than a week from today



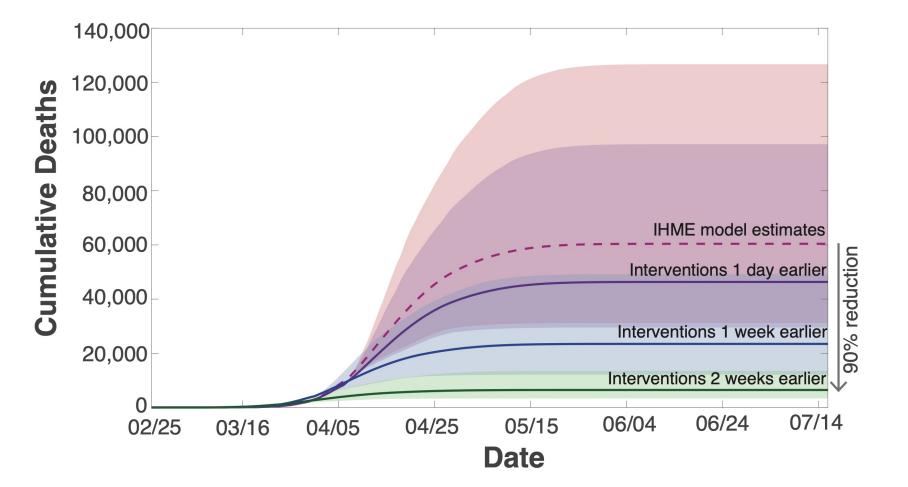




Timing of Mitigation Measures Looking Back

90% of US deaths averted in first wave if we had started full mitigation on March 2 instead of March 16

Note averted may only be postponed depending on second wave etc and how that is mitigated







What Have We Learned From Mathematical Models?

Models are not crystal balls

- Early models predicted areas where infection was likely widespread before large numbers of cases were observed
- Models have contributed to estimating (time-varying) reproductive numbers, case fatality ulletrates, and the scale of underreporting
- Models suggested early the extent of transmission prior to onset of any symptoms ullet
- Models suggested how long the virus had been circulating in a community ullet
- Models of heterogeneous spread provide insight into the extent of when herd immunity \bullet may have an impact on transmission
- Models assessed the impact of travel restrictions
- Models suggest complex interaction between the impacts of contact tracing, isolation and ulletquarantining, and social distancing (and social bubbles)



Calculation of the CFR or IFR?

- CFR—Case Fatality Rate = proportion of *confirmed cases* that do not survive
 - Denominator (# confirmed cases) influenced by testing patterns and frequency—likely an undercount
 - Confirmed cases are more likely to be symptomatic and therefore at higher risk of poor \bullet outcomes
 - Numerator (# COVID-19 deaths) is an undercount since (i) some COVID-19 infections will be hospitalized, not yet recovered, with still some risk of death; that is all unresolved infections are counted as recoveries, and (ii) reporting delays
- IFR—Infection Fatality Rate = proportion of *infected* that do not survive
 - Denominator mot reported correctly in ascertainment systems—have to rely on additional seroprevalence surveys
 - Numerator suffers from same problems as for IFR \bullet

Statistical methods and models have been developed to tackle these issues



Both have issues if deaths from COVID-19

The IHME "Model"

- Quality of fatality reporting: underreporting and delayed reporting (most death counts refer only to deaths in hospital)
- Farr's Law (epidemic case counts follow a bell-shaped curve) and fallacy (eg HIV)
- Consistent of mortality curves across regions
- Assumption of same effects of social distancing everywhere
- Estimation of uncertainty
- Volatility of projections day-to-day

As of Dec 8, the current estimate for cumulative US deaths is 538,893 (by Apr 1, 2021) but as high as 770,234 if mandates are eased, or as low as 472,679 with universal mask wearing



IHME updated their model on May 4: US estimate for deaths immediately went from 72,433 to 134,475

This is when IHME started to incorporate SEIR modelling

The IHME "Model"

Annals of Internal Medicine[®]

LATEST ISSUES CHANNELS CME/MOC IN THE CLINIC **JOURNAL CLUB** WEB EXCLUSIVES **AUTHOR INFO**

IDEAS AND OPINIONS | 14 APRIL 2020

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; Britta L. Jewell, PhD

Article, Author, and Disclosure Information

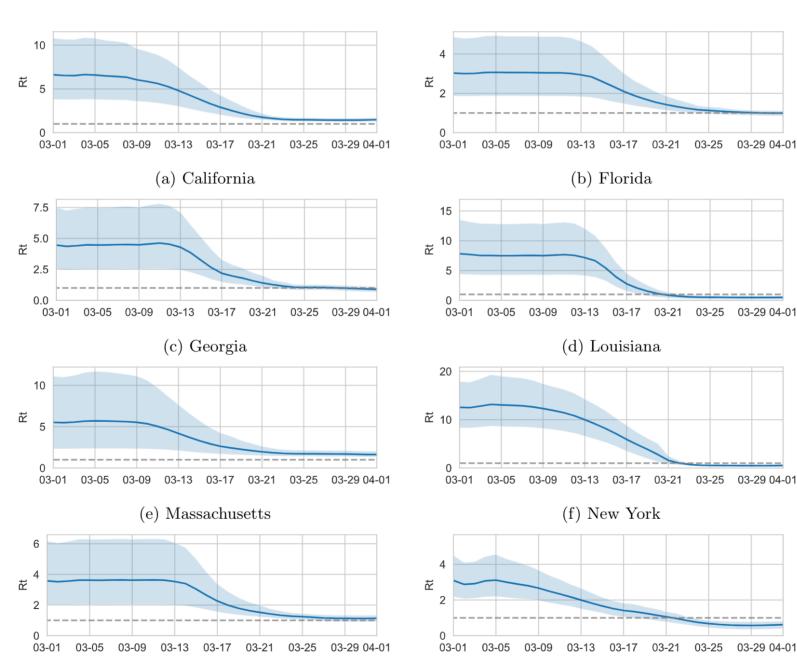


A recent modeling analysis by the Institute for Health Metrics and Evaluation (IHME) (1) projecting deaths due to coronavirus disease 2019 (COVID-19) has attracted considerable attention, including from the U.S. government (2). The model used COVID-19 mortality projections to estimate hospital bed requirements and deaths. We agree with qualitative conclusions that demand for hospital beds may exceed capacity and efforts to enhance mitigation policies and surge planning are essential. Data endorse shelter-in-place orders and suggest that these measures must remain while awaiting advances in surveillance, treatment, and vaccines.



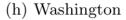
Mobility Data Analysis by State

Apple mobility data: work of Emily Fox, Carlos Guestrin, Andy Miller, Nick Foti, Joseph Lewnard and Nick Jewell

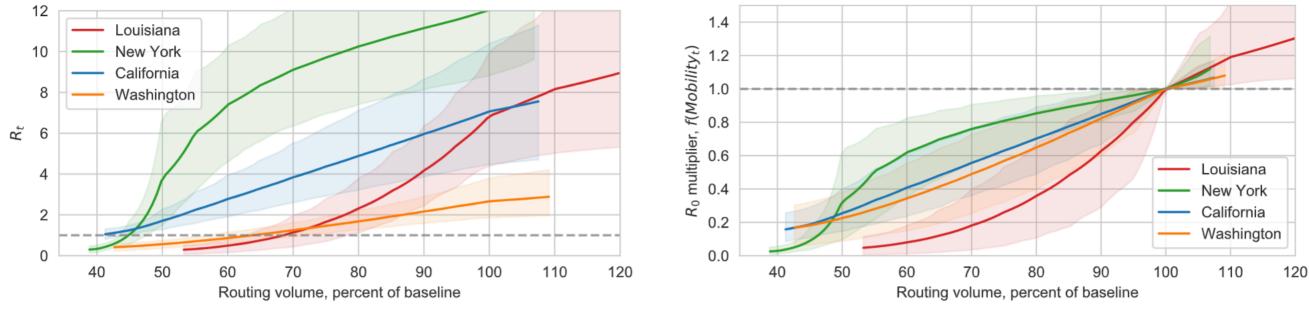


(g) Texas





Mobility Data Analysis by State



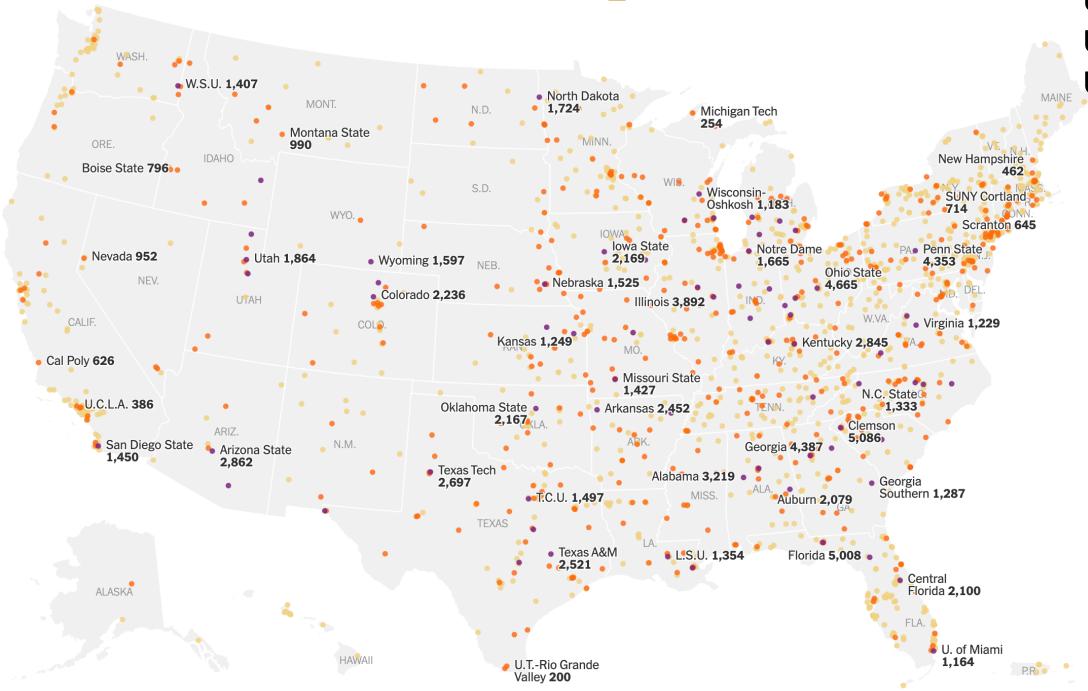
(a) Effective R vs. mobility

(b) R_0 multiplier vs. mobility

Apple mobility data: work of Emily Fox, Carlos Guestrin, Andy Miller, Nick Foti, Joseph Lewnard and Nick Jewell



Universities and Colleges



U. of Michigan (2,545)
U. of Wisconsin (4,193)
U. of N. Carolina (1,274)
U. of Illinois (3,892)

As of Nov. 19, 2020



University of California, Berkeley

Semester Name

(AII)

Date

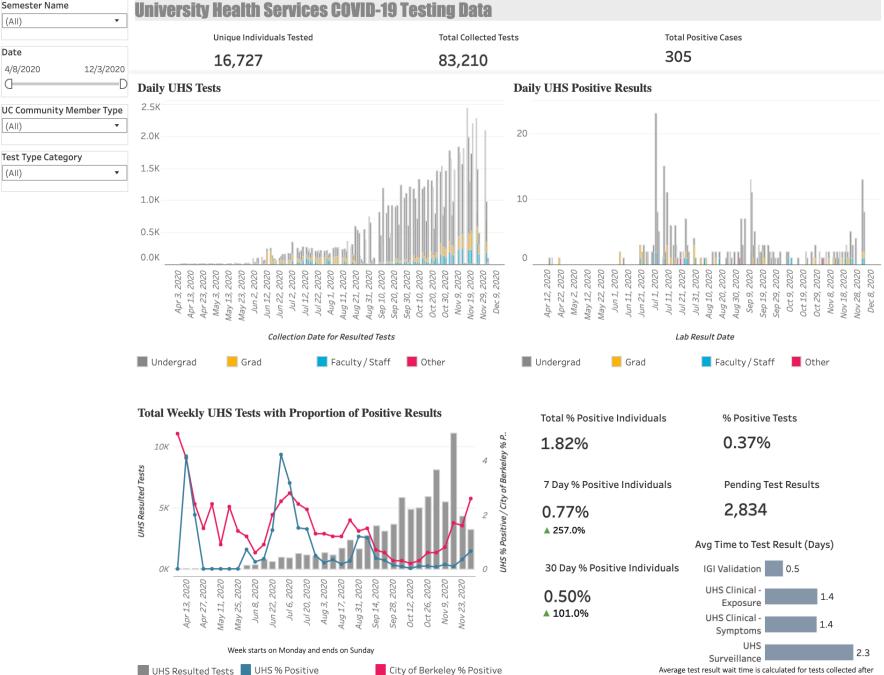
Ð

(AII)

(AII)

4/8/2020

- Classes
- Dorms
- Bathrooms
- Testing
- **Contact tracing** lacksquare
- Quarantining & isolation ullet
- Sewage testing ${}^{\bullet}$
- Masks
- Ventilation
- Signage
- Education



*Data includes PCR tests performed at UHS only. Indeterminate results have been excluded from analysis. Data include both symptomatic and surveillance testing. Positive test results reported by result date. Tests are typically resulted within 3 days of collection. Filters do not apply to pending result counts. Filters return tests resulted within selected time period for selected populations and test types. Positivity rates are based on date resulted. Weeks start on Sunday. Spring 20 is equal to 4/1/20-5/15/20, Summer 20 is equal to 5/16/20-8/18/20, Fall 20 is equal to 8/19/20-12/31/20. Data is refreshed daily.



10/31/2020

Average test result wait time is calculated for tests collected after

COVID-19 and the "Economy"

- All economies are interconnected and are being badly hit
- In history, pandemics have had a significant impact on human society and governments, \bullet perhaps even more so than wars (although the two often go hand in hand as wars have a tendency to promote epidemics and vice versa) (see *Pandemics and the Shape of* Human History, Elizabeth Kolbert, The New Yorker, Mar 30, 2020, or for a longer treatment *Epidemics and Society: From the Black Death to the Present*, Frank Snowden, Yale University Press, 2019)
 - The Justinianic plague and the Roman Empire
 - New World pandemic(s) and eradication of populations and development of the slave trade



The 1918-19 influenza pandemic and WWI (and WWII?)

COVID-19 and the "Economy"

- Clearly, public health the economy, but the reverse causal relationship is much less well understood although a poor economy has many significant impacts on other public health issues during a pandemic
- Opening up the economy *m* increased risk of COVID-19 infection
- No serious economic models that interlink with (SEIR) models of transmission dynamics
- In different communities, what are the "sweet spots" of maximizing economic activity (and minimizing unemployment) that do not raise community transmission above levels that mitigation measures can be effective?

Real risk of the economic slowdown is not the short-term effects but longer terms effects of inflation and political instability



Natural Experiments and Learning

- Clearly, many natural experiments are occurring regarding school openings, mitigation ۲ measures, etc
 - How do we link the outcomes of these natural experiments to outcomes
 - What are the data systems that should be set up?
- These natural experiments almost always involve group interventions with infectious ٠ disease outcomes also measured at the group level
- Not enough attempts to learn from these natural experiments, or to randomize (other ٠ than for vaccine and treatment trials, of course) *Experiment Aversion*
 - Randomized experiment of opening gyms in Oslo (Helsingen *et al*, medR χ iv)
 - Attempted randomized school re-openings in Norway (not implemented)
 - Which Interventions Work Best in a Pandemic (Haushofer and Metcalf, Science 5) \bullet Jun 2020)
 - Danish Mask Study





Randomization and Natural Experiment Designs

- We have significant *equipoise* (costs, benefits and timing)
- Stepped wedge designs that rollout interventions over time
 - Complex time dynamics, many NPIs may wane in efficacy over time ullet
- Spillover effects
- Data collection systems for outcomes
 - Infection counts, hospital admissions etc measure at end of intervention? ullet
 - Measurements of economic and/or psychological costs--when and how? \bullet





Incidence/Seroprevalence Surveys

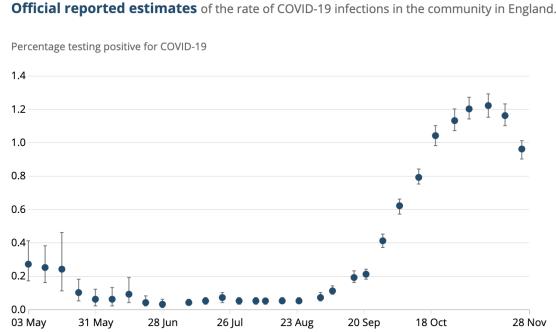
- Consistent methods of assessing scale and scope of infections—the cornerstone of ۲ Epidemiology
 - Diagnostic testing (who is getting tested and why)
 - Surveillance testing
 - Seroprevalence surveys
 - Hospitalizations
 - Deaths—and excess mortality •
- Testing in surveillance studies plays a different role from its use as a control measure ۲
- **Statistical Issues** ۲
 - **Reporting delays**
 - Selection bias (what is the provenance of the data?) •
 - Measurement error (even or especially in death certificates)



Who is getting infected and when? Where are the fires, and how fast are they spreading?

National Studies: UK

- Coronavirus (COVID-19) Infection Survey UK—Office for National Statistics
- Fully launched 9 October, 2020
- Starting with ~ 21,000 individuals from a national probability sample
- Participants provide nasal and throat swabs
- About 10% of the sample will ultimately be asked to provide blood samples for antibody testing

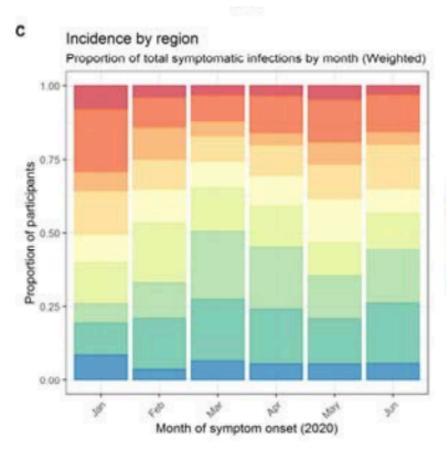


Source: Office for National Statistics - Coronavirus (COVID-19) Infection Survey



National Studies: UK

- REACT-1: Imperial College
- Monthly results since May 2020
- ~ 150,000 participants from a national probability sample each month
- Participants provide nasal and throat swabs
- PCR testing





Region

North East North West Yorkshire and The Humber East Midlands East of England London South East South West

National Studies: UK

- **REACT-2:** Imperial College
- National probability sample of ~100,000 adults
- Antibody tests using fingerprick samples

Table shows results from June-July 2020

Table 1: Prevalence of SARS-CoV-2 IgG antibodies: crude, adjusted for test performance, and weighted to the population (18+ years) of England, by sociodemographic characteristics

	Total	Total tests	Crude prevalence
	antibody	(with valid	%
	positive	results)	[95% confidence
			intervals]
England	5544	99908	5.6 [5.4-5.7]
Sex			
Male	2405	43825	5.5 [5.3-5.7]
Female	3139	56083	5.6 [5.4-5.8]
Age			
18-24	463	6499	7.1 [6.5-7.8]
25-34	930	13366	7.0 [6.5-7.4]
35-44	964	17052	5.7 [5.3-6.0]
45-54	1255	20634	6.1 [5.8-6.4]
55-64	1131	20404	5.5 [5.2-5.9]
65-74	568	15543	3.7 [3.4-4.0]
75+	233	6410	3.6 [3.2-4.1]
Ethnicity			
White	4827	92737	5.2 [5.1-5.3]
Mixed	106	1347	7.9 [6.5-9.4]
Asian ²	369	3658	10.1 [9.2-11.1]
Black ³	135	900	15.0 [12.8-17.5]
Other	79	762	10.4 [8.4-12.7]



Prevalence %	Weigh ted ¹	
adjusted for test	prevalence %	
[95% confidence	[95% confidence	
intervals]	intervals]	
5.0 [4.8-5.2]	6.0 [5.8-6.1]	
4.9 [4.7-5.2]	6.2 [5.9-6.4]	
5.1 [4.8-5.3]	5.8 [5.5-6.0]	
6.9 [6.2-7.7]	7.9 [7.3-8.5]	
6.7 [6.2-7.2]	7.8 [7.4-8.3]	
5.1 [4.7-5.6]	6.1 [5.7-6.6]	
5.6 [5.3-6.0]	6.4 [6.0-6.9]	
5.0 [4.6-5.4]	5.9 [5.5-6.4]	
2.7 [2.4-3.1]	3.2 [2.8-3.6]	
2.7 [2.2-3.3]	3.3 [2.9-3.8]	
4.6 [4.4-4.8]	5.0 [4.8-5.2]	
7.8 [6.2-9.7]	8.9 [7.1-11.1]	
10.5 [9.3-11.7]	11.9 [11.0-12.8]	
16.4 [13.8-19.4]	17.3 [15.8-19.0]	
10.8 [8.4-13.7]	12.3 [10.2-14.7]	

San Francisco Bay Area PCR/Serological Survey

The overall goals of this study are to:

1) better understand the short- and long-term prevalence and the spread of COVID-19 in Bay Area communities in order to inform virus-containment measures, assess epidemiologic characteristics of the virus including the probability of symptomatic or asymptomatic infection, and create predictive models of COVID-19 spread in the region 2) examine the participant characteristics (e.g., age, sex, underlying medical conditions), household, genetic, environmental, and viral factors that may affect the risk of infection and/or modify the manifestation of symptoms and other outcomes; 3) assess the sensitivity and specificity of fingerprick whole blood in commercially available rapid serological tests, and filter paper blood spots vs. serum from venous blood draw

tested by ELISA;

4) assess immune protection against COVID-19 associated with responses to previous infection with SARS-CoV-2 and/or endemic seasonal coronaviruses.



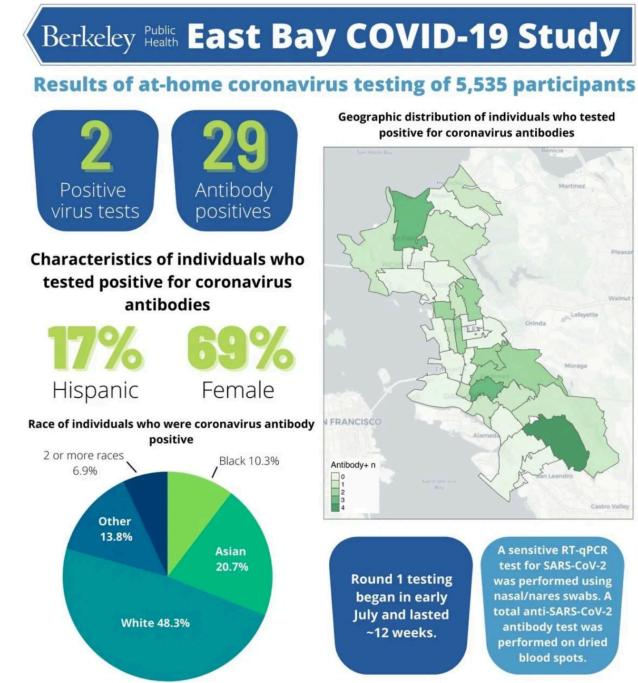


San Francisco Bay Area PCR/Serological Survey

- Not probability sample •
- Designed to provide estimates for ethnic subgroups
- Designed to assess longitudinal antibody evolution
- Preliminary results show about 2% seropositivity

Lisa Barcellos, Eva Harris, Joe Lewnard, Nick Jewell







Farmworkers in Monterey County, CA

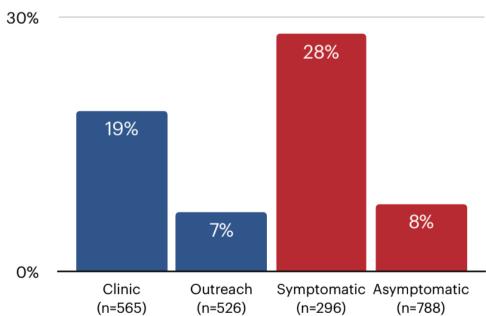
- Clinics, Community Testing events
 Household Complexes
- ~1,000 farmworker participants
- Throat swab and blood samples
- Not probability sample
- 13% tested positive at recruitment from July—November, 2020

Brenda Eskenazi, Ana Maria Mora, Joe Lewnard, Max Cuevas





Farmworkers in Monterey County, CA



Prevalence of infection: PCR

ANTIBODY TEST RESULTS

For this preliminary report, we have received antibody results from those tested up through the end of October (n=740). Antibody results from farmworkers who tested negative for active infection revealed that antibody prevalence has increased over time from 9.1% in July, to 12.5% in August, 20.2% in September, and 19.4% in October.



Other Issues

- Herd Immunity Threshold \bullet
 - Definition (depends on R_0)
 - final size is greater than herd immunity threshold
 - Threshold will vary due to other mitigation measures •
- National Epidemiological Modeling Strategy (see recent report OPCAST report for US) \bullet
- Testing—pooled testing ۲
- **Bi-directional contact tracing** ۲
- **Excess deaths**
- Causal Inference methods?
- Vaccines
 - Distribution •
 - Ongoing assessment of efficacy (test-negative design?)
 - Ongoing assessment of adverse events



Causal Inference?

Does Contact Tracing Work? Quasi-Experimental Evidence from an Excel Error in England*

Thiemo Fetzer[†] Thomas Graeber[‡]

November 24, 2020

Abstract

Does contact tracing work? **Quasi-experimental** evidence from an **Excel error in** England

CAGE working paper no. 521

November 2020

Thiemo Fetzer

Thomas Graeber

Contact tracing has been a central pillar of the public health response to the COVID-19 pandemic. Yet, contact tracing measures face substantive challenges in practice and well-identified evidence about their effectiveness remains scarce. This paper exploits quasi-random variation in COVID-19 contact tracing. Between September 25 and October 2, 2020, a total of 15,841 COVID-19 cases in England (around 15 to 20% of all cases) were not immediately referred to the contact tracing system due to a data processing error. Case information was truncated from an Excel spreadsheet after the row limit had been reached, which was discovered on October 3. There is substantial variation in the degree to which different parts of England areas were exposed - by chance - to delayed referrals of COVID-19 cases to to the contact tracing system. We show that more affected areas subsequently experienced a drastic rise in new COVID-19 infections and deaths alongside an increase in the positivity rate and the number of test performed, as well as a decline in the performance of the contact tracing system. Conservative estimates suggest that the failure of timely contact tracing due to the data glitch is associated with more than 125,000 additional infections and over 1,500 additional COVID-19related deaths. Our findings provide strong quasi-experimental evidence for the effectiveness of contact tracing.

Summary

- Another newly emergent infectious disease (there will be more and this is not the 'big one')
- A continually evolving epidemiologic situation and response
- Gaps remain in our global capacity to prepare for, predict, detect, and respond to newly emerging infections **m** invest in public health infrastructure!
- Mathematical models and statistical analysis of emerging data have crucial roles to play but *significant* effort and caution is needed in interpretation and

dissemination



End with hope . . .

WHERE HOPE COMES FROM

It comes from heartache. And it grows like the lone sapling from the ashes of loss. And it carves its way out of the heart of tragedy and its heavy cost. And it rises like a soldier thought lost returning home to his mother. And it smiles like the calm, clear sky following weeks of one storm after another.

And maybe this is why when Pandora opened the box that carried such calamities which inflicted all of mankind, gentle hope emerged from it too.

> What else helps us overcome suffering if not by giving hope a chance to bloom.

> > Nikita Gill



