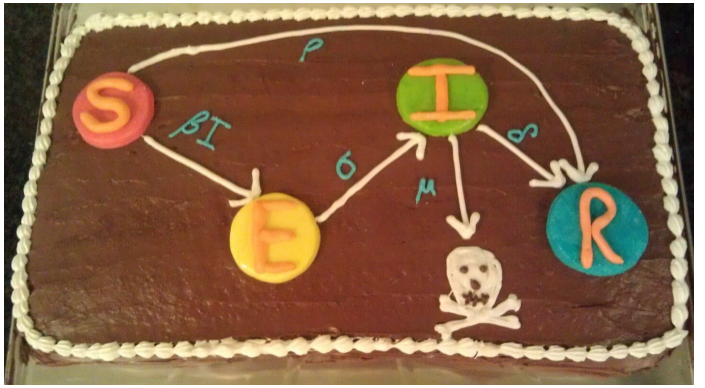
The Surprising Power of Napkin Math

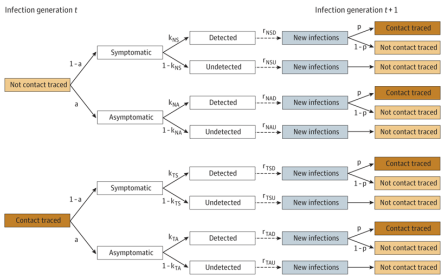
Why Simple Models Matter (More) in 2024

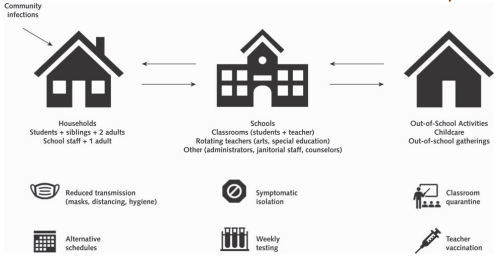
Alyssa Bilinski, PhD
Brown University

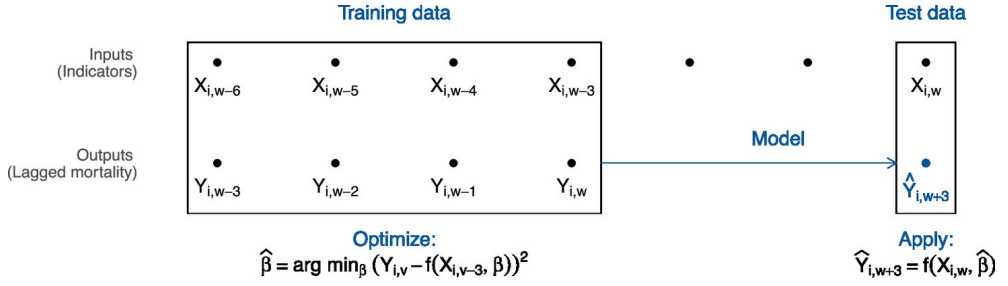
Modeling for Policy
November 2024

Models are powerful tools for prediction and policy.









Direct policy engagement!
(local, state, federal, international)

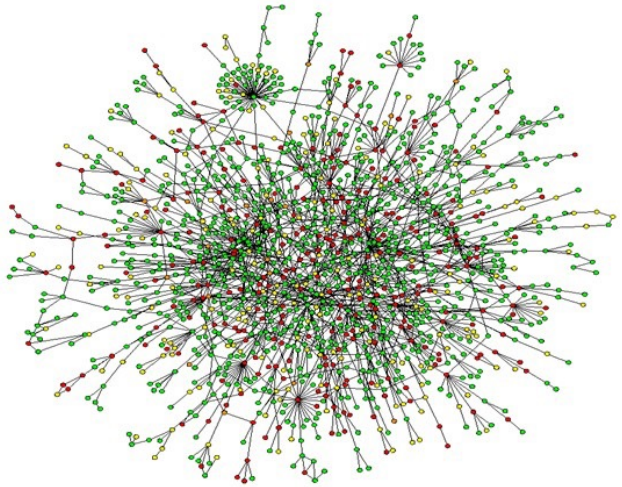
JAMA NO 2020, 2022; Annals of IM 2021, 2022; JAMA Peds 2022; PNAS 2023

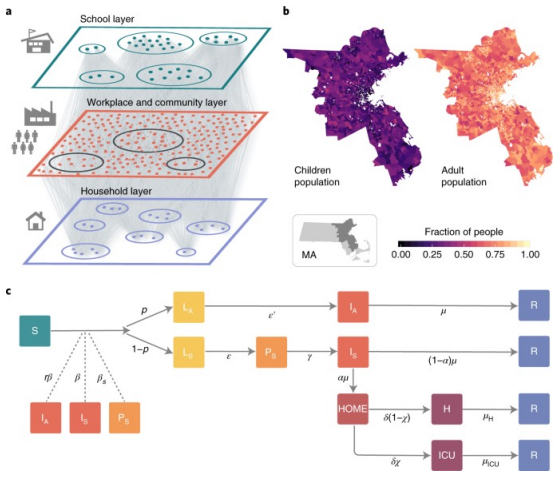
My takeaways

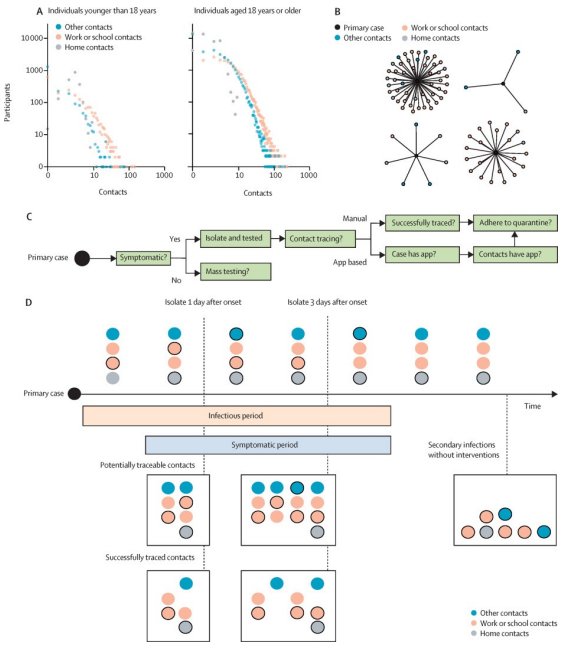
1. **Model with policymakers.** Why stop after COVID-19?

2. **Simple models are both incredibly powerful and often undervalued.**

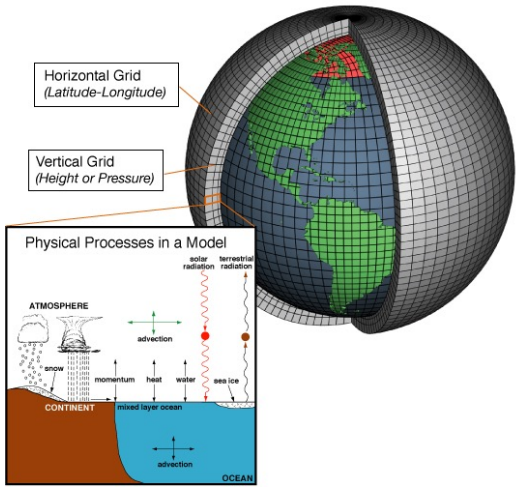
Modern models are often complex.







Modern models are often complex.



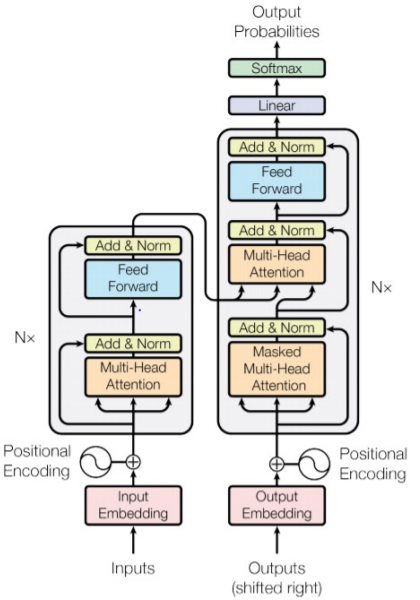
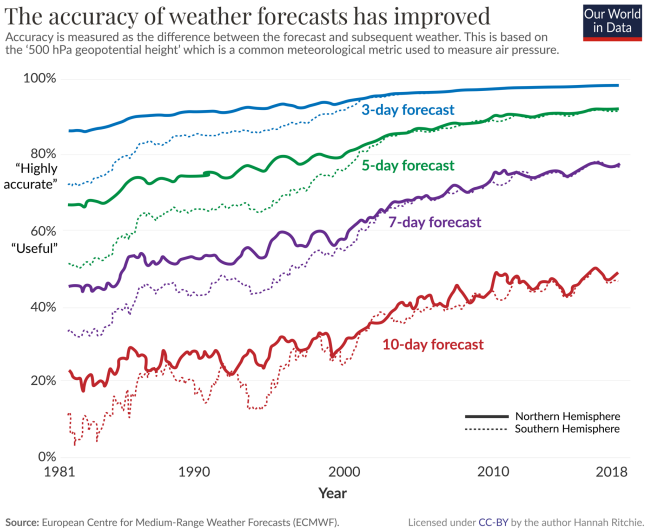
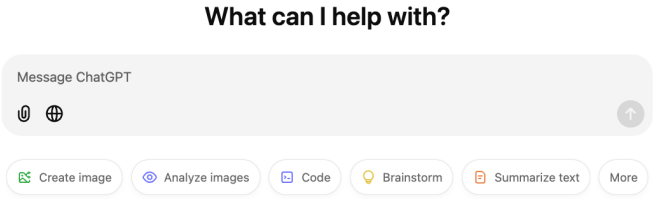


Figure 1: The Transformer - model architecture.

With stunning results…





But complexity comes with costs for modelers…

1. **Production time.** Even with high-performance computing, complex models take time to develop and run.

2. **Data collection.** Granular models require granular data.

3. **Interpretability.** What exactly is my model doing?
   → catching typos and thinkos
   → making policy recommendations

…and these costs matter for policymakers.

1. **Production time.**
2. **Data collection.**

→ **Slower availability of modeling evidence.**

3. **Interpretability.** →

Trust
Transportability (not forecasts)
→ context-specific expertise
Value judgments
→ asymmetric error costs
Understanding divergent models

Complex models can close conversations.

At the same time…

I realized that skilled colleagues – both researchers and policymakers – drew heavily on the ability to **rapidly generate, manipulate, and explain simple models** – with **actionable results**.

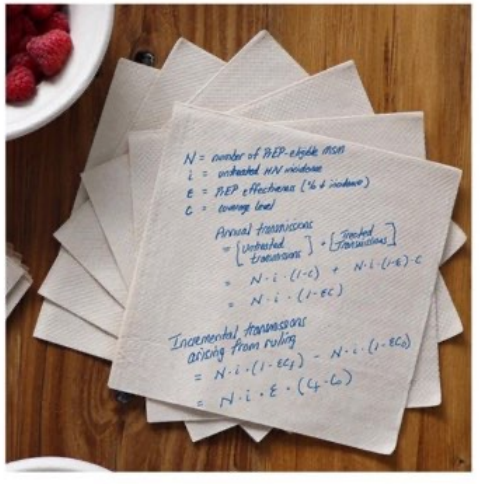
But in contrast to the many classes available to learn about complex simulations and machine learning, I had never seen this formally taught.

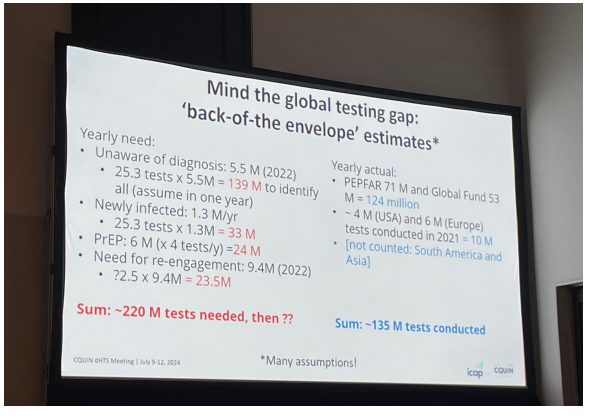
In fact, we didn't even have **vocabulary** for it.

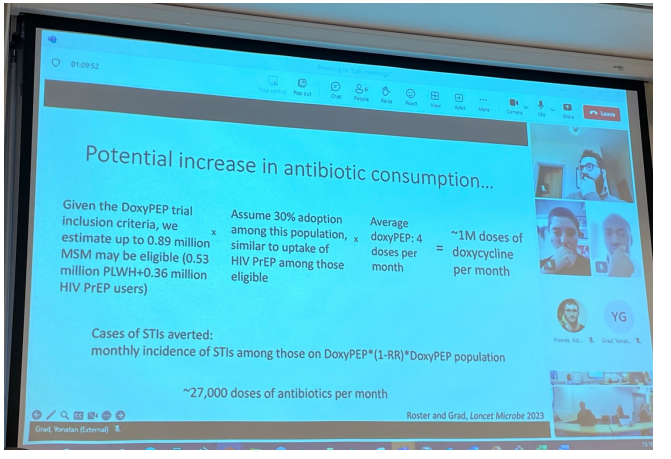
Enter napkin math.

What is napkin math?

Napkin math is an approach to analytic thinking that starts with the **simplest model**, adding complexity as needed to inform decision-making.







Though the math is simple (mainly **multiplication** and **division**, no cluster computing here), the process may not be.

Why use napkin math?

1. For many policy questions, **we may not need a complex model**. Napkin math can get us the answer with sufficient precision **to make a decision.**

2. Even when complex models are needed, **napkin math can help** us to *develop*, *check*, *interpret*, and *compare* models and understand common *prediction errors*.

Course

- >300 participants
- Techniques
- Examples
- Practice

Research and Practice

Collaborators




Josh Salomon (Stanford)



Jeff Imai-Eaton (Harvard)


Meagan Fitzpatrick (UMD)


David Paltiel (Yale)

Today

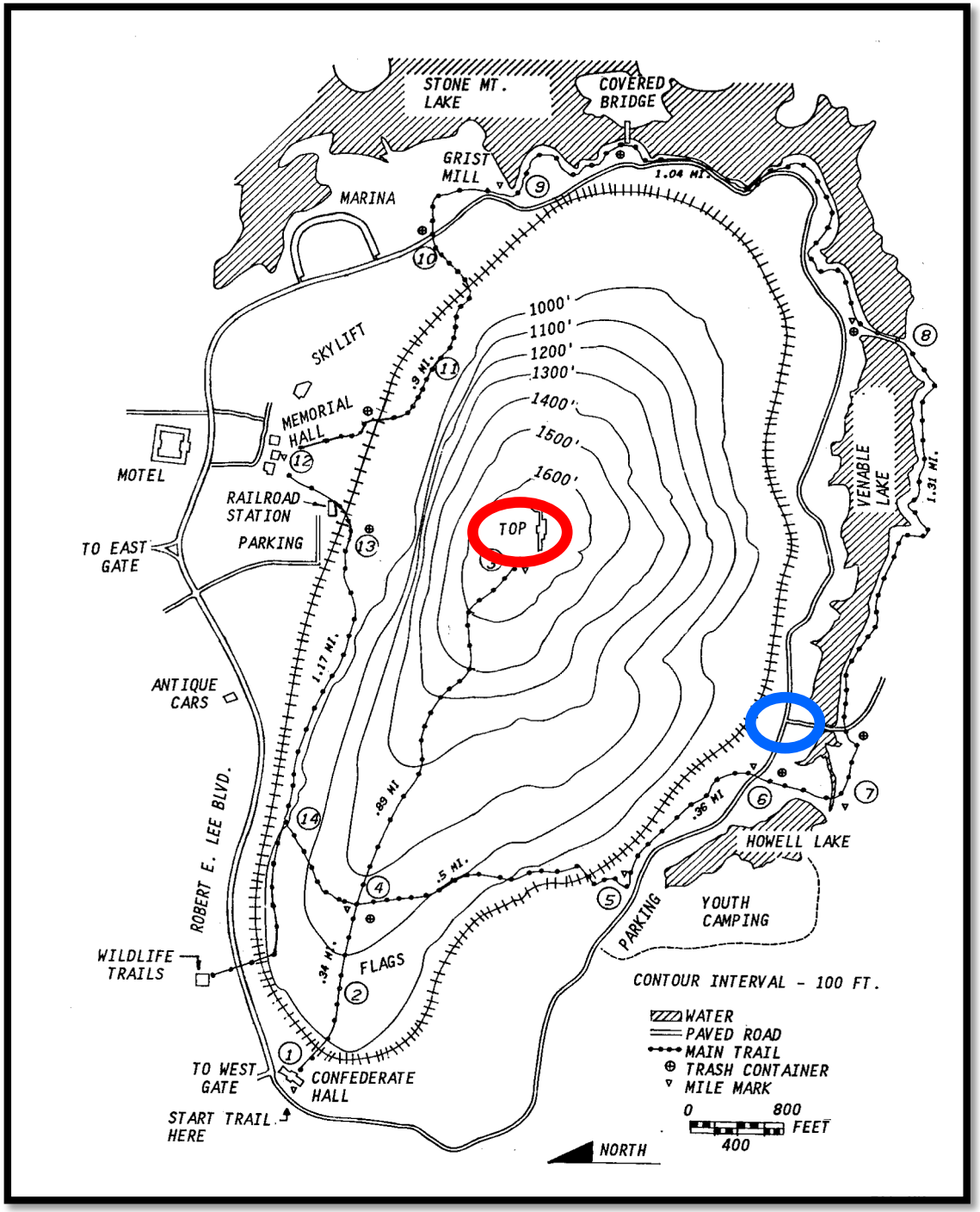
I will highlight how **napkin math** is a powerful tool for:

1. **Answering policy questions (“Aiming off”)**
2. **Building complex models (“What did I do?”)**
3. **Using complex models (“What do we do?”),**

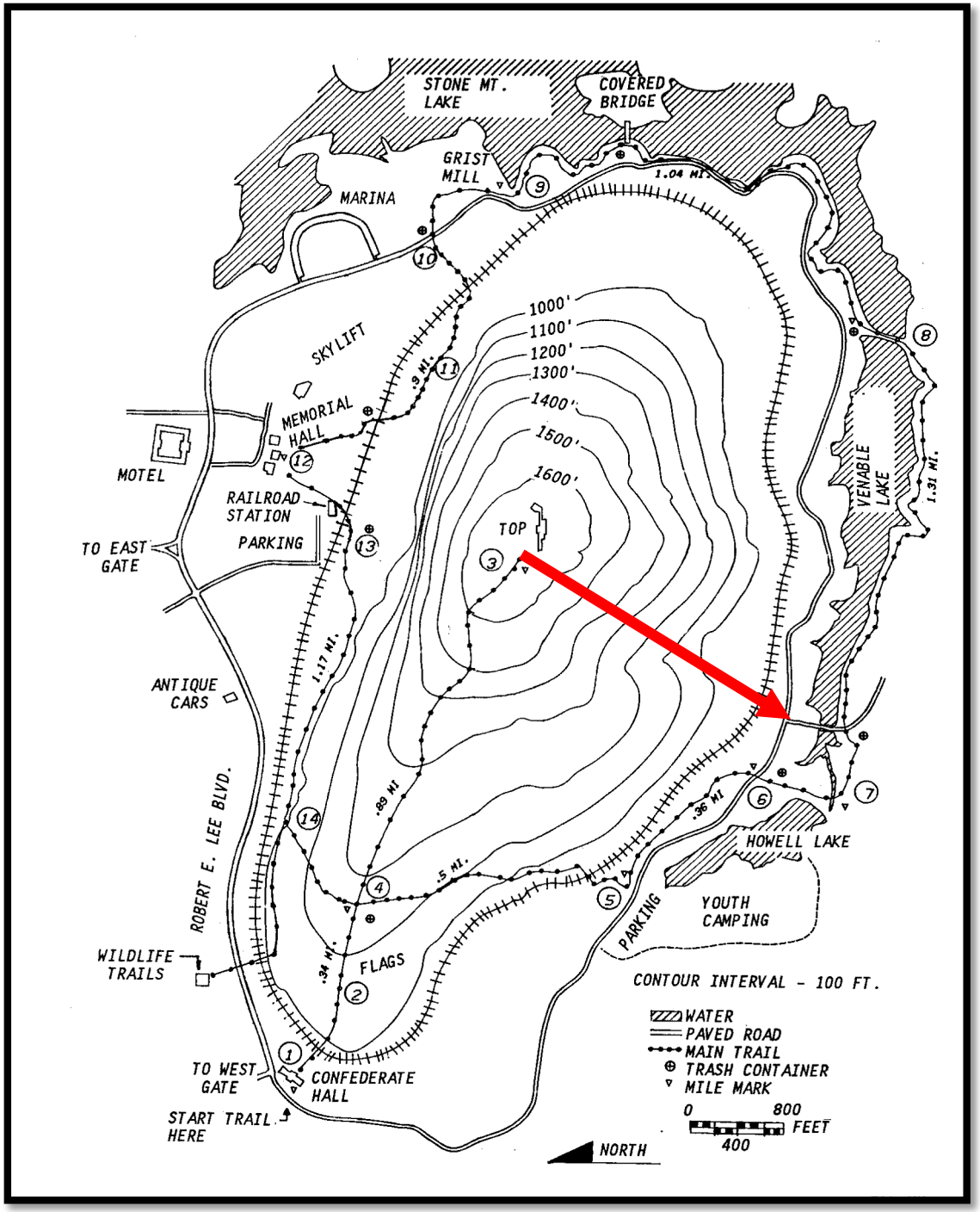
Aiming off

Even if the world is complicated,
your decision may not be.

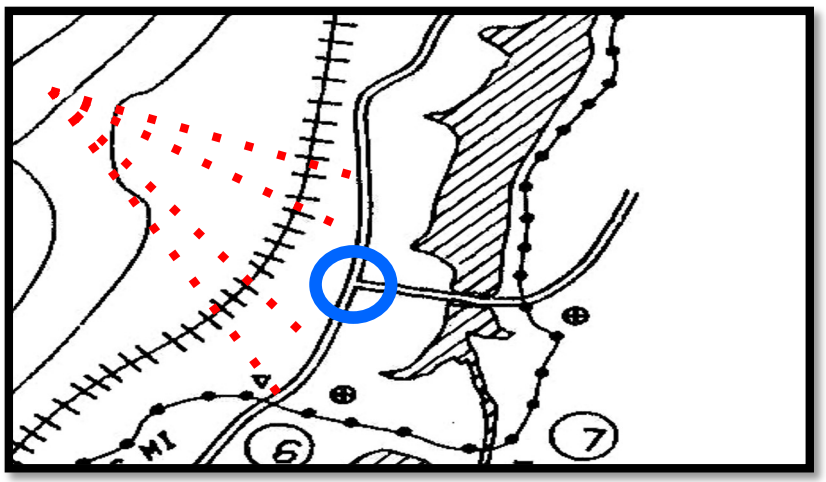
Aiming off



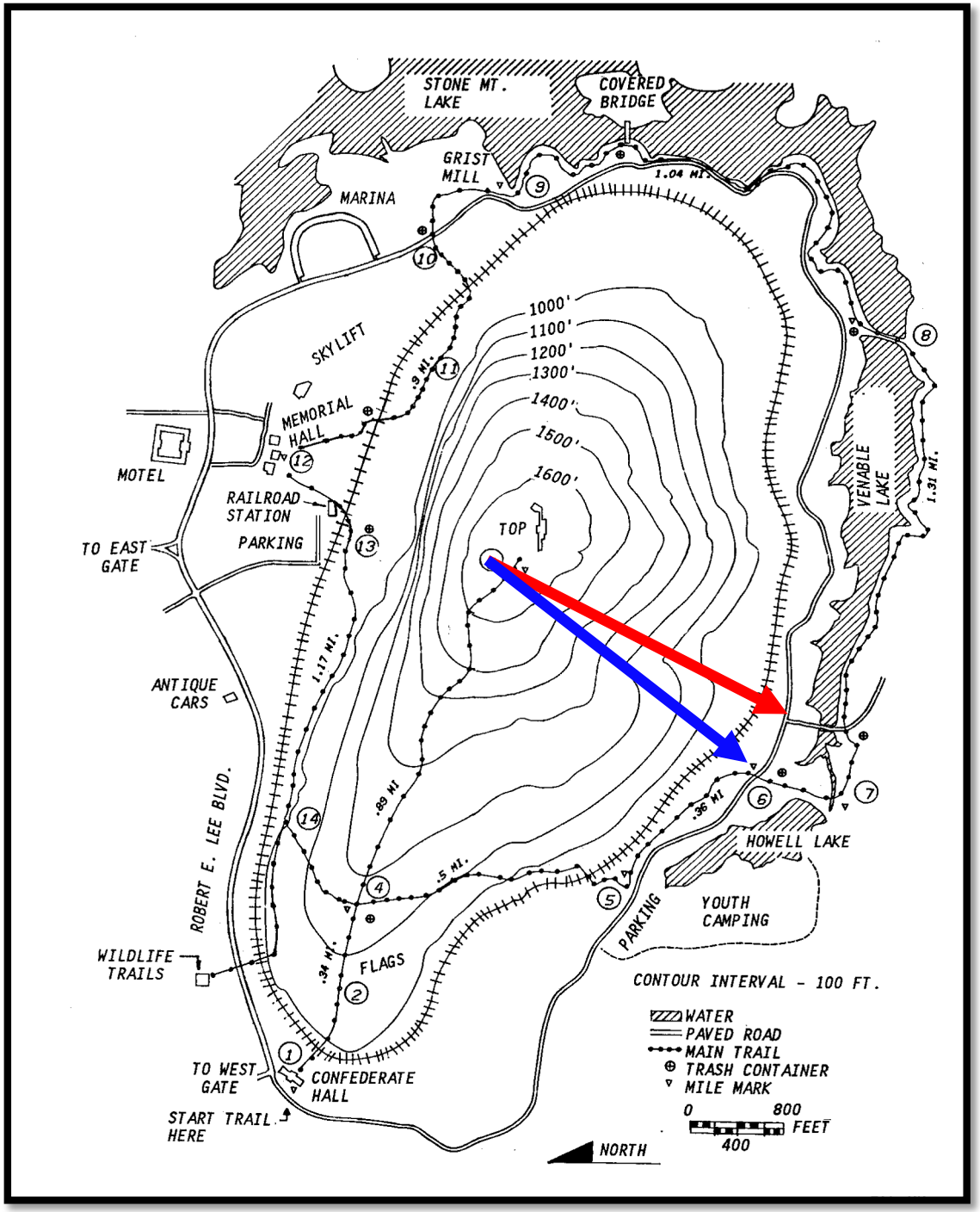
Should you follow your best compass bearing?



In a heavily wooded area, chances of hitting your target directly are low.



Consider...


STONE MT. LAKE
COVERED BRIDGE
GRIST MILL
MARINA
1.04 MI.
SKYLIFT
1000'
1100'
1200'
1300'
1400'
1500'
1600'
.9 MI.
MEMORIAL HALL
MOTEL
RAILROAD STATION
PARKING
TO EAST GATE
TOP
VENABLE LAKE
1.31 MI.
ANTIQUE CARS
1.17 MI.
ROBERT E. LEE BLVD.
.89 MI
.36 MI
HOWELL LAKE
.5 MI.
YOUTH CAMPING
PARKING
WILDLIFE TRAILS
.34 MI.
FLAGS
CONTOUR INTERVAL - 100 FT.
WATER
PAVED ROAD
MAIN TRAIL
TRASH CONTAINER
MILE MARK
0
800
FEET
400
TO WEST GATE
CONFEDERATE HALL
START TRAIL HERE
NORTH

Aiming off

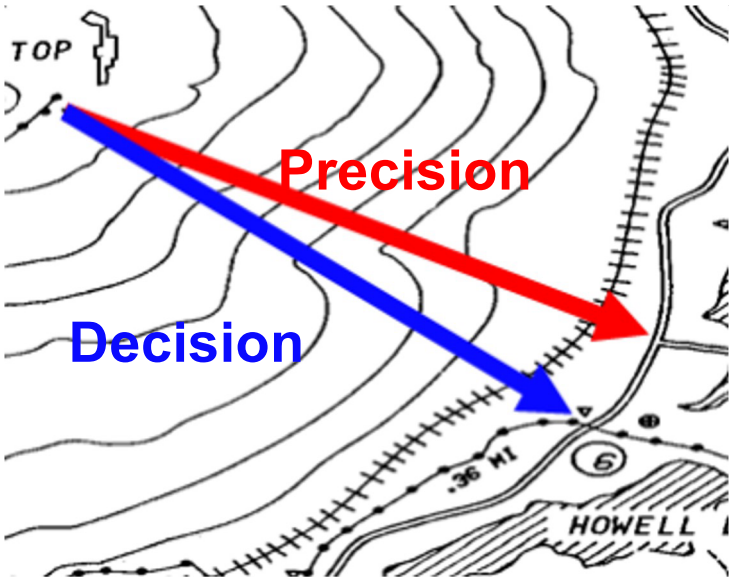
A fundamental strategy for wilderness navigation

A particularly useful technique in bad weather or when your view of the destination is blocked by tree cover or the contour of the land.

Aiming off

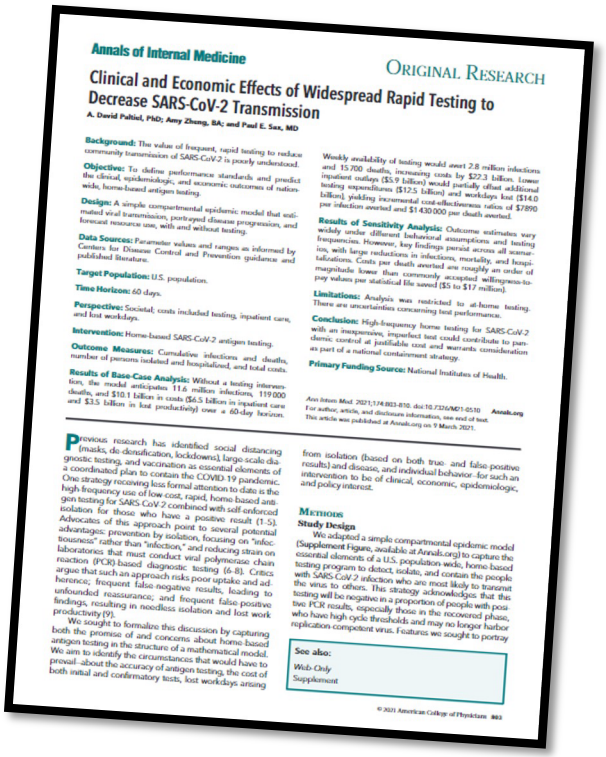
Also, a useful strategy for decision making under uncertainty: **bias your inputs.**

Illustrates a theme for today: Managing the decision/ precision tradeoff.



Application

What would it cost and what would we get by mailing rapid, antigen-based tests for SARS-CoV-2 to every household in the US?



Annals of Internal Medicine

ORIGINAL RESEARCH

Clinical and Economic Effects of Widespread Rapid Testing to Decrease SARS-CoV-2 Transmission

A. David Paltiel, PhD; Amy Zheng, BA; and Paul E. Sax, MD

Background: The value of frequent, rapid testing to reduce community transmission of SARS-CoV-2 is poorly understood.

Objective: To define performance standards and predict the clinical, epidemiologic, and economic outcomes of nationwide, home-based antigen testing.

Design: A simple compartmental epidemic model that estimated viral transmission, portrayed disease progression, and forecast resource use, with and without testing.

Data Sources: Parameter values and ranges as informed by Centers for Disease Control and Prevention guidance and published literature.

Target Population: U.S. population.

Time Horizon: 60 days.

Perspective: Societal; costs included testing, inpatient care, and lost workdays.

Intervention: Home-based SARS-CoV-2 antigen testing.

Outcome Measures: Cumulative infections and deaths, number of persons isolated and hospitalized, and total costs.

Results of Base-Case Analysis: Without a testing intervention, the model anticipates 11.6 million infections, 119 000 deaths, and $10.1 billion in costs ($6.5 billion in inpatient care and $3.5 billion in lost productivity) over a 60-day horizon. Weekly availability of testing would avert 2.8 million infections and 15 700 deaths, increasing costs by $27.3 billion. Lower inpatient outlays ($5.9 billion) would partially offset additional testing expenditures ($12.5 billion) and workdays lost ($14.0 billion), yielding incremental cost-effectiveness ratios of $7890 per infection averted and $1 430 000 per death averted.

Results of Sensitivity Analysis: Outcome estimates vary widely under different behavioral assumptions and testing frequencies. However, key findings persist across all scenarios, with large reductions in infections, mortality, and hospitalizations. Costs per death averted are roughly an order of magnitude lower than commonly accepted willingness-to-pay values per statistical life saved ($5 to $17 million).

Limitations: Analysis was restricted to at-home testing. There are uncertainties concerning test performance.

Conclusion: High-frequency home testing for SARS-CoV-2 with an inexpensive, imperfect test could contribute to pandemic control at justifiable cost and warrants consideration as part of a national containment strategy.

Primary Funding Source: National Institutes of Health.

Ann Intern Med. 2021;174:803-810. doi:10.7326/M21-0530 Annals.org
For author, article, and disclosure information, see end of text.
This article was published at Annals.org on 9 March 2021.

Previous research has identified social distancing (masks, de-densification, lockdowns), large-scale diagnostic testing, and vaccination as essential elements of a coordinated plan to contain the COVID-19 pandemic. One strategy receiving less formal attention to date is the high-frequency use of low-cost, rapid, home-based antigen testing for SARS-CoV-2 combined with self-enforced isolation for those who have a positive result (1–5). Advocates of this approach point to several potential advantages: prevention by isolation, focusing on "infectiousness" rather than "infection," and reducing strain on laboratories that must conduct viral polymerase chain reaction (PCR)–based diagnostic testing (6–8). Critics argue that such an approach risks poor uptake and adherence; frequent false-negative results, leading to unfounded reassurance; and frequent false-positive findings, resulting in needless isolation and lost work productivity (9).

We sought to formalize this discussion by capturing both the promise of and concerns about home-based antigen testing in the structure of a mathematical model. We aim to identify the circumstances that would have to prevail—about the accuracy of antigen testing, the cost of both initial and confirmatory tests, lost workdays arising from isolation (based on both true- and false-positive results) and disease, and individual behavior—for such an intervention to be of clinical, economic, epidemiologic, and policy interest.

Methods

Study Design

We adapted a simple compartmental epidemic model (Supplement Figure, available at Annals.org) to capture the essential elements of a U.S. population–wide, home-based testing program to detect, isolate, and contain the people with SARS-CoV-2 infection who are most likely to transmit the virus to others. This strategy acknowledges that this testing will be negative in a proportion of people with positive PCR results, especially those in the recovered phase, who have high cycle thresholds and may no longer harbor replication-competent virus. Features we sought to portray

See also:
Web-Only
Supplement


The New York Times
Why You Still Might Want to Have a Home Covid Test on Hand
At-home rapid Covid-19 tests can offer unique benefits for weddings, parties, travel or for households with children or at-risk adults.
COVID-19 Ag CARD
Associated Press Images
Abbott
COVID-19 Ag CARD
ellume
COVID-19 home test
IntelliSwab
COVID-19 RAPID TEST
HOME TEST
HealthAffairs
YALE UNIVERSITY
COVID-19
Topics
Journal
Blog
Podcasts
HEALTH AFFAIRS BLOG
RELATED TOPICS: COVID-19 | QUARANTINE | PANDEMICS | CAREGIVERS
Screening To Prevent SARS-CoV-2 Outbreaks: Saliva-Based Antigen Testing Is Better Than The PCR Swab
A. David Paltiel, Rochelle P. Walensky
SEPTEMBER 11, 2020
10.1377/hblog20200909.430047
ADD TO FAVORITES
SHARE
At first blush, you might reasonably dismiss the rapid, saliva-based antigen assay for coronavirus. What possible purpose can a test serve in a pandemic if it misses, say, 30 of every 100 infections it encounters, leaving those cases undetected, untreated, and uncontained? What good is a test that creates that much false reassurance and causes exactly the wrong
We Can End The Pandemic TODAY
Rapid tests can help stop the spread of COVID-19
The New York Times
OPINION
A Cheap, Simple Way to Control the Coronavirus
With easy-to-use tests, everyone can check themselves every day.
The New York Times
THE MORNING NEWSLETTER
Where Are the Tests?
Other countries are awash in Covid tests. The U.S. is not.
Would at-home COVID-19 testing be too expensive?
Example of a COVID-19 paper strip test
MED CRAM
Michael Mina, MD, PhD
Harvard School of Public Health

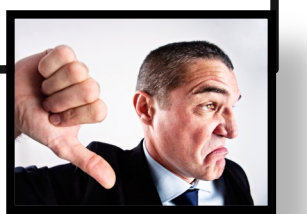
The evidence was already there…

Comparative cost-effectiveness of SARS-CoV-2 testing strategies in the USA: a modelling study

Despite the intimidating upfront costs, mass testing with rapid surveillance tests coupled with strict but relatively short isolation of confirmed cases is recommended to health authorities and local governments as a cost-effective strategy for mitigating the unprecedented threat of the COVID-19 pandemic, before safe and efficacious vaccines can be widely administered or efficacious drugs become available.

…but skeptics worried:

- poor uptake
- imperfect adherence
- frequent false-negatives
- frequent false-positives



"Even if" analysis

Stacking the deck against mass testing:

- as many as 75% of tests go straight into the garbage can;
- as many as 75% of positive test findings are simply ignored;
- each day, up to 33% of those in isolation abandon and return to the community;
- test specificity 95% (best guess: 98.5%);
- test sensitivity 80% (best guess: ~100%);
- test cost $5 (best guess: $0.20)
- value of a statistical life $5.3M (range $5.3-$15.6M)

Results

	No Testing	Testing (Base Case)	Testing (Worst Case)
Total infections	11,600,000	8,810,000	11,000,000
Total Deaths	119,000	103,000	116,000
Total Costs ($ billions)	10.1	32.4	24.1
Infections averted		2,830,000	634,000
Deaths averted		15,700	3,390
Cost per death averted		1,430,000	4,140,000

Paltiel, et al. Ann Intern Med 2021. DOI: 10.7326/M21-0510

→ Engage a debate

Re-testing to confirm HIV diagnosis before ART initiation

Clinical Infectious Diseases

BRIEF REPORT

The Cost of Not Retesting: Human Immunodeficiency Virus Misdiagnosis in the Antiretroviral Therapy "Test-and-Offer" Era

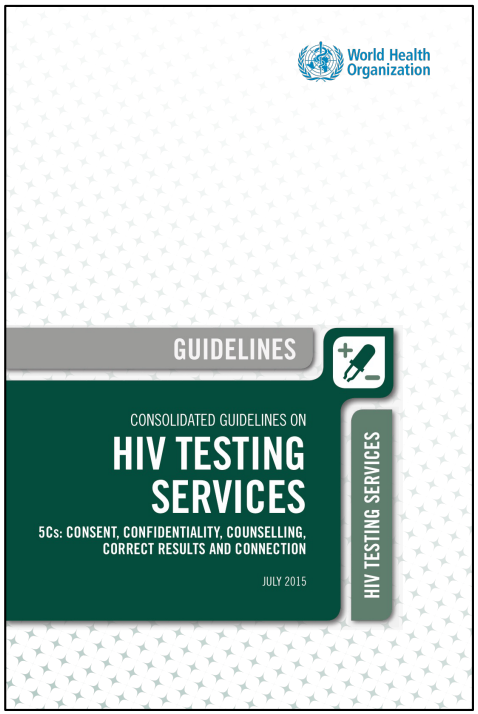
Jeffrey W. Eaton,[1] Cheryl C. Johnson,[2] and Simon Gregson[1,3]

[1]Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom; [2]HIV Department, World Health Organization, Geneva, Switzerland; and [3]Biomedical Research and Training Institute, Harare, Zimbabwe

Eaton, Johnson, Gregson Clin Infect Dis 2017; 65:522-5

Re-testing to confirm HIV diagnosis before ART initiation

- Longstanding WHO recommendation to re-test people with HIV prior to lifelong ART initiation to confirm HIV positive status
  - Increased importance under 'test-and-start'
- Substantial evidence that HIV misdiagnosis occurs in several global settings
  - Malawi 2015: 4.6% of people referred for ART were subsequently found to be HIV-negative when retested
- Recommendation very poorly implemented
  - Out of 48 national HTS policies, only 2 countries mentioned re-testing (circa 2016)



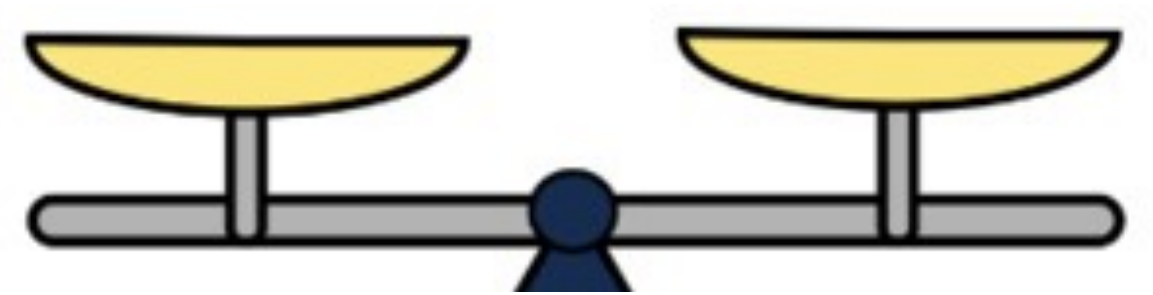
Re-testing to confirm HIV diagnosis before ART initiation

Reasons why HIV misdiagnosis is bad

- Individual and family consequences
- Legal exposure to the provider
- Undermines confidence in health system
- Lifelong ART for someone HIV-negative is expensive

Reasons cited for not implementing confirmation testing recommendation

- HIV testing algorithms 'highly accurate' (>99.5% specificity)
- Perceived high cost of re-testing everyone before ART
- Health worker burden and capacity constraints
- Uncertainty how / where to implement



Re-testing to confirm HIV diagnosis before ART initiation

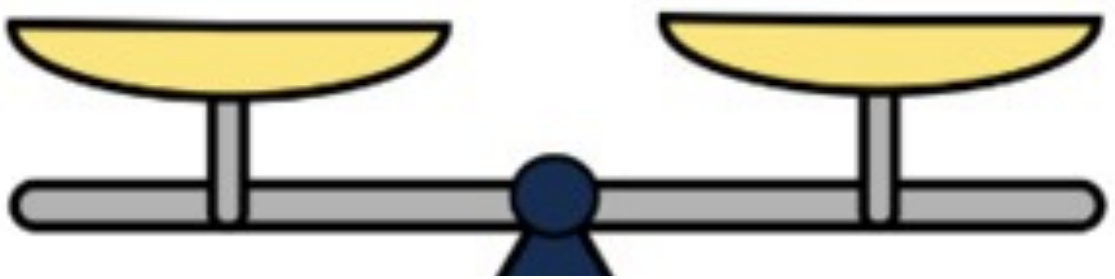
Reasons why HIV misdiagnosis is bad

- Individual and family consequences
- Legal exposure to the provider
- Undermines confidence in health system
- Lifelong ART for someone HIV-negative is expensive

Challenge: very hard to quantify the 'cost' of these (potentially infinite?)

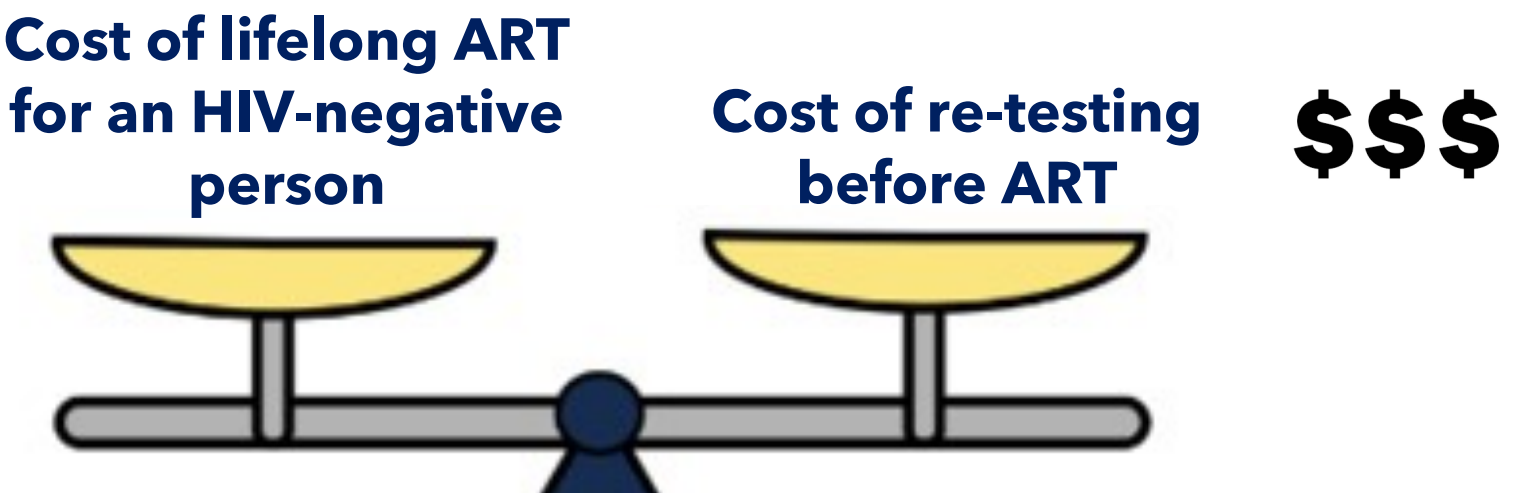
This we can quantify!

Cost of re-testing before ART (~2x ???) **$$$**



Is the benefit of **avoiding the lifelong ART cost for misdiagnosed HIV-negative** alone sufficient to offset the **cost of re-testing**?

- *Establish a lower-bound on the other unquantifiable costs*



Re-testing to confirm HIV diagnosis before ART initiation

Clinical Infectious Diseases

BRIEF REPORT

The Cost of Not Retesting: Human Immunodeficiency Virus Misdiagnosis in the Antiretroviral Therapy "Test-and-Offer" Era

Jeffrey W. Eaton,[1] Cheryl C. Johnson,[2] and Simon Gregson[1,3]

[1]Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom; [2]HIV Department, World Health Organization, Geneva, Switzerland; and [3]Biomedical Research and Training Institute, Harare, Zimbabwe

This problem turned out to be really 'boring'… (The best kind of modeling problem)

Eaton, Johnson, Gregson Clin Infect Dis 2017; 65:522-5

Re-testing to confirm HIV diagnosis before ART initiation

Clinical Infectious Diseases

BRIEF REPORT

The Cost of Not Retesting: Human Immunodeficiency Virus Misdiagnosis in the Antiretroviral Therapy "Test-and-Offer" Era

Jeffrey W. Eaton,[1] Cheryl C. Johnson,[2] and Simon Gregson[1,3]

[1]Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom; [2]HIV Department, World Health Organization, Geneva, Switzerland; and [3]Biomedical Research and Training Institute, Harare, Zimbabwe

Eaton, Johnson, Gregson Clin Infect Dis 2017; 65:522-5

Key realization from 5 minutes sketching out a model:

- Number testing **HIV-negative** is the main driver of HIV testing program costs
- **Confirmation re-testing involves only people who tested HIV positive**
  - Small and reducing fraction; not anywhere near 2x testing cost
- Cost and personnel time for HIV testing is low compared to providing ART
  - *Didn't need to consider the things that were really hard to quantify*

→ To avoid quantifying nebulous parameters

Pregnant enrollment in randomized clinical trials

SINS OF OMISSION

Model-based Estimates of the Health Effects of Excluding Pregnant Participants from Randomized Controlled Trials[1]

Alyssa Bilinski, PhD · Natalia Emanuel, PhD · Andrea Ciaranello, MD, MPH

Pregnant people are excluded from drug development RCTs by default, with the objective of protecting them and their children.

But…what happens then?

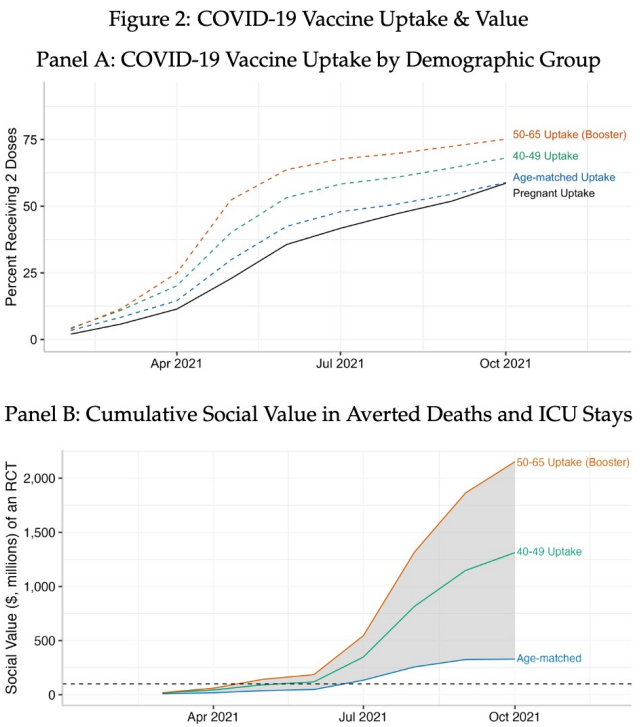
1) **Reduction in beneficial use:** Some people are hesitant to take medications that would benefit them because of a lack of evidence (1/3 reduction).

2) **Pregnant individuals can still opt to take most medications:** Still exposed to side effects (24% still take)

Pregnant enrollment in randomized clinical trials

Unless a drug is harmful AND pre-clinical evidence is sufficient to curtail its use, the current system is the worst of all worlds.

Limits benefits while still incurring harms



7% of all maternal deaths in 2021 (20% of maternal COVID-19 deaths)

→ Projecting across diverse potential future scenarios

Cancer screening

Original Investigation

August 28, 2023

Estimated Lifetime Gained With Cancer Screening Tests
A Meta-Analysis of Randomized Clinical Trials

Michael Bretthauer, MD, PhD[1]; Paulina Wieszczy, MSc, PhD[1,2]; Magnus Løberg, MD, PhD[1]; et al

Author Affiliations | Article Information

JAMA Intern Med. Published online August 28, 2023. doi:10.1001/jamainternmed.2023.3798

Results In total, 2 111 958 individuals enrolled in randomized clinical trials comparing screening with no screening using 6 different tests were eligible. Median follow-up was 10 years for computed tomography, prostate-specific antigen testing, and colonoscopy; 13 years for mammography; and 15 years for sigmoidoscopy and FOBT. The only screening test with a significant lifetime gain was sigmoidoscopy (110 days; 95% CI, 0-274 days). There was no significant difference following mammography (0 days: 95% CI, −190 to 237 days), prostate cancer screening (37 days; 95% CI, −37 to 73 days), colonoscopy (37 days; 95% CI, −146 to 146 days), FOBT screening every year or every other year (0 days; 95% CI, −70.7 to 70.7 days), and lung cancer screening (107 days; 95% CI, −286 days to 430 days).

Cancer screening

Original Investigation

August 28, 2023

Estimated Lifetime Gained With Cancer Screening Tests
A Meta-Analysis of Randomized Clinical Trials

Michael Bretthauer, MD, PhD[1]; Paulina Wieszczy, MSc, PhD[1,2]; Magnus Løberg, MD, PhD[1]; et al

Author Affiliations | Article Information

JAMA Intern Med. Published online August 28, 2023. doi:10.1001/jamainternmed.2023.3798

Results In total, 2 111 958 individuals enrolled in randomized clinical trials comparing screening with no screening using 6 different tests were eligible. Median follow-up was 10 years for computed tomography, prostate-specific antigen testing, and colonoscopy; 13 years for mammography; and 15 years for sigmoidoscopy and FOBT. The only screening test with a significant lifetime gain was sigmoidoscopy (110 days; 95% CI, 0-274 days). There was no significant difference following mammography (0 days: 95% CI, −190 to 237 days), prostate cancer screening (37 days; 95% CI, −37 to 73 days), colonoscopy (37 days; 95% CI, −146 to 146 days), FOBT screening every year or every other year (0 days; 95% CI, −70.7 to 70.7 days), and lung cancer screening (107 days; 95% CI, −286 days to 430 days).

Conclusions and Relevance The findings of this meta-analysis suggest that current evidence does not substantiate the claim that common cancer screening tests save lives by extending lifetime, except possibly for colorectal cancer screening with sigmoidoscopy.

Pull out your napkin…

Many estimates suggest that breast cancer screening is cost effective with 500-2000 screened per death averted (say, 10 years of life gained).

What does this translate to in population life expectancy gains among people eligible for screening?

1/500 x 10 x 365 = 7.3 days
1/2000 x 10 x 365 = 1.8 days
Heck, 1/100 x 10 x 365 = 37 days

Pull out your napkin…

Many estimates suggest that breast cancer screening is cost effective with 500-2000 screened per death averted (say, 10 years of life gained).

What does this translate to in population life expectancy gains among people eligible for screening?

1/500 x 10 x 365 = 7.3 days
1/2000 x 10 x 365 = 1.8 days
Heck, 1/100 x 10 x 365 = 37 days

What does this mean for screening?

for sigmoidoscopy and FOBT. The only screening test with a significant lifetime gain was sigmoidoscopy (110 days; 95% CI, 0-274 days). There was no significant difference following mammography (0 days: 95% CI, –190 to 237 days), prostate cancer screening (37 days; 95% CI, –37 to 73 days), colonoscopy (37 days; 95% CI, –146 to 146 days), FOBT screening every year or every other year (0 days; 95% CI, –70.7 to 70.7 days), and lung cancer screening (107 days; 95% CI, –286 days to 430 days).

Population-level estimates are **underpowered** to detect screening effects, with **clinically meaningfully impacts**.

→ **Are we using the right outcome?**

Aiming off

When it works, you're have an **actionable insight**. You have a **bound**.

…even when inputs are contested, hard to measure, or uncertain.

The Bad: You can aim off into an unhelpful answer.

What did I do?

Using simple models to develop complex ones

Is your complex model working?

This is a deceptively challenging task…and critical for trust.

Start with coherence checks.

- "Does this make sense?"

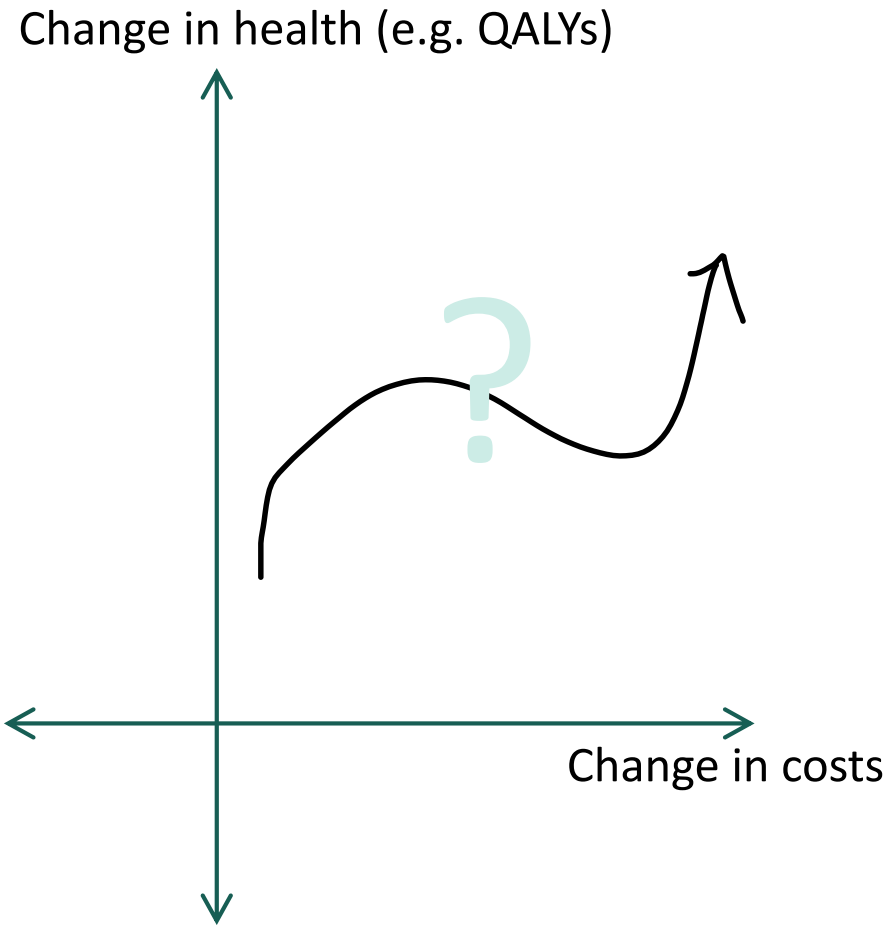
Then, apply formal testing.

- Systematic code evaluation

Benchmarking against linearity

Let's picture some intervention comparisons in the typical cost-effectiveness plane

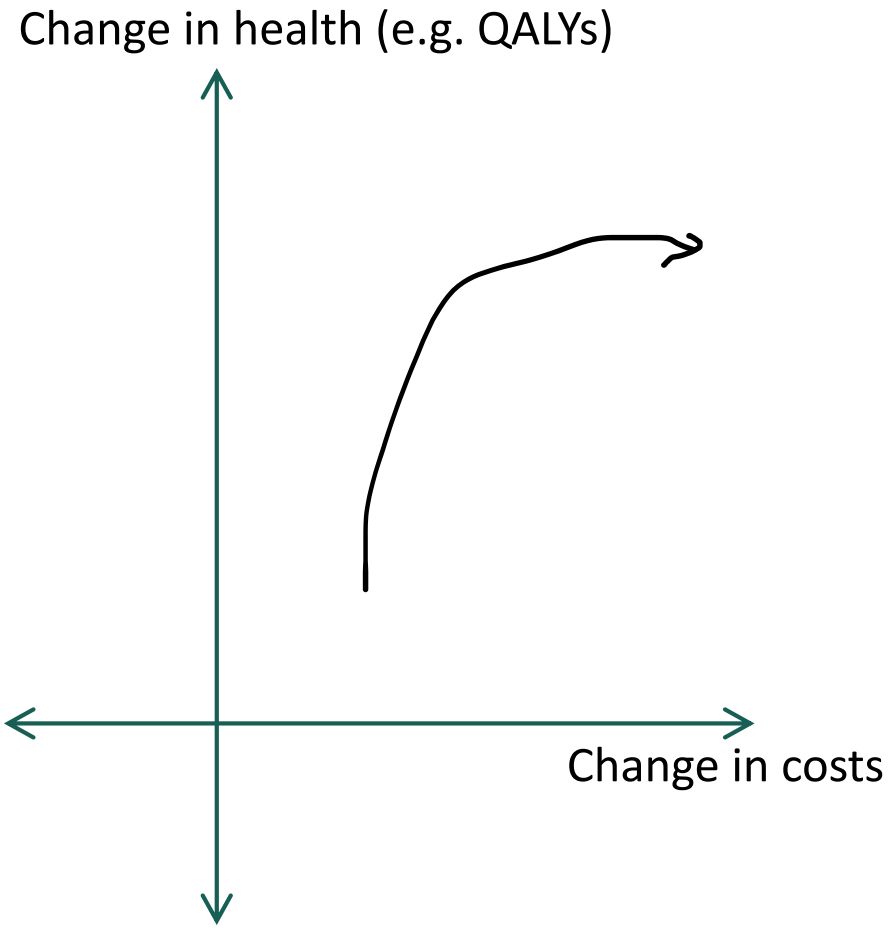
How do costs and health benefits change as we 'expand' an intervention?



Time to play … "Is this thing linear?"

Expected shape of expansion path for:

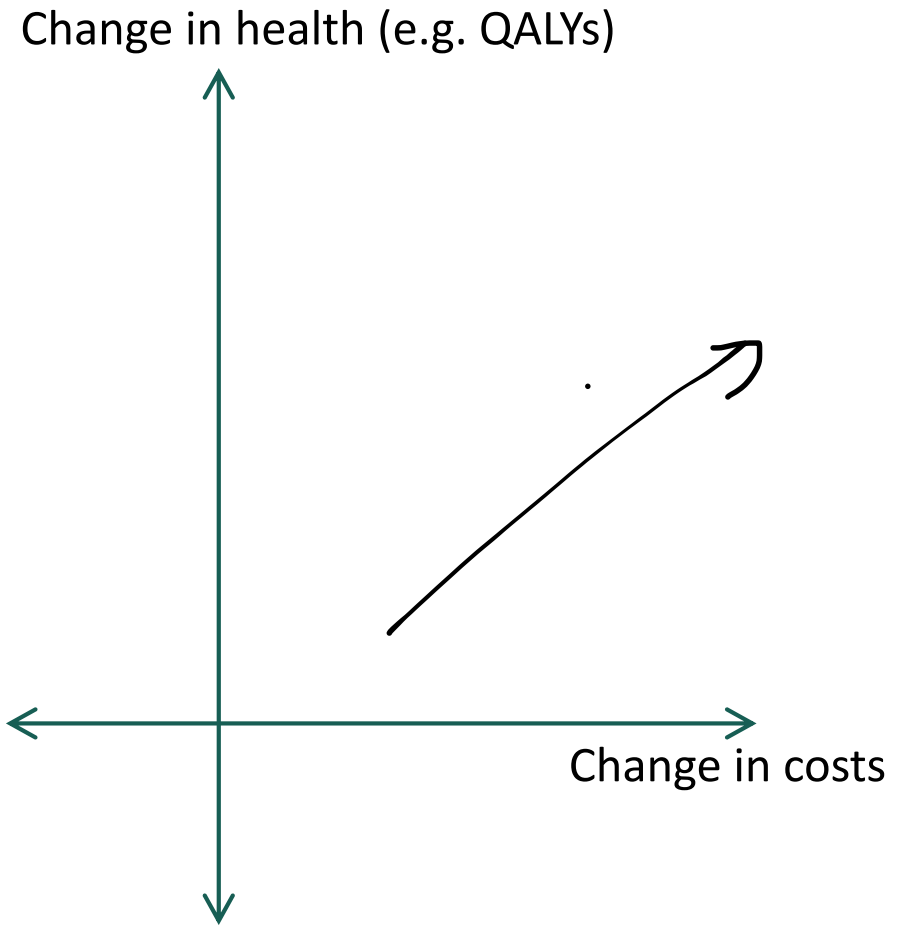
A. Cervical cancer screening, with frequency going from q10y to q1y



Time to play … "Is this thing linear?"

Expected shape of expansion path for:

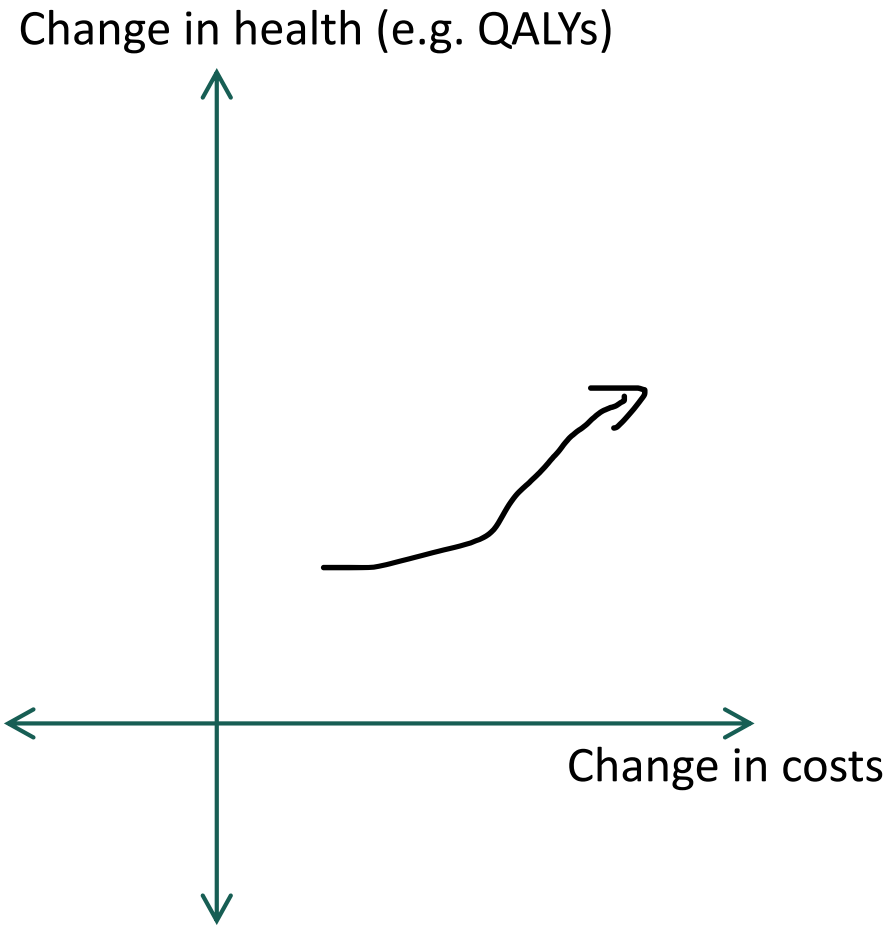
A. Cervical cancer screening, with frequency going from q10y to q1y
B. Hypertension treatment, with adherence going from 20% to 80%



Time to play … "Is this thing linear?"

Expected shape of expansion path for:

A. Cervical cancer screening, with frequency going from q10y to q1y
B. Hypertension treatment, with adherence going from 20% to 80%
C. Covid-19 vaccination with coverage going from 20% to 80%

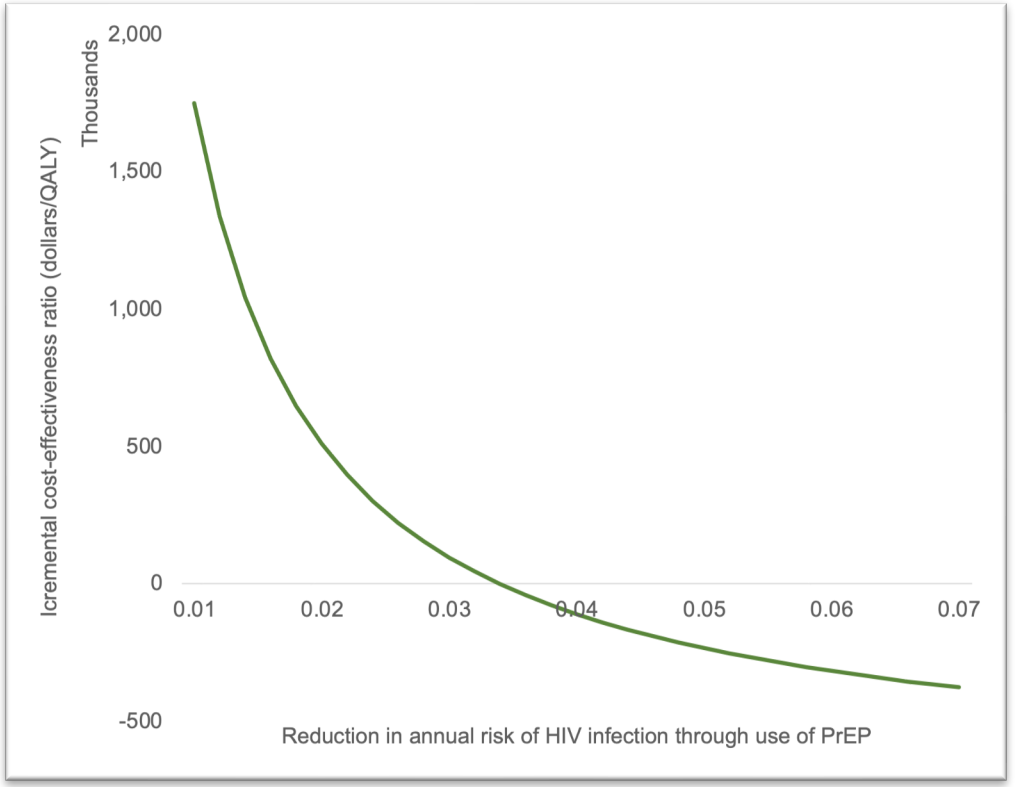


Benchmarking against linearity – let's zoom out

- When are expansion paths in the cost-effectiveness plane nonlinear?
  - Costs and effects are disproportionately affected by changes in scale
- What can make costs nonlinear?
  - Fixed costs spread out over units of production
  - Variable costs that are not constant in scale
- What can make health effects nonlinear?
  - Redundancy of effort
  - Transmission dynamics

Mini models… PrEP edition

- CDC is interested in reconsidering the way it estimates the population-level need for HIV pre-exposure prophylaxis (PrEP).
- One relevant question is how the cost-effectiveness of PrEP relates to underlying HIV risk.
- We can estimate this relationship with a very simple model.
- … but does this look right?



Direction? Shape? Magnitude?

Mini models… PrEP edition

Some assumptions…

PrEP cost per person:

$10,331

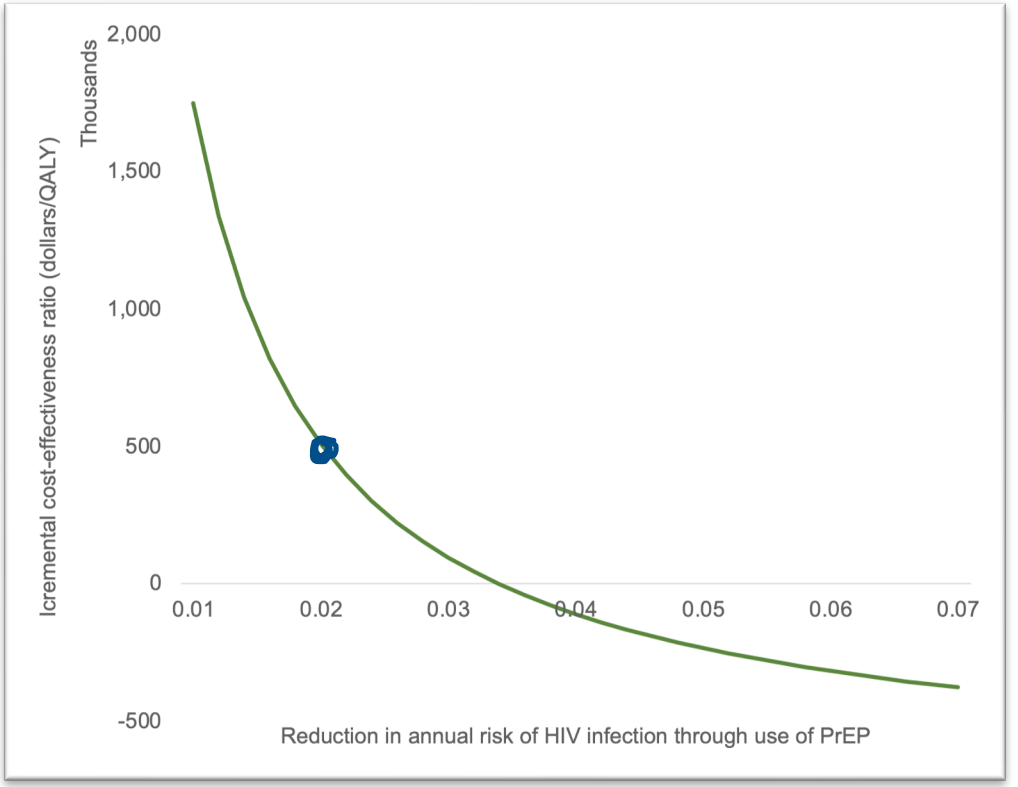
Lifetime cost of HIV:

$305,000

QALYs lost per HIV infection:

2.4

Start with a point on the curve and do a little napkin math



(Think of x-axis as expected infections averted per person taking PrEP for one year)

Mini models… PrEP edition

Some napkin math…

What is the number needed to treat to prevent one infection at 2% risk reduction? **50**

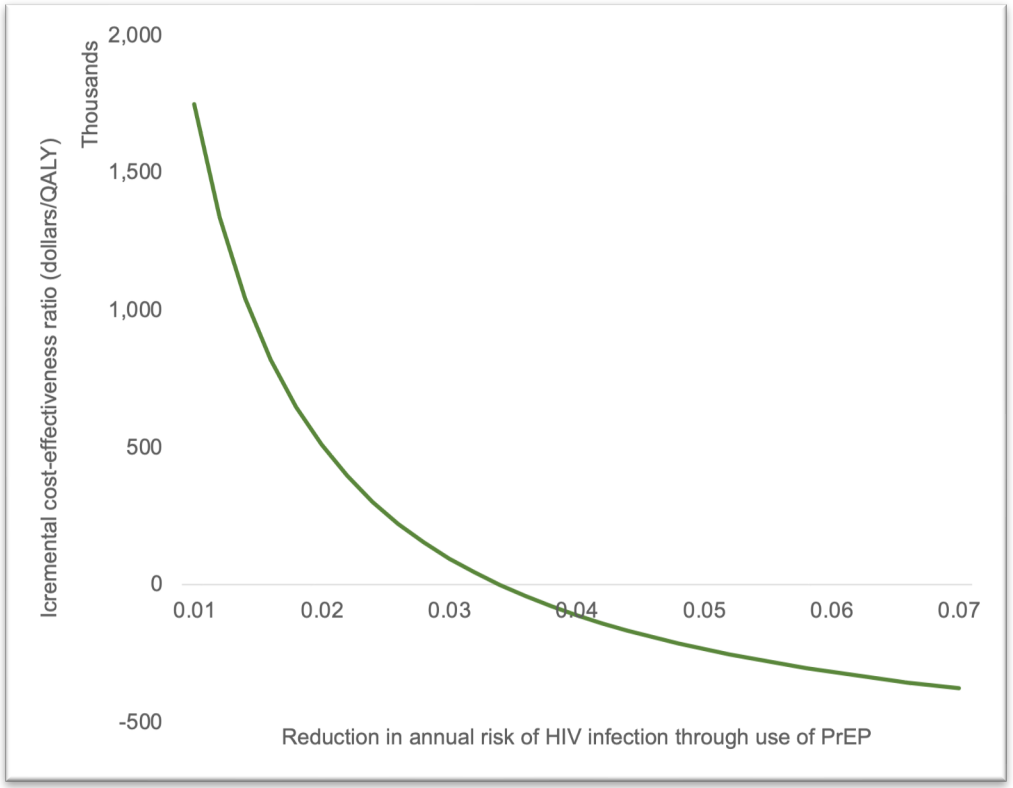
→ What is the PrEP cost to prevent an infection? **$10K*50**

→ What is the cost net of savings from averting an infection?

$10K*50-305K

→ So, roughly … what is the incremental cost per QALY at 2% risk reduction?

$$\frac{change\ in\ costs}{change\ in\ DALYs} \rightarrow \frac{\mathbf{10K * 50 - 305K}}{\mathbf{2.4}} \mathbf{\sim 80K}$$



Mini models… PrEP edition

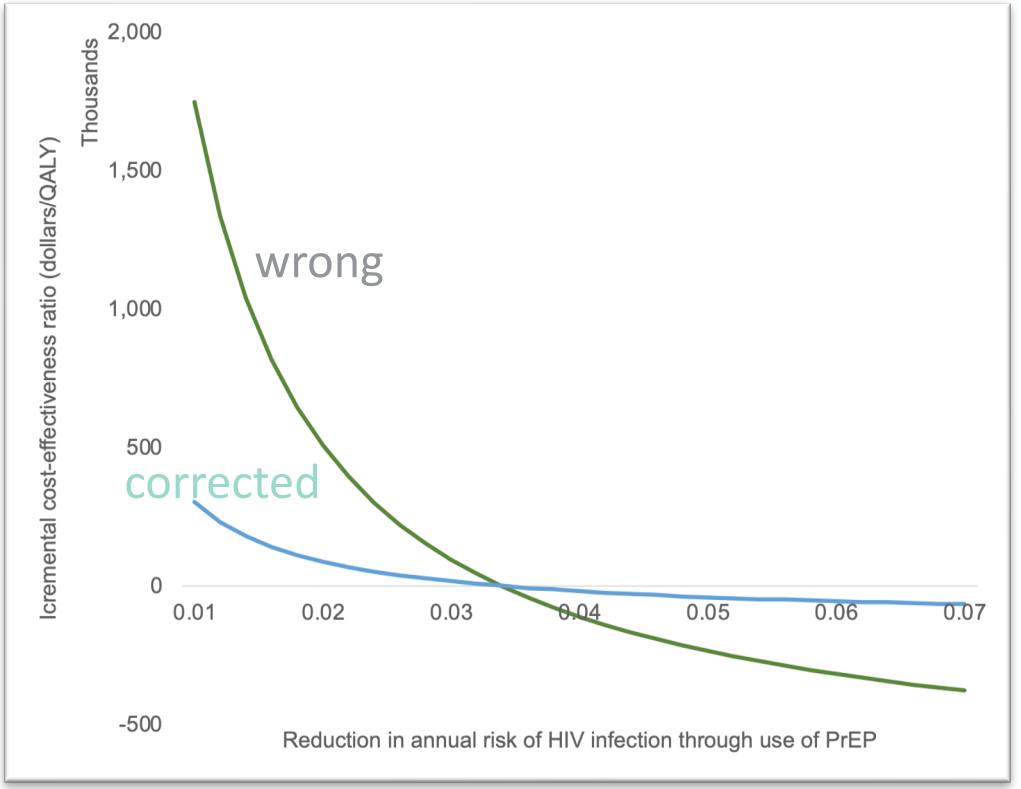
Some napkin math…

What is the number needed to treat to prevent one infection at 2% risk reduction?

→ What is the PrEP cost to prevent an infection?

→ What is the cost net of savings from averting an infection?

→ So, roughly … what is the incremental cost per QALY at 2% risk reduction?



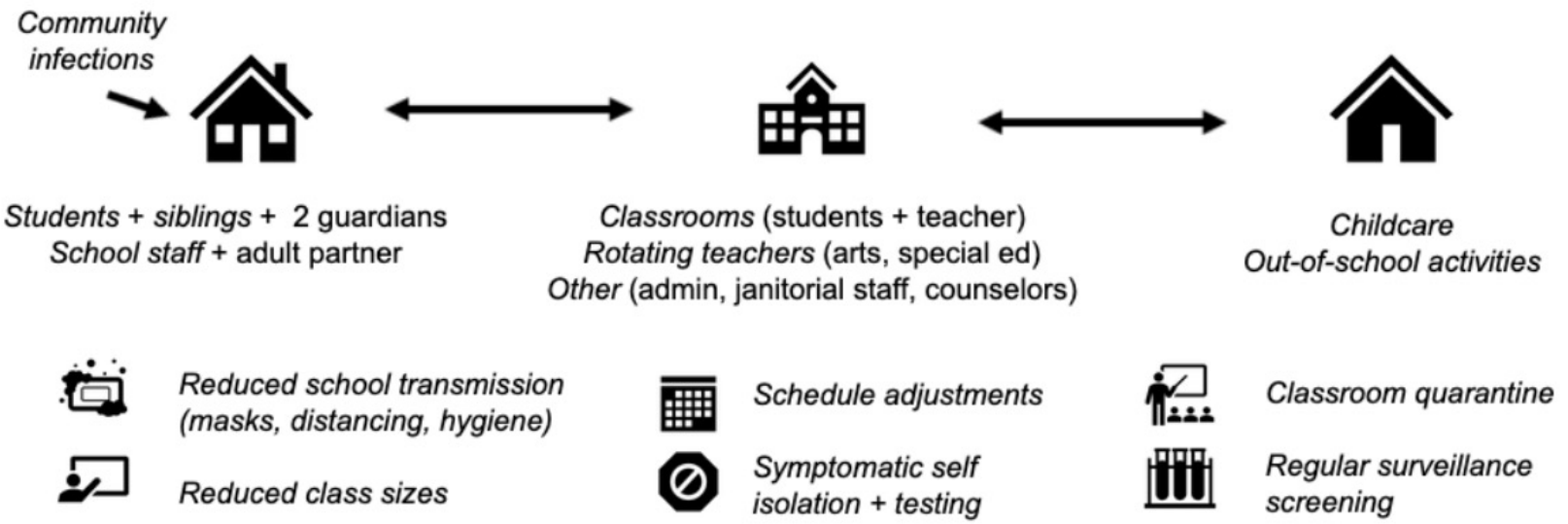
General tips for interrogating interim results

- Look at lots of XY plots
  - › You already do this in various ways (the CE plane is one, sensitivity analyses are another)
  - › …Now do more!

- Why collapse into 2D?
  - › Visualizing in >2 dimensions is hard.
  - › We try it anyway, but…

- Work backward from your summary results
  - › Look at intermediate health outcomes
  - › Disaggregate costs by major categories

Testing complex models



Testing complex models

Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

Parameter	Values	Source
Full day in-school symptomatic adult-to-adult secondary attack rate (unmitigated)		
Wild-type	2.0%	Bilinski et al,[16] 2021; Doyle et al,[24] 2021[a]
Alpha	3.5%	Davies et al,[25] 2021[a]
Delta	7.0%	Singanayagam et al,[26] 2021; Dougherty et al,[27] 2021; National Centre for Immunisation Research and Surveillance,[28] 2021[a]
Attack rate multipliers by location and duration of contact (relative to full day in-school contact)		
At-home contacts	2[b]	Assumption based on documented increased attack rates in the home (Thompson et al,[29] 2021) and increased time in close proximity
Brief contacts at school (random and specials classes)	0.125[b]	Assumed to last 1 period out of an 8-period day, with infection risk proportional to time
Brief contacts at school (staff-staff contacts)	0.25[b]	Assumed to last 1 period out of an 8-period day, but with higher risk from closer proximity (eg, break room)
Contacts between households (eg, childcare)	1	Assumption; in-school mitigation measures are not applied to these contacts
Infectiousness (relative to symptomatic adults)		
Student (in-school and asymptomatic at-home)	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Asymptomatic adult	0.5[b]	Byambasuren et al,[30] 2020; He et al,[31] 2020
Student (symptomatic at-home)	1	Paul et al,[32] 2021
Overdispersion multiplier (for adults)	Lognormal distribution (0.84, 0.3)/0.84[b]	Kerr et al,[22] 2020; Endo et al,[33] 2020
Susceptibility (relative to adults)		
Student	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Length of latent and incubation periods and infection (days)		
Time from exposure to infectious (latent period)	Maximum of gamma distribution (5.8, 0.95) minus normal distribution (2, 0.4); 1[b]	Lauer et al,[18] 2020; He et al,[19] 2020; Li et al,[20] 2020; Gatto et al,[21] 2020
Time from exposure to symptoms (if symptoms occur) (incubation period)	Gamma distribution (5.8, 0.95)[b]	Lauer et al,[18] 2020; Li et al,[20] 2020
Duration of infectious period	Lognormal distribution (5, 2)[b]	Li et al,[20] 2020; Kerr et al,[22] 2020; He et al,[19] 2020; Firth et al,[23] 2020[c]
Probability clinical/symptomatic infection		
Probability of asymptomatic infection		
Student	0.4[b]	Fontanet et al,[34] 2021; Stein-Zamir et al,[35] 2020
Adult	0.2[b]	Byambasuren et al,[30] 2020
Probability of subclinical infection, including asymptomatic		
Student	0.8[b]	Han et al,[36] 2021
Adult	0.4[b]	Upper bound of estimate from Byambasuren et al,[30] 2020
Polymerase chain reaction test characteristics		
Sensitivity (during infectious period)	0.9 (asymptomatic testing); 1 (symptomatic testing)[b]	Atkeson et al,[37] 2021; Larremore et al,[38] 2021; Cevik et al,[39] 2021; Wyllie et al,[40] 2020; Kojima et al,[41] 2021
Test turnaround time, d	1[b]	Assumption
Weekly screening parameters		
Testing uptake (fraction of school screened each week)	90%[b]	Assumption
Testing day	Monday[b]	Assumption
Hospitalization risk after SARS-CoV-2 infection		
Student (unvaccinated)	0.1%	US Centers for Disease Control and Prevention,[42] 2021; Delahoy et al,[43] 2021[a]
Adult (unvaccinated)	2.4%	US Centers for Disease Control and Prevention,[42] 2021[a]
All (vaccinated)	0%	Rosenberg et al,[44] 2021[a]
Vaccine uptake		
Student	0%, 25%, 50%, and 70% (base case); 90% (sensitivity analysis)	Assumption
Adult	70% (base case); 50% and 90% (sensitivity analysis)	US Centers for Disease Control and Prevention,[45] 2021
Vaccine effectiveness		
All individuals	70% reduction in infection risk (base case); 25%, 50%, and 90% (sensitivity analysis)	Rosenberg et al,[46] 2021; Keehner et al,[47] 2021; Fowlkes et al,[48] 2021; Puranik et al,[49] 2021; Zeng et al,[50] 2021[a]
Risk of exposure in wider local community		
Observed local incidence rate	0-50 cases per 100 000 residents per d	Assumption
Actual incidence of infections within immediate school community sourced from wider local community	3 × observed local incidence rate	Assumption

[a] eMethods 1 in the Supplement includes an explanation of how these parameters were derived from the listed sources.

[b] Baseline parameter from Bilinski, et al.[16]

[c] This value was set to match the generation time implied by observed estimates of the serial interval and presymptomatic transmission, without assuming waning infectiousness.

Testing complex models

Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

Parameter	Values	Source
Full day in-school symptomatic adult-to-adult secondary attack rate (unmitigated)		
Wild-type	2.0%	Bilinski et al,[16] 2021; Doyle et al,[24] 2021[a]
Alpha	3.5%	Davies et al,[25] 2021[a]
Delta	7.0%	Singanayagam et al,[26] 2021; Dougherty et al,[27] 2021; National Centre for Immunisation Research and Surveillance,[28] 2021[a]
Attack rate multipliers by location and duration of contact (relative to full day in-school contact)		
At-home contacts	2[b]	Assumption based on documented increased attack rates in the home (Thompson et al,[29] 2021) and increased time in close proximity
Brief contacts at school (random and specials classes)	0.125[b]	Assumed to last 1 period out of an 8-period day, with infection risk proportional to time
Brief contacts at school (staff-staff contacts)	0.25[b]	Assumed to last 1 period out of an 8-period day, but with higher risk from closer proximity (eg, break room)
Contacts between households (eg, childcare)	1	Assumption; in-school mitigation measures are not applied to these contacts
Infectiousness (relative to symptomatic adults)		
Student (in-school and asymptomatic at-home)	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Asymptomatic adult	0.5[b]	Byambasuren et al,[30] 2020; He et al,[31] 2020
Student (symptomatic at-home)	1	Paul et al,[32] 2021
Overdispersion multiplier (for adults)	Lognormal distribution (0.84, 0.3)/0.84[b]	Kerr et al,[22] 2020; Endo et al,[33] 2020
Susceptibility (relative to adults)		
Student	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Length of latent and incubation periods and infection (days)		
Time from exposure to infectious (latent period)	Maximum of gamma distribution (5.8, 0.95) minus normal distribution (2, 0.4); 1[b]	Lauer et al,[18] 2020; He et al,[19] 2020; Li et al,[20] 2020; Gatto et al,[21] 2020
Time from exposure to symptoms (if symptoms occur) (incubation period)	Gamma distribution (5.8, 0.95)[b]	Lauer et al,[18] 2020; Li et al,[20] 2020
Duration of infectious period	Lognormal distribution (5, 2)[b]	Li et al,[20] 2020; Kerr et al,[22] 2020; He et al,[19] 2020; Firth et al,[23] 2020[c]
Probability clinical/symptomatic infection		
Probability of asymptomatic infection		
Student	0.4[b]	Fontanet et al,[34] 2021; Stein-Zamir et al,[35] 2020
Adult	0.2[b]	Byambasuren et al,[30] 2020
Probability of subclinical infection, including asymptomatic		
Student	0.8[b]	Han et al,[36] 2021
Adult	0.4[b]	Upper bound of estimate from Byambasuren et al,[30] 2020
Polymerase chain reaction test characteristics		
Sensitivity (during infectious period)	0.9 (asymptomatic testing); 1 (symptomatic testing)[b]	Atkeson et al,[37] 2021; Larremore et al,[38] 2021; Cevik et al,[39] 2021; Wyllie et al,[40] 2020; Kojima et al,[41] 2021
Test turnaround time, d	1[b]	Assumption
Weekly screening parameters		
Testing uptake (fraction of school screened each week)	90%[b]	Assumption
Testing day	Monday[b]	Assumption
Hospitalization risk after SARS-CoV-2 infection		
Student (unvaccinated)	0.1%	US Centers for Disease Control and Prevention,[42] 2021; Delahoy et al,[43] 2021[a]
Adult (unvaccinated)	2.4%	US Centers for Disease Control and Prevention,[42] 2021[a]
All (vaccinated)	0%	Rosenberg et al,[44] 2021[a]
Vaccine uptake		
Student	0%, 25%, 50%, and 70% (base case); 90% (sensitivity analysis)	Assumption
Adult	70% (base case); 50% and 90% (sensitivity analysis)	US Centers for Disease Control and Prevention,[45] 2021
Vaccine effectiveness		
All individuals	70% reduction in infection risk (base case); 25%, 50%, and 90% (sensitivity analysis)	Rosenberg et al,[46] 2021; Keehner et al,[47] 2021; Fowlkes et al,[48] 2021; Puranik et al,[49] 2021; Zeng et al,[50] 2021[a]
Risk of exposure in wider local community		
Observed local incidence rate	0-50 cases per 100 000 residents per d	Assumption
Actual incidence of infections within immediate school community sourced from wider local community	3 × observed local incidence rate	Assumption

[a] eMethods 1 in the Supplement includes an explanation of how these parameters were derived from the listed sources.

[b] Baseline parameter from Bilinski, et al.[16]

[c] This value was set to match the generation time implied by observed estimates of the serial interval and presymptomatic transmission, without assuming waning infectiousness.

???

Testing complex models

Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

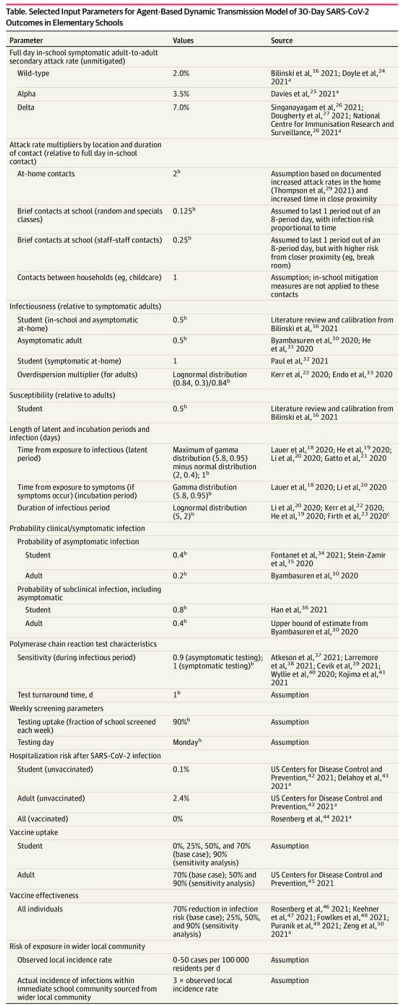


Table. Selected Input Parameters for Agent-Based Dynamic Transmission Model of 30-Day SARS-CoV-2 Outcomes in Elementary Schools

Parameter	Values	Source
Full day in-school symptomatic adult-to-adult secondary attack rate (unmitigated)		
Wild-type	2.0%	Bilinski et al,[16] 2021; Doyle et al,[24] 2021[a]
Alpha	3.5%	Davies et al,[25] 2021[a]
Delta	7.0%	Singanayagam et al,[26] 2021; Dougherty et al,[27] 2021; National Centre for Immunisation Research and Surveillance,[28] 2021[a]
Attack rate multipliers by location and duration of contact (relative to full day in-school contact)		
At-home contacts	2[b]	Assumption based on documented increased attack rates in the home (Thompson et al,[29] 2021) and increased time in close proximity
Brief contacts at school (random and specials classes)	0.125[b]	Assumed to last 1 period out of an 8-period day, with infection risk proportional to time
Brief contacts at school (staff-staff contacts)	0.25[b]	Assumed to last 1 period out of an 8-period day, but with higher risk from closer proximity (eg, break room)
Contacts between households (eg, childcare)	1	Assumption; in-school mitigation measures are not applied to these contacts
Infectiousness (relative to symptomatic adults)		
Student (in-school and asymptomatic at-home)	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Asymptomatic adult	0.5[b]	Byambasuren et al,[30] 2020; He et al,[31] 2020
Student (symptomatic at-home)	1	Paul et al,[32] 2021
Overdispersion multiplier (for adults)	Lognormal distribution (0.84, 0.3)/0.84[b]	Kerr et al,[22] 2020; Endo et al,[33] 2020
Susceptibility (relative to adults)		
Student	0.5[b]	Literature review and calibration from Bilinski et al,[16] 2021
Length of latent and incubation periods and infection (days)		
Time from exposure to infectious (latent period)	Maximum of gamma distribution (5.8, 0.95) minus normal distribution (2, 0.4); 1[b]	Lauer et al,[18] 2020; He et al,[19] 2020; Li et al,[20] 2020; Gatto et al,[21] 2020
Time from exposure to symptoms (if symptoms occur) (incubation period)	Gamma distribution (5.8, 0.95)[b]	Lauer et al,[18] 2020; Li et al,[20] 2020
Duration of infectious period	Lognormal distribution (5, 2)[b]	Li et al,[20] 2020; Kerr et al,[22] 2020; He et al,[19] 2020; Firth et al,[23] 2020[c]
Probability clinical/symptomatic infection		
Probability of asymptomatic infection		
Student	0.4[b]	Fontanet et al,[34] 2021; Stein-Zamir et al,[35] 2020
Adult	0.2[b]	Byambasuren et al,[30] 2020
Probability of subclinical infection, including asymptomatic		
Student	0.8[b]	Han et al,[36] 2021
Adult	0.4[b]	Upper bound of estimate from Byambasuren et al,[30] 2020
Polymerase chain reaction test characteristics		
Sensitivity (during infectious period)	0.9 (asymptomatic testing); 1 (symptomatic testing)[b]	Atkeson et al,[37] 2021; Larremore et al,[38] 2021; Cevik et al,[39] 2021; Wyllie et al,[40] 2020; Kojima et al,[41] 2021
Test turnaround time, d	1[b]	Assumption
Weekly screening parameters		
Testing uptake (fraction of school screened each week)	90%[b]	Assumption
Testing day	Monday[b]	Assumption
Hospitalization risk after SARS-CoV-2 infection		
Student (unvaccinated)	0.1%	US Centers for Disease Control and Prevention,[42] 2021; Delahoy et al,[43] 2021[a]
Adult (unvaccinated)	2.4%	US Centers for Disease Control and Prevention,[42] 2021[a]
All (vaccinated)	0%	Rosenberg et al,[44] 2021[a]
Vaccine uptake		
Student	0%, 25%, 50%, and 70% (base case); 90% (sensitivity analysis)	Assumption
Adult	70% (base case); 50% and 90% (sensitivity analysis)	US Centers for Disease Control and Prevention,[45] 2021
Vaccine effectiveness		
All individuals	70% reduction in infection risk (base case); 25%, 50%, and 90% (sensitivity analysis)	Rosenberg et al,[46] 2021; Keehner et al,[47] 2021; Fowlkes et al,[48] 2021; Puranik et al,[49] 2021; Zeng et al,[50] 2021[a]
Risk of exposure in wider local community		
Observed local incidence rate	0-50 cases per 100 000 residents per d	Assumption
Actual incidence of infections within immediate school community sourced from wider local community	3 × observed local incidence rate	Assumption

[a] eMethods 1 in the Supplement includes an explanation of how these parameters were derived from the listed sources.

[b] Baseline parameter from Bilinski, et al.[16]

[c] This value was set to match the generation time implied by observed estimates of the serial interval and presymptomatic transmission, without assuming waning infectiousness.

Formal testing

Design tests to evaluate that all use cases perform correctly…

…which sounds great until you realize that tests require us to **anticipate model output**.

Testing complex models

What Goes In Must Come Out: Functional testing for complex simulation models

Alyssa Bilinski,[1] Luke Massa,[2] Andrea Ciaranello[3], Meagan C. Fitzpatrick[4], John Giardina[3]

1) Collate input parameters
- Table 1 + structural parameters

2) Define and track intermediate outputs
- Add *intermediate "napkin" outputs* for each input that can reverse engineer input behavior

3) Run and report test results over different input combinations

Testing complex models

Overdispersion parameter

- Wildtype COVID-19 was "overdispersed" → heterogeneous individual infectiousness.

- Implemented as a multiplier on individual attack rate

To track attack rates:

1) Track number of contacts per day in each setting
2) Track number of new infections per day in each setting

Testing complex models

Overdispersion parameter

- Wildtype COVID-19 was "overdispersed" → heterogeneous individual infectiousness.

- Implemented as a multiplier on individual attack rate

To track attack rates:
1) Track number of contacts per day in each setting
2) Track number of new infections per day in each setting

We observed a slight underestimate in attack rate:
1) only in households
2) only with overdispersion turned on

In rare (<1/2000) cases, overdispersion could push the attack rate > 1 → NA.

Testing complex models

1. Even if you can't "see" bugs as easily, napkin thinking can still be useful for informing complex models.

2. "Napkins" can make model testing more transparent.

Table 1 Checks

Parameter	Target Value	Observed Value	Relative Difference	Approach to tracking
Baseline Attack Rate	0.070	0.070	<1%	The model code tracks the total number of contacts for each type of interaction between infected and susceptible individuals (e.g., at-home contact between an asymptomatic adult and child) and the total number of infections resulting from those contacts. In order to recover an estimate for each attack rate multiplier from these trackers, we used a three-step process. First, for all types of interactions that involved a particular multiplier (e.g., at-home attack rate multiplier), we calculated the number of infections that we would expect to see in the absence of that multiplier by multiplying the tracked number of contacts for those interactions by all the other relevant multipliers (e.g., for at-home interaction between an asymptomatic adult and child, multiply total number of contacts by the baseline attack rate, asymptomatic adult infectiousness multiplier, and child susceptibility multiplier, but not the at-home multiplier); we calculated the total number of expected infections across all types of interactions involving the particular multiplier. Second, we calculated the total number of tracked actual infections across those interactions. Third, we divided this total number of actual infections by the total number of infections we would expect in the absence of the multiplier (calculated in the first step). If the model code has implemented the multiplier correctly, this quotient will equal (in the limit) the particular multiplier parameter we want to recover.
At-school mitigation multiplier	0.500	0.501	<1%	
At-school mitigation multiplier	1.000	1.002	<1%	
At-home attack rate multiplier	2.000	2.015	<1%	
Brief contact multiplier	0.125	0.125	<1%	
Staff-staff contact multiplier	2.000	1.996	<1%	
Child care contact multiplier	1.000	1.000	<1%	
Child infectiousness multiplier	0.500	0.501	<1%	
Asymptomatic adult infectiousness multiplier	0.500	0.497	<1%	
Symptomatic child at-home infectiousness multiplier	2.000	2.020	<1%	
Child susceptibility multiplier	0.500	0.500	<1%	

				Tracker names (in abm_code.R): person.days.at.risk.home.parents, etc. (for number of contacts per interaction type) and the location, source, adult, and source_symp variables (for number of infections from each interaction type). These trackers are summarized., respectively, at the end of the model run in the risk_ct_sympA_A_home, etc. and inf_ct_sympA_A_home, etc. variables.
Latent Period (days)	3.046	3.045	<1%	Each day in the model, it is checked who is infected but not infectious (latent), infected but not symptomatic (incubation), or infectious at home. For people who meet the criteria, 1 is added to a tracker for the latent, incubation, or infectious period. Individuals infected in the broader community are not included in this tracker, because they can become infected in the "start-up" period in the model. The infectious period at home is tracked instead of other infectious period metrics (e.g., infectious days at school), since the other metrics are impacted by policies like testing and quarantine. The code for these policies is tested using the "structural checks" described below. *Tracker names (in abm_code.R):* exposed_not.inf_days, exposed_not.symp_days, and inf_home_days, to track latent, incubation, and infectious periods, respectively.
Incubation Period (days)	5.010	5.012	<1%	
Infectious Period (days)	5.051	5.051	<1%	
Probability of asymptomatic infection (child)	0.400	0.398	<1%	Total number of infected individuals who are not flagged as symptomatic divided by total number of infected individuals. *Tracker names (in abm_code.R):* symp
Probability of asymptomatic infection (adult)	0.200	0.200	<1%	
Probability of subclinical infection (child)	0.800	0.795	<1%	Total number of infected individuals who are flagged as subclinical divided by total number of infected individuals. *Tracker names (in abm_code.R):* sub_clin
Probability of subclinical infection (adult)	0.400	0.400	<1%	

Screening Test Sensitivity	0.900	0.900	<1%	Total number of true positive tests divided by total number of tests conducted. *Tracker names (in abm_code.R):* pcr_tp_count for number of true positive tests and test_ct for total number of tests conducted.
Screening Test Uptake	0.900	0.900	<1%	Total number of tests conducted divided by total number of times individuals were eligible (i.e., in-school on a screening day). *Tracker names (in abm_code.R):* test_ct for total number of tests conducted and test_regular_eligible for number of times individuals were eligible.
Hospitalization Rate (unvaccinated child)	0.001000	0.000999	<1%	First, the total number of infections tracked in the model was multiplied by the fraction of the susceptible population (i.e., unvaccinated or "non-effective" vaccination) that is unvaccinated. Then, the total number of hospitalized individuals was divided by this number of infections in unvaccinated individuals to recover the hospitalization rate. *Tracker names (in abm_code.R):* children for total number of children infected, adult for total number of teachers infected, family for total number of adult family members infected, hosp_child for total number of children hospitalized, and hosp_adult for the total number of adults hospitalized.
Hospitalization Rate (unvaccinated adult)	0.024000	0.023996	<1%	
Vaccine uptake (student)	0.250	0.250	<1%	The total number of individuals flagged as vaccinated was divided by the total number of individuals. *Tracker names (in abm_code.R):* vacc
Vaccine uptake (teacher)	0.700	0.700	<1%	
Vaccine uptake (family)	0.700	0.700	<1%	
Vaccine effectiveness	0.700	0.700	<1%	The total number of individuals flagged as "not susceptible" was divided by the total number of individuals flagged as vaccinated. *Tracker names (in abm_code.R):* susp for susceptible individuals and vacc for vaccinated individuals.

Local Incidence Rate (cases per residents per day)	0.000150	0.000150	<1%	The total number of individuals infected in the wider community was divided by the product of the total number of days run and the total number of individuals in the model. When an individual was infected within the model (e.g., at school, not in the wider community), the number of days remaining in the model run after their infection was subtracted from the denominator of this fraction. *Tracker names (in abm_code.R):* child.start.count and adult.start.count for total number of individuals infected in the wider community, and child.community.risk.days and adult.community risk days for the number of days individuals were at risk of infection from the wider community.
Local Incidence Rate (cases per residents per day)	0.000750	0.000749	<1%	
Local Incidence Rate (cases per residents per day)	0.001500	0.001497	<1%	

Structural Checks

Household Contact Structure

Model run stopped if the household members contacted by each infected individual do not match the list of all uninfected and susceptible individuals in that household (line 403 in abm_code.R).

Model run stopped if an infected individual does not contact anyone in their household, but there are uninfected individuals in that household (line 412 in abm_code.R).

Model run is stopped if an infected individual who was not infected in the wider community does not contact any household members (line 1286 in abm_code.R).

Model run is stopped if an individual infected in the wider community infects an individual in their household (line 1311 in abm_code.R).

In-School Transmissions

Model run is stopped if an infectious individual is at school on a weekend day (line 1326 in abm_code.R).

Model run is stopped if individuals infected at school were not supposed to be present at school on that day (lines 1345-1364 in abm_code.R). The list of individuals present in school on each given day is determined by a separate testing function that takes into account the quarantine, isolation, and testing policy structure in the model (lines 807-866 in abm_code.R).

Classroom Contact Structure

Model run stopped if the classroom members contacted by each infected individual do not match the uninfected and susceptible individuals in their classroom present in school on a given day (lines 441-457

What do I do?

Using simple models to act on complex ones

Contact tracing models

Research Letter | Public Health

August 21, 2020

Modeling Contact Tracing Strategies for COVID-19 in the Context of Relaxed Physical Distancing Measures

Alyssa Bilinski, MS[1]; Farzad Mostashari, MD[2]; Joshua A. Salomon, PhD[3]

Article | Published: 05 August 2020

Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19

Alberto Aleta, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani ✉, Esteban Moro ✉ & Yamir Moreno ✉

ARTICLES | VOLUME 20, ISSUE 10, P1151-1160, OCTOBER 2020

Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study

Adam J Kucharski, PhD • Petra Klepac, PhD • Andrew J K Conlan, PhD • Stephen M Kissler, PhD • Maria L Tang, MMath • Hannah Fry, PhD • et al. Show all authors

Models may not reach the same qualitative conclusions.

Research Letter | Public Health

August 21, 2020

Modeling Contact Tracing Strategies for COVID-19 in the Context of Relaxed Physical Distancing Measures

Alyssa Bilinski, MS[1]; Farzad Mostashari, MD[2]; Joshua A. Salomon, PhD[3]

Similar to other models,[5,6] our estimates imply that contact tracing could support partial relaxation of physical distancing measures but not a full return to levels of contact before lockdown.

Article | Published: 05 August 2020

Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19

Alberto Aleta, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro & Yamir Moreno

syndrome coronavirus 2 (SARS-CoV-2) transmission in the Boston metropolitan area. We find that a period of strict social distancing followed by a robust level of testing, contact-tracing and household quarantine could keep the disease within the capacity of the healthcare system while enabling the reopening of economic activities. Our results show that a response

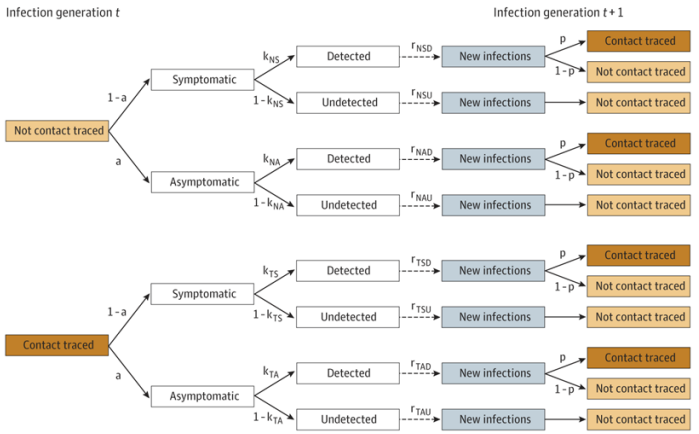
ARTICLES | VOLUME 20, ISSUE 10, P1151-1160, OCTOBER 2020

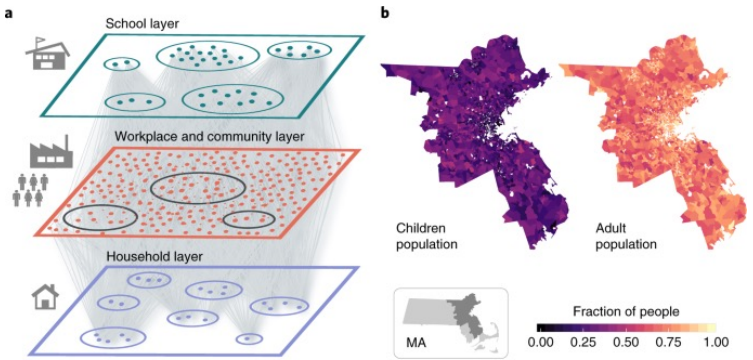
Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study

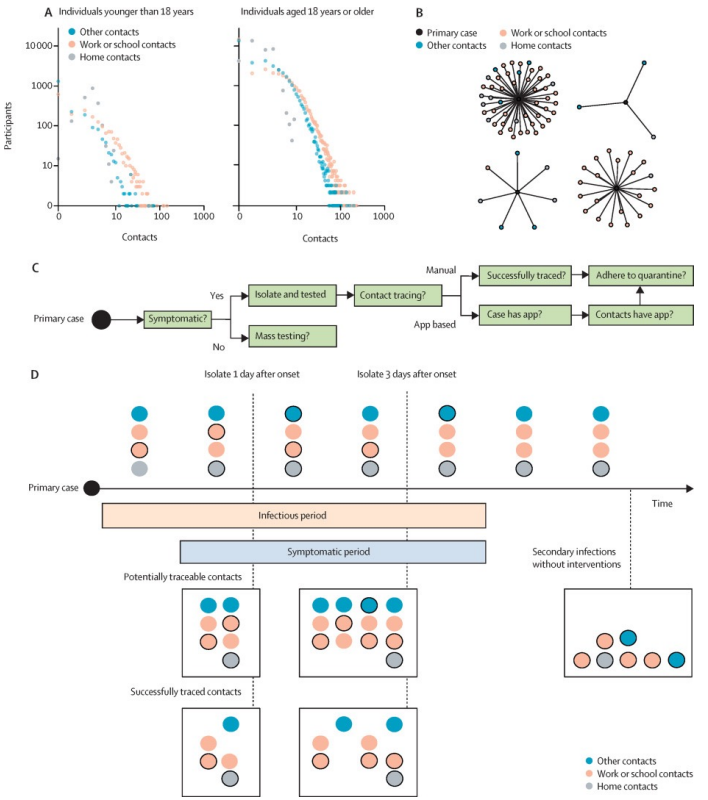
Adam J Kucharski, PhD • Petra Klepac, PhD • Andrew J K Conlan, PhD • Stephen M Kissler, PhD • Maria L Tang, MMath • Hannah Fry, PhD • et al. Show all authors

our analysis estimated that a high proportion of cases would need to self-isolate and a high proportion of their contacts to be successfully traced to ensure an effective reproduction number lower than 1 in the absence of other measures. If combined with moderate physical distancing measures, self-isolation and contact tracing would be more likely to achieve control of severe acute respiratory syndrome coronavirus 2 transmission.

And models might look quite different…







And models might look quite different…

1) Structure

- Compartmental vs. agent based

2) Deterministic vs. stochastic

3) Data/calibration

- On which populations (and subpopulations) were the model trained?

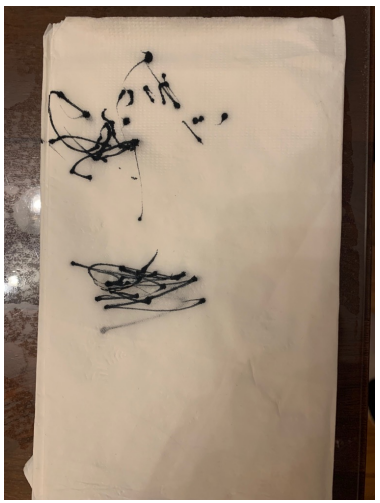
4) Parameter choices

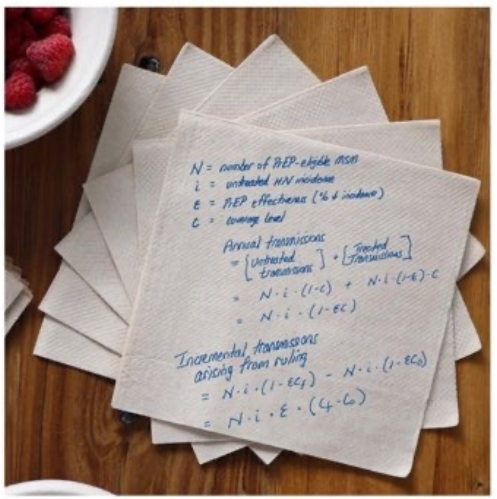
- What levels of testing/tracing/isolation can be realistically achieved?

It's easy to fall back on simple heuristics.

1) The **most complex model** is the 'best model.'
2) The model **calibrated on data most similar to my population** is the 'best model.'

Can we do better?





Step 1: Write a (napkin) model.

For epidemic infectious diseases, we often focus on the effective reproduction number **R_t**. This represents the **average number of new infections generated per infectious individual**.

- If $R_t > 1$, incident cases grow exponentially.
- If $R_t < 1$, incident cases decline exponentially.
- If $R_t = 1$, incident cases stay constant.

Step 1: Write a (napkin) model.

R_t = avg contacts x prob of transmission per contact

- If you increase contacts by 2, R_t doubles.
- If you increase probability of transmission by 2, R_t doubles.

(Sidebar: infectious diseases are one of the more complicated cases. Often, you have even simpler linear processes.)

Step 1: Write a (napkin) model.

Key question: Does test/tracing/isolation push $R_t < 1$?

Our simple model:

- initial R_0/R_t without contact tracing
- fraction of cases detected
- fraction of contacts traced
- percentage of traced contacts who isolate

Multiply!

Step 2: Fill in parameters.

	Bilinski et. al.	Kucharski et. al.
R_t without tracing		
Percentage detected		
Percentage traced		
Percentage of traced contacts who isolate		
Additional social distancing		
R_t with tracing		

Step 2: Fill in parameters.

	Bilinski et. al.	Kucharski et. al.
R_t without tracing	2.5	
Percentage detected	<50%	
Percentage traced	50%	
Percentage of traced contacts who isolate	1 (varied)	
Additional social distancing	(varied)	
R_t with tracing	1.9	

2.5*(.5 (detected)*.5 (traced)+.5 (undetected)) ~ 1.9

Step 2: Fill in parameters.

	Bilinski et. al.	Kucharski et. al.
R_t without tracing	2.5	2.6
Percentage detected	<50%	90% of symptomatic (75%) → 63%
Percentage traced	50%	100% HH (n = 2) 90% school, 79% work (n = 11) 52% other (n = 17) HH attack rate: 20% Other attack rate: 6%
Percentage of traced contacts who isolate	1 (varied)	90% (but detected 20% into disease course)
Additional social distancing	(varied)	Symptomatic self-isolation – reduces transmission by ½
R_t with tracing	1.9	1.1

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus manual contact tracing of all contacts	Yes	All	100%	21%	0·94	64%
Self-isolation plus HHQ plus app-based tracing	Yes	All	53%	30%	1·4	47%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus app-based tracing	Yes	All	90% school, 79% work, and 52% other with manual tracing; 53% with app tracing	23%	1	61%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus limit to four daily contacts with other individuals	Yes	All	90% school, 79% work, and 52% other	21%	0·93	64%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus app-based tracing plus limit to four daily contacts with other individuals	Yes	All	90% school, 79% work, and 52% other with manual tracing; 53% with app tracing	20%	0·87	66%
Mass testing of 5% of population per week	No	NA	NA	49%	2·5	2%

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus manual contact tracing of all contacts	Yes	All	100%	21%	0·94	64%
Self-isolation plus HHQ plus app-based tracing	Yes	All	53%	30%	1·4	47%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus app-based tracing	Yes	All	90% school, 79% work, and 52% other with manual tracing; 53% with app tracing	23%	1	61%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus limit to four daily contacts with other individuals	Yes	All	90% school, 79% work, and 52% other	21%	0·93	64%
Self-isolation plus HHQ plus manual contact tracing of acquaintances plus app-based tracing plus limit to four daily contacts with other individuals	Yes	All	90% school, 79% work, and 52% other with manual tracing; 53% with app tracing	20%	0·87	66%
Mass testing of 5% of population per week	No	NA	NA	49%	2·5	2%

Can we get to 1.7?

$R_t = 2.6$

- Asymptomatic: 30% of cases (1/2 transmission risk)

$R_t \sim 1.6$ among asymp
$R_t \sim 3.1$ among symp

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus manual contact tracing of all	Yes	All	100%	21%	0·94	64%

Can we get to 1.7?

$R_t = 2.6$

- Asymptomatic: 30% of cases (1/2 transmission risk)

$R_t \sim 1.6$ among asymp
$R_t \sim 3.1$ among symp

- Symptomatic:
  - 90% isolate
  - Pre-isolation: ~1/2 of transmission

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus manual contact tracing of all contacts	Yes	All	100%	21%	0·94	64%

Can we get to 1.7?

$R_t = 2.6$

- Asymptomatic: 30% of cases (1/2 transmission risk)

$R_t \sim 1.6$ among asymp
$R_t \sim 3.1$ among symp

- Symptomatic:
  - 90% isolate
  - Pre-isolation: ~1/2 of transmission

Asymp: 0.3*1.6 +
Undetected symp: 0.7*0.1*3.1 +
Detected symp: 0.7*0.9*3.1*0.5 = 1.7

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus manual contact tracing of all contacts	Yes	All	100%	21%	0·94	64%

Can we get to 0.94?

Asymp: 0.3*1.6 +
Undetected symp: 0.7*0.1*3.1 +
Detected symp: 0.7*0.9*3.1**X**

But what is X?

We still cut transmission in half from isolation → 0.5

Contact tracing cuts transmission by a factor of 0.72 among contacts.
→ X = 0.5*0.28

This gets us **0.97!**

Step 3: Let's compare.

Table 3

Mean reduction in R_{eff} under different control measures

	Self-Isolation	Contact tracing	Non-HH contacts who are potentially traceable (%)	Cases who have R>1 (%)	R_{eff}	Mean reduction in R_{eff}
No control	No	No	NA	50%	2·6	0%
Self-isolation within home	Yes	No	NA	40%	1·8	29%
Self-isolation outside home	Yes	NA	NA	37%	1·7	35%
Self-isolation plus HHQ	Yes	HH	NA	35%	1·6	37%
Self-isolation plus HHQ plus work or school contact tracing	Yes	HH and work or school	100%	27%	1·2	53%
Self-isolation plus HHQ plus manual contact tracing of acquaintances	Yes	All	90% school, 79% work, and 52% other	26%	1·1	57%
Self-isolation plus HHQ plus	Yes	All	100%	21%	0·94	64%

Can we get to 1.1?

Asymp: 0.3*1.6 +
Undetected symp: 0.7*0.1*3.1 +
Detected symp: 0.7*0.9*3.1**X**

Contact tracing cut transmission by ~90% among pre-isolation contacts. → .5*.25

What fraction are traced?
1-(.1*11*.06+.5*17*.06)/(2*.2 + 11*.06+17*.06) ~.72
→ X = .5*(.72*.28 + .28)

This gets us **1.17!**

We felt comfortable telling policymakers to focus on community testing and tracing rates to predict contact tracing effectiveness.

So, now let's try it with bunch of complicated models averaged together.

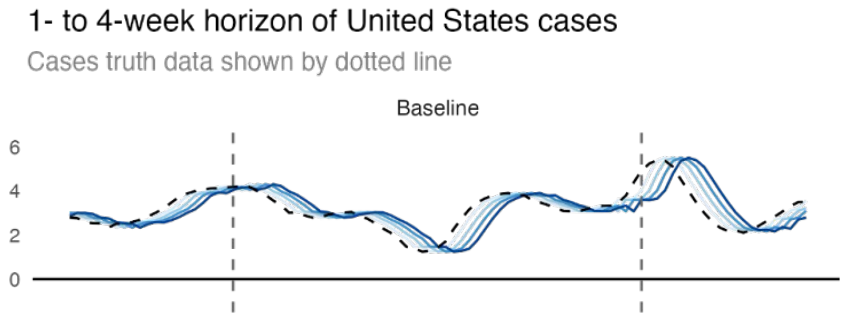
COVID ForecastHub ensemble

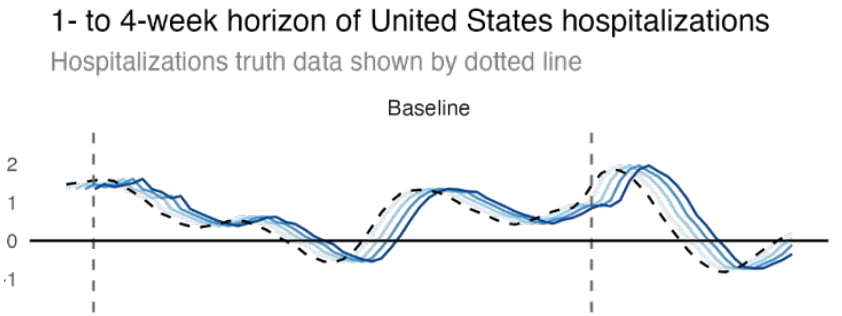
COVID-19 **ForecastHub**

Ensemble models average together forecasts of cases, hospitalizations, and deaths submitted by teams.

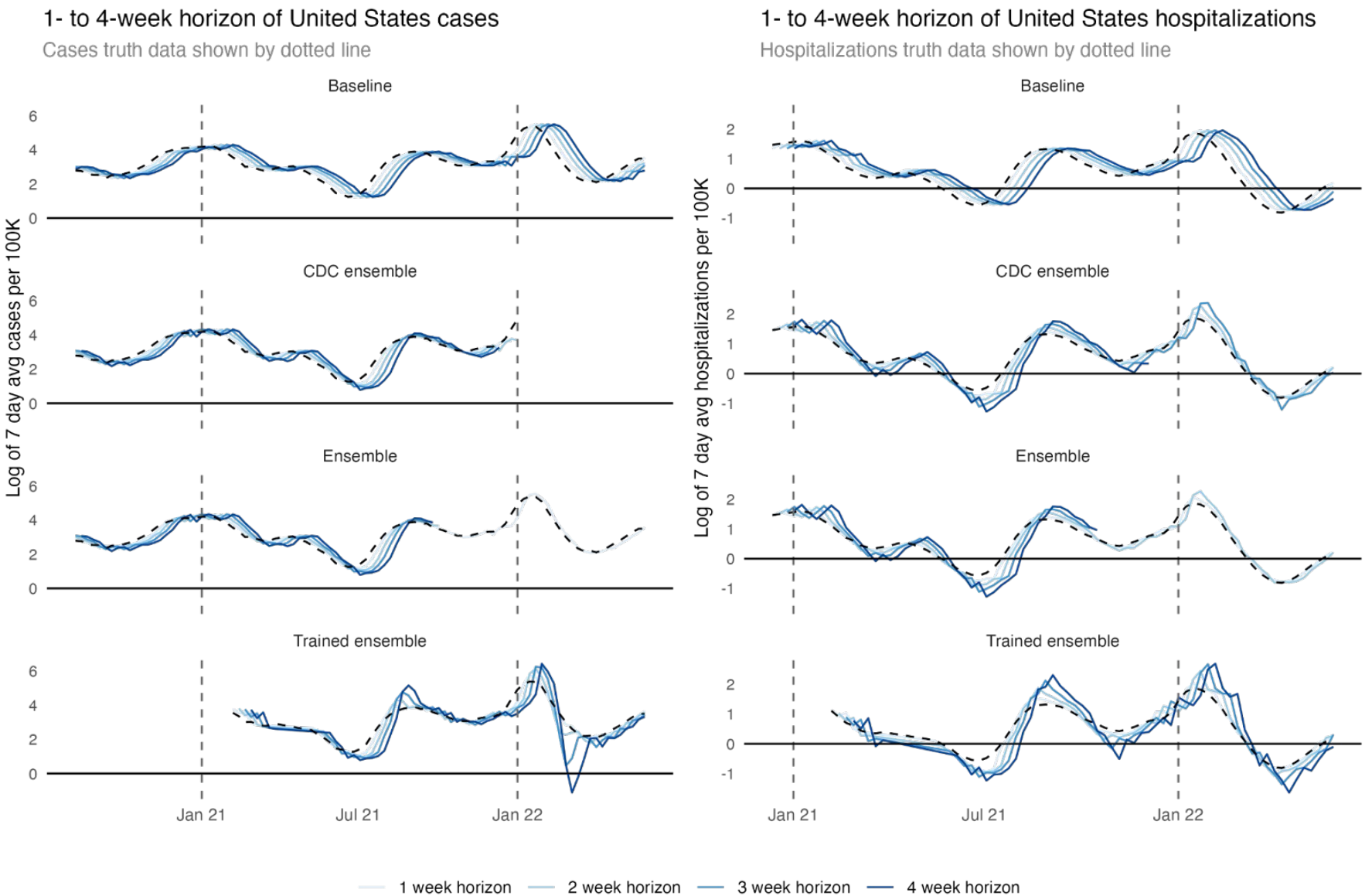
What is the simplest thing ensemble models could be doing?

How do simple and complex models compare?

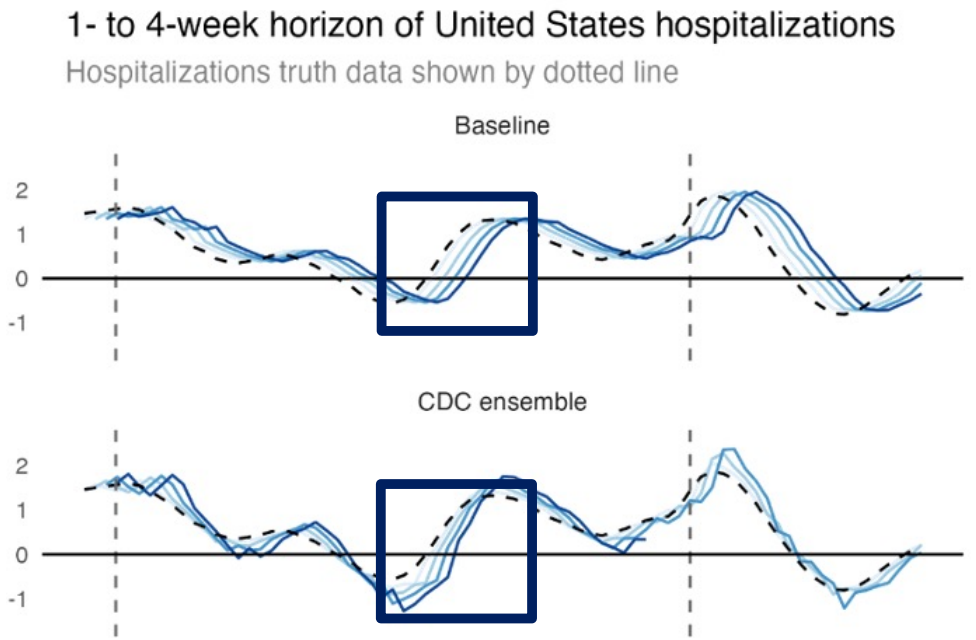




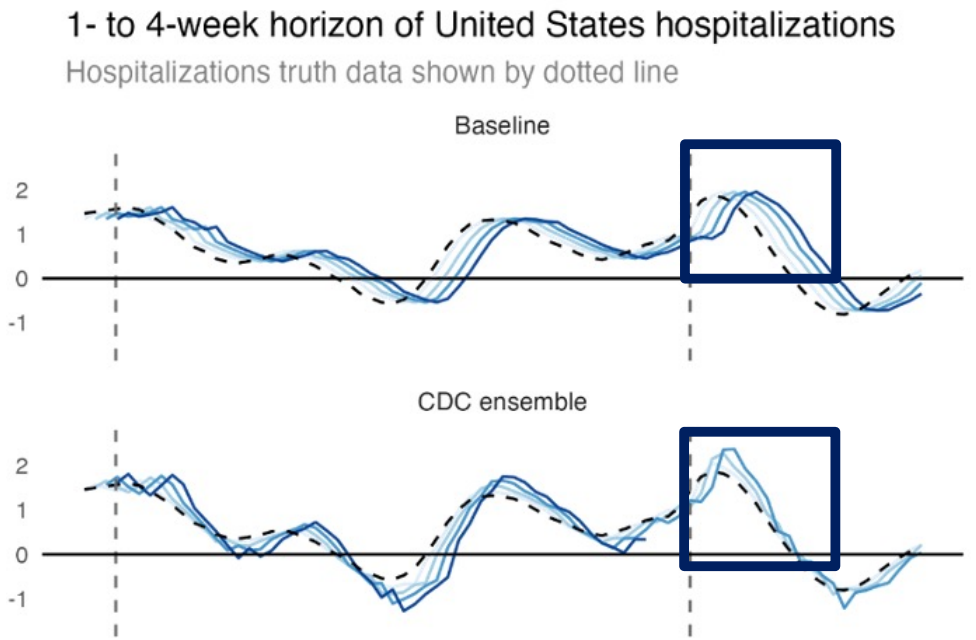
How do simple and complex models compare?



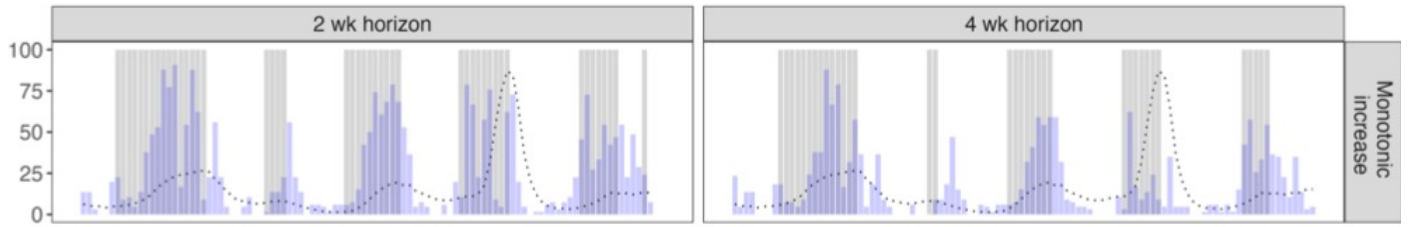
How do simple and complex models compare?

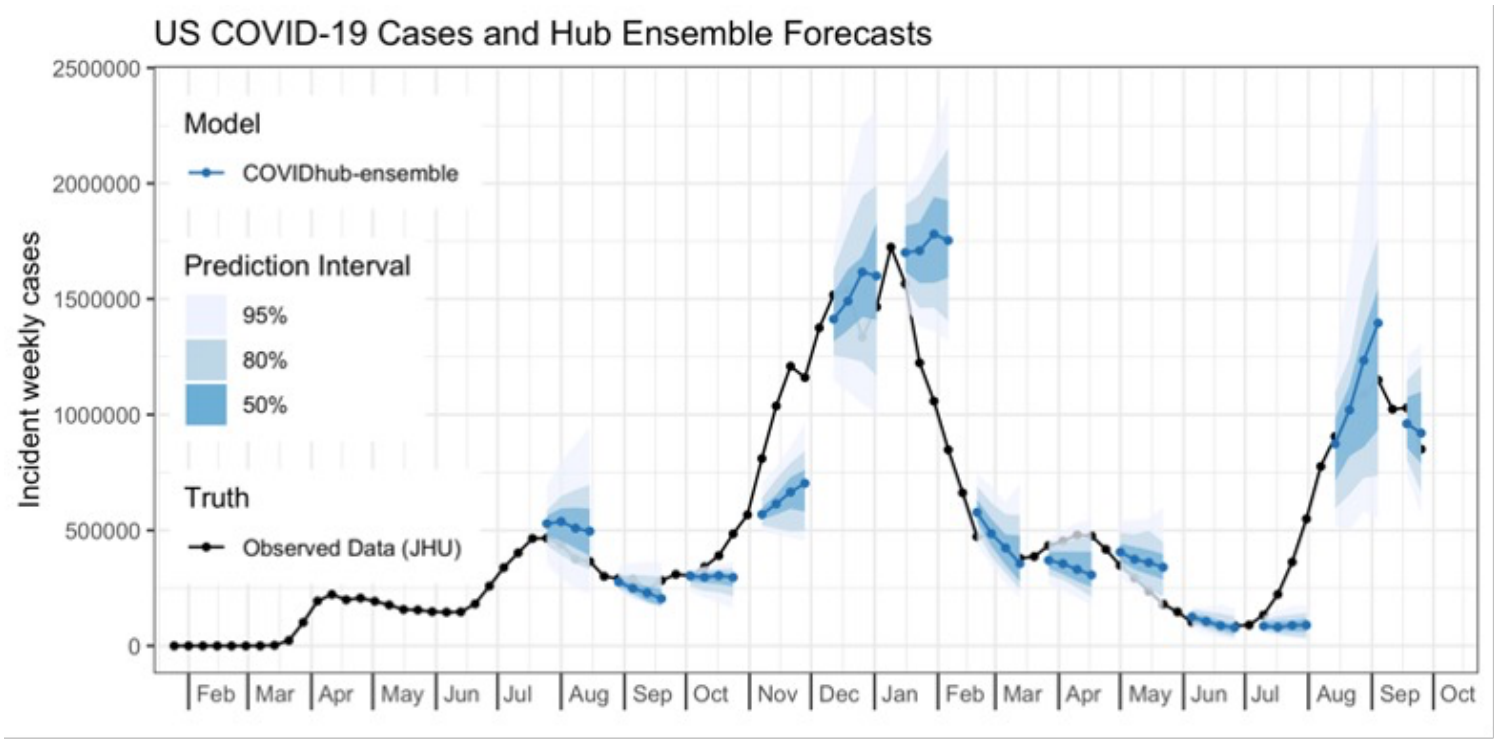


How do simple and complex models compare?



Helps us understand how models are likely to struggle





There is a strong temptation to evaluate decisions based on outcomes.

There is a strong temptation to evaluate decisions based on outcomes.

Quality of decision vs. quality of outcome

		Quality of Outcome	
		Good	**Bad**
Quality of Decision	**Good**	You decided well and did well	You were unlucky
	Bad	You were lucky	You decided poorly and did poorly

Napkin math for decision evaluation

1. What did (or could) you have known when making the decision?
2. What was (or should have been) your objective function?
3. What was the expected value of your decision?

1976 swine flu (H1N1)

- In 1976, there was a swine flu (H1N1) outbreak on a military base that killed 1 and hospitalized 13.
- President Ford announced a plan to vaccinate every in the country.
- 43 million individuals were vaccinated in 10 weeks.
- Some (perhaps a few hundred) experienced Guillain-Barre Syndrome (1/100K increased risk).

And no pandemic materialized...

1976 swine flu (H1N1)

Swine Flu Fiasco

By Harry Schwartz

The sorry debacle of the swine flu vaccine program provides a fitting end point to the misunderstandings and misconceptions that have marked Government approaches to health care during the last eight years, when Washington power has been shared between a Republican White House and a Democratic Congress.

Last February and March, on the flimsiest of evidence, President Ford and the Congress were panicked into believing that the country stood at the threshold of a killer flu epidemic, one that might claim millions of lives as did the much-cited influenza pandemic of 1918-1919.

Today, there is no sign whatsoever of anything approaching a swine flu epidemic; but there is growing apprehension that the millions of dollars of Federal money spent and the vast vaccination program pushed with all of Washington's energies may have resulted in the death of some persons and sickened many more. In short, there seem to have been significant costs without any visible benefits.

Any reasonable effort to assign responsibility for this state of affairs knows comparatively little about the origin and spread of influenza epidemics. In a sense, the Public Health Service and the Center for Disease Control reacted as the Pentagon tends to do. Both health agencies assumed the worst that could happen and urged action on that worst assumption, just as the Pentagon traditionally wants to have forces capable of fighting three major wars simultaneously.

● The self-interest of the Government health bureaucracy, which saw in the swine flu threat the ideal chance to impress the nation with the capabilities of saving money and lives by preventing disease. The Center for Disease Control in particular has long wanted to increase the size of its empire and multiply its budget by becoming the Government center for health education and disease prevention. Funds used for that purpose inevitably take money away from those whose job is actually to treat sick people. But the potentials of health education and disease prevention are still unproved—and perhaps only moderate at best.

It is possible, of course, that the country will still have a swine flu epidemic. But more and more expert opinion is shifting to the idea that such an epidemic, if it comes at all,

The Presidential Public Health Failure History Forgot

In the 1970s, the federal government attempted to vaccinate every American against the swine flu. It did not go well.

By Jeffrey Young

All medicines can cause side effects, including vaccines. Immunization programs like those for polio and smallpox are designed with that in mind, and with an aim toward protecting far more people than they harm. In the case of swine flu, there were only risks and no benefits, because an outbreak never occurred.

IMMUNE RESPONSE | VACCINE

The fiasco of the 1976 ‘swine flu affair’

Stepping back…

$$Expected\ value = P(Pandemic) * Value\ of\ vaccination\ in\ pandemic + P(No\ pandemic) * Value\ of\ vaccination\ without\ pandemic$$

$$Expected\ value = 1\% * 300K\ lives\ saved\ \left(\frac{1}{2}\ COVID\ deaths\ scaled\ to\ 1976\ population\right) + 99\% * Value\ of\ vaccination\ without\ pandemic$$

$$Expected\ value = 3bn\ (assuming\ \$10m\ VSL) + 99\% * Value\ of\ vaccination\ without\ pandemic$$

$$Expected\ value = 3bn\ (assuming\ \$10m\ VSL) - 99\% * 38M\ (Guillan - Barre\ syndrome)$$

We shouldn't think of this as a failure!!

Expect to be wrong sometimes…and pick your direction.

Adaptive metrics for an evolving pandemic: A dynamic approach to area-level COVID-19 risk designations

Alyssa M. Bilinski[a,1], Joshua A. Salomon[b], and Laura A. Hatfield[c]

States

	Neutral				Don't cry wolf (0.5x FN)				Better safe than sorry (0.5x FP)				
	Training	Training MR	Test	Test MR	Training	Training MR	Test	Test MR	Training	Training MR	Test	Test MR	
Adaptive: CHO	88	3	80	10	87	5	86	2	90	3	75	14	>1 death/100K/wk
Adaptive: CHOZ	88	3	83	5	88	5	87	2	90	2	80	10	
Adaptive: CHOD	86	6	77	17	85	9	84	4	87	5	74	16	
Adaptive: HZ	89	1	83	5	89	3	87	1	91	3	81	8	
Simplified adaptive: HZ	86	6	82	8	85	12	84	5	89	5	82	4	
Community Levels	64	44	71	24	76	28	72	29	53	62	70	24	
Z	80	15	81	8	81	19	79	12	80	25	83	3	
CHO	86	7	68	41	87	5	63	58	84	11	73	22	
HO	85	7	68	41	87	6	63	58	84	11	73	22	
CH	87	5	56	52	86	9	47	67	90	3	68	26	
H	83	15	56	52	84	19	68	45	88	8	69	26	
C	86	5	45	60	86	9	41	67	90	2	62	33	
Prevalence	68		41		68		41		68		41		

When unpacking any decision…

Quality of decision vs. quality of outcome

		Quality of Outcome: Good	Quality of Outcome: Bad
Quality of Decision	Good	You decided well and did well	You were unlucky
Quality of Decision	Bad	You were lucky	You decided poorly and did poorly

More likely in high-uncertainty contexts!!

Conclusions

Takeaways

Modelers

1. Especially for **policy decisions**, start with the **simplest model**.

2. Even if you don't stay there, **benchmark** your complex model **against a baseline** – both for evaluating performance and for interpretability.

3. Teaching napkin math – and **building model intuition** – is just as important as teaching high-performance computing.

Takeaways

Policymakers

1. Ask questions about **methods** and **mechanisms**, rather than just results.

2. Understand key sensitivities.

For Both

1. Good policy modeling **starts a conversation** rather than ending it.

And one more…

How often do you see statements like …

"We need more research"

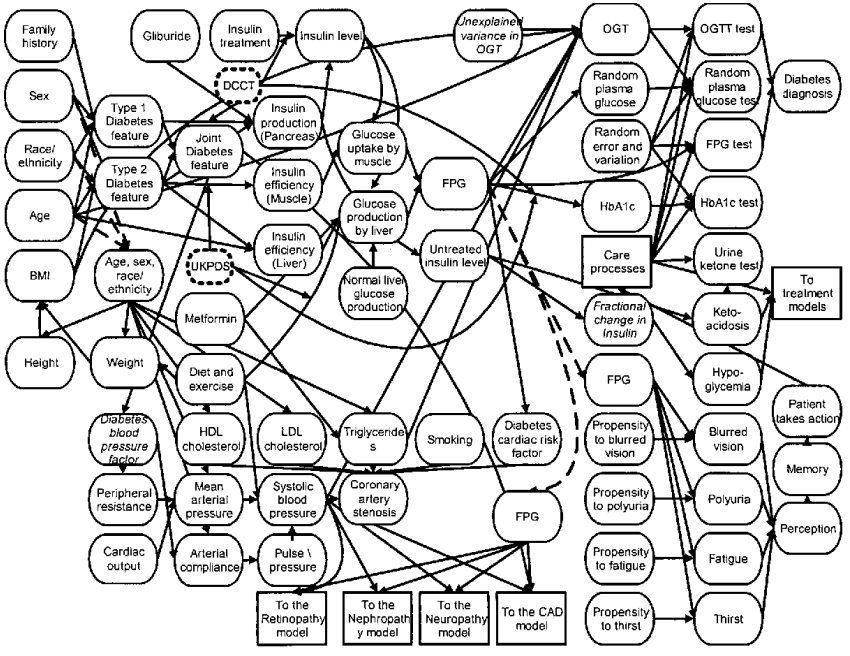
"There is insufficient evidence to draw definitive conclusions"

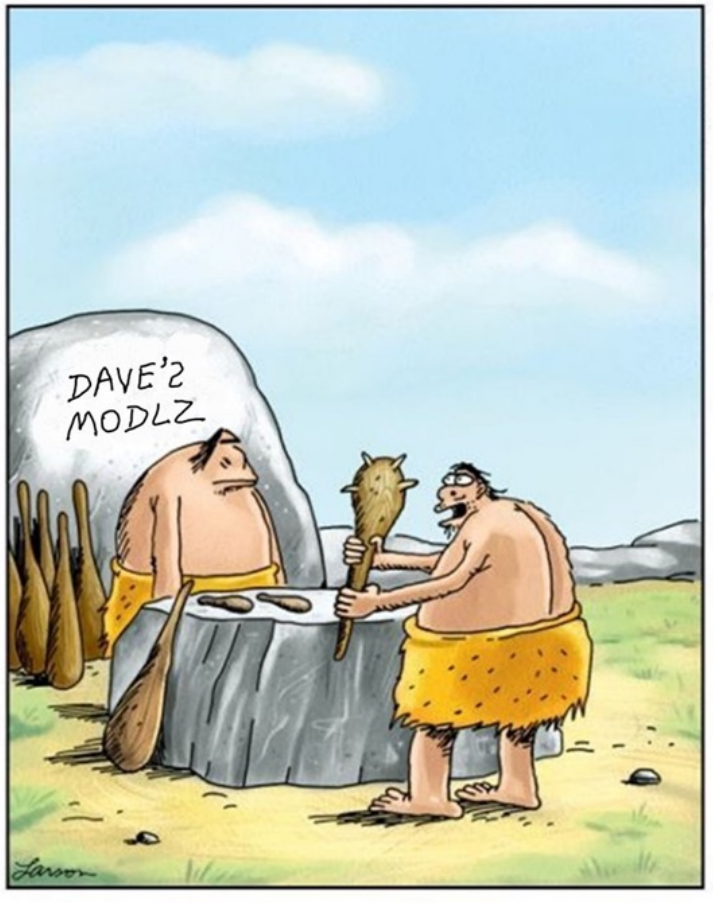
Complexity is neither a license to do nothing…

A favorite means of escaping the solution to any problem is to declare it too complex for solution. This absolves us from attempting solution…

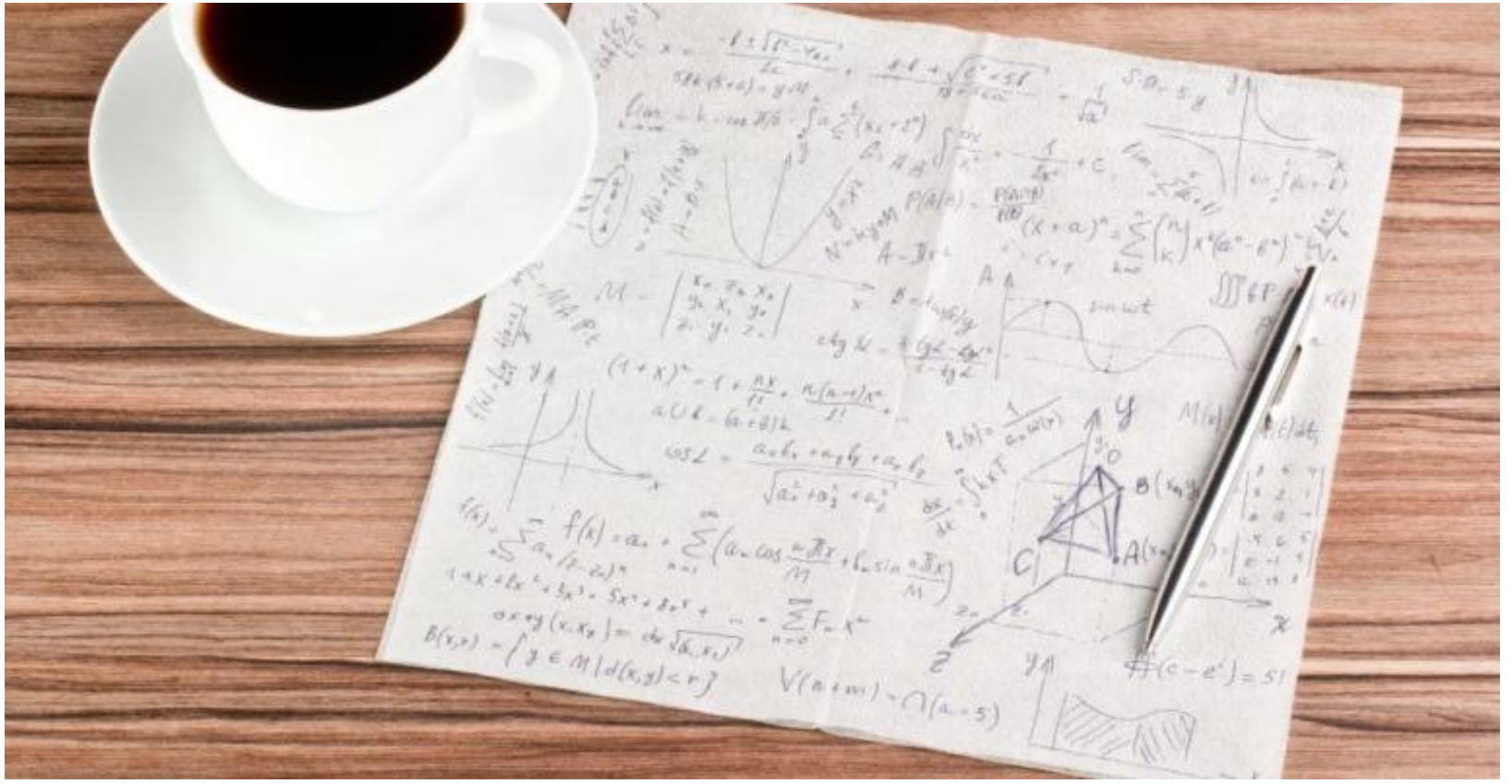
–Pearl S. Buck

…nor an excuse to over-specify

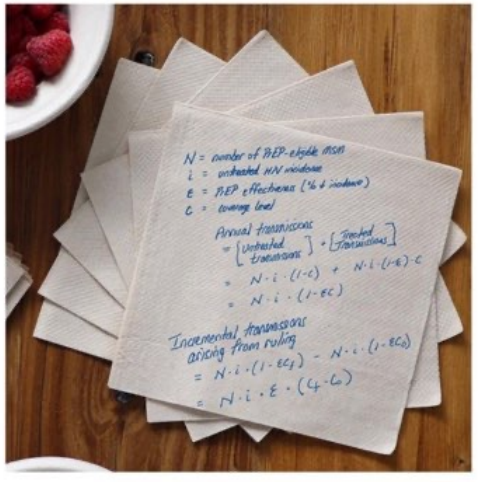


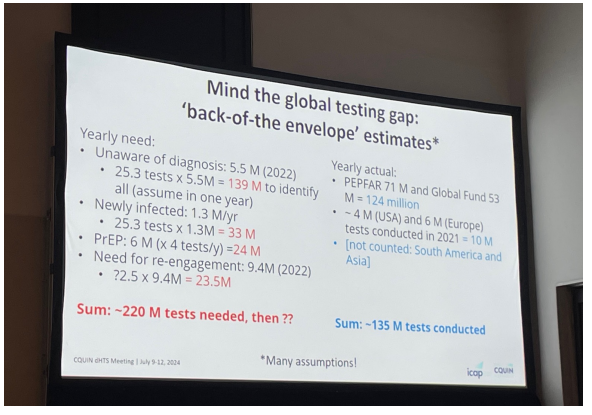


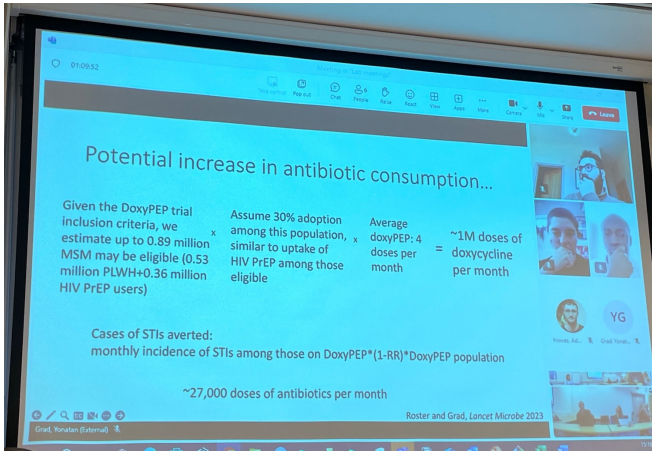
"No, no. ... Not this one. Too many bells and whistles."



We can calibrate investigations and our synthesis of the evidence – without sacrificing scientific rigor – to the needs they are intended to serve. Complex models help us learn about the world, but sometimes a little **napkin math** can help us with this task. Indeed, sometimes a napkin is all that is needed.


N = number of PrEP-eligible MSM
i = untreated HIV incidence
E = PrEP effectiveness (% ↓ incidence)
C = coverage level
Annual transmissions
= [Untreated transmissions] + [Treated transmissions]
= N·i·(1-C) + N·i·(1-E)·C
= N·i·(1-EC)
Incremental transmissions averted from ruling
= N·i·(1-EC1) - N·i·(1-EC0)
= N·i·E·(C1-C0)


Mind the global testing gap:
'back-of-the envelope' estimates*
Yearly need:
Unaware of diagnosis: 5.5 M (2022)
25.3 tests x 5.5M = 139 M to identify all (assume in one year)
Newly infected: 1.3 M/yr
25.3 tests x 1.3M = 33 M
PrEP: 6 M (x 4 tests/y) =24 M
Need for re-engagement: 9.4M (2022)
?2.5 x 9.4M = 23.5M
Sum: ~220 M tests needed, then ??
Yearly actual:
PEPFAR 71 M and Global Fund 53 M = 124 million
~ 4 M (USA) and 6 M (Europe) tests conducted in 2021 = 10 M
[not counted: South America and Asia]
Sum: ~135 M tests conducted
*Many assumptions!
icap
CQUIN


Potential increase in antibiotic consumption...
Given the DoxyPEP trial inclusion criteria, we estimate up to 0.89 million MSM may be eligible (0.53 million PLWH+0.36 million HIV PrEP users)
x
Assume 30% adoption among this population, similar to uptake of HIV PrEP among those eligible
x
Average doxyPEP: 4 doses per month
=
~1M doses of doxycycline per month
Cases of STIs averted:
monthly incidence of STIs among those on DoxyPEP*(1-RR)*DoxyPEP population
~27,000 doses of antibiotics per month
Roster and Grad, Lancet Microbe 2023

Questions?

How did models do?

Strengths

- *Shed light on "what if" scenarios in time-sensitive contexts when "doing nothing" was a risky choice*
- *Qualitative insights*

16 March 2020 *Imperial College COVID-19 Response Team*

Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand

Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani.

SCIENCE ADVANCES | RESEARCH ARTICLE

CORONAVIRUS

Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening

Daniel B. Larremore[1,2*], Bryan Wilder[3], Evan Lester[4,5], Soraya Shehata[5,6], James M. Burke[4], James A. Hay[7,8], Milind Tambe[3], Michael J. Mina[7,8,9*†], Roy Parker[2,4,6,10*†]

JAMA Network | Open

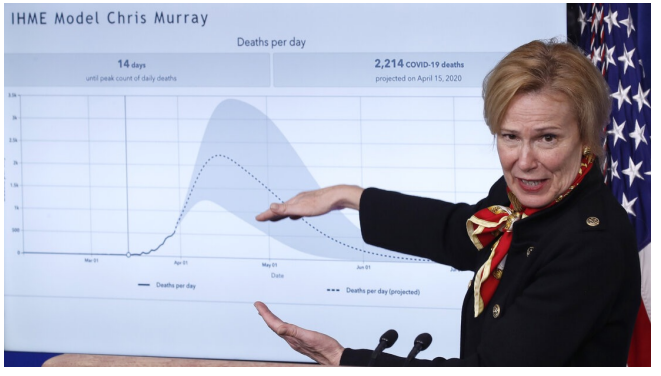
Original Investigation | Public Health

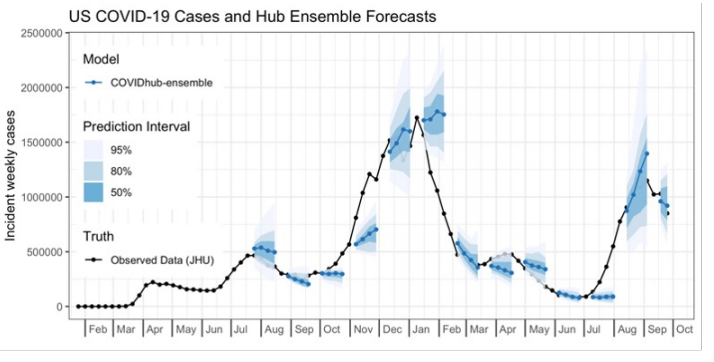
Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States

A. David Paltiel, PhD; Amy Zheng, BA; Rochelle P. Walensky, MD, MPH

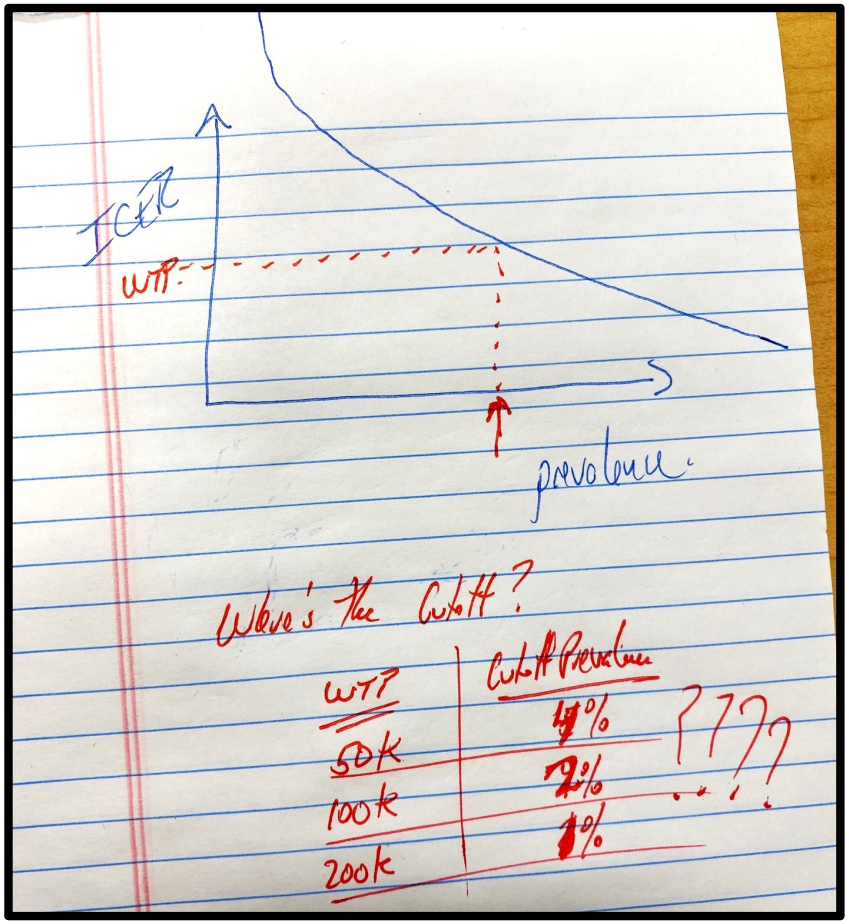
Weaknesses

- *Wrong in important ways!*



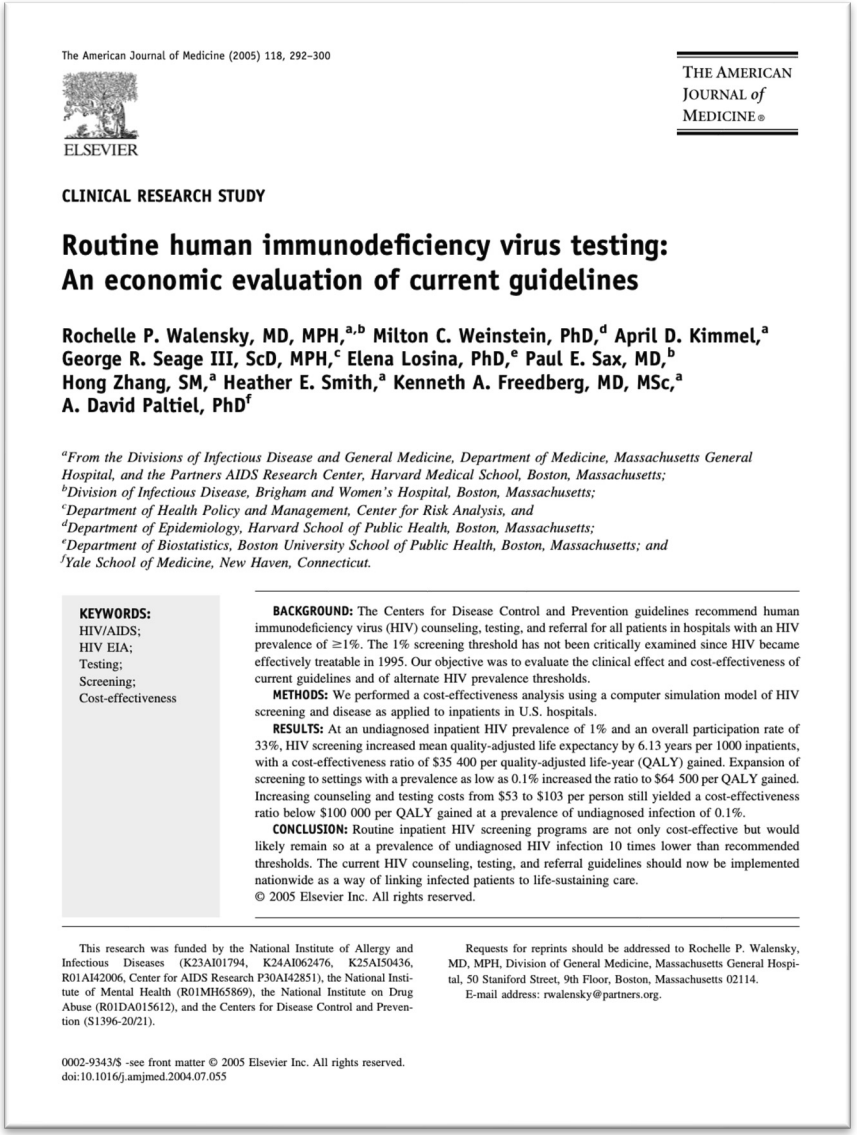


Preliminary sketches do not always pan out



Routine testing for HIV

- A model-based assessment of the cost-effectiveness of expanded HIV testing in the US.
- Key question: How does the cost-effectiveness estimate vary as a function of the prevalence of undetected HIV in the population?



The American Journal of Medicine (2005) 118, 292–300

ELSEVIER

THE AMERICAN JOURNAL of MEDICINE®

CLINICAL RESEARCH STUDY

Routine human immunodeficiency virus testing: An economic evaluation of current guidelines

Rochelle P. Walensky, MD, MPH,[a,b] Milton C. Weinstein, PhD,[d] April D. Kimmel,[a] George R. Seage III, ScD, MPH,[c] Elena Losina, PhD,[e] Paul E. Sax, MD,[b] Hong Zhang, SM,[a] Heather E. Smith,[a] Kenneth A. Freedberg, MD, MSc,[a] A. David Paltiel, PhD[f]

[a]*From the Divisions of Infectious Disease and General Medicine, Department of Medicine, Massachusetts General Hospital, and the Partners AIDS Research Center, Harvard Medical School, Boston, Massachusetts;*
[b]*Division of Infectious Disease, Brigham and Women's Hospital, Boston, Massachusetts;*
[c]*Department of Health Policy and Management, Center for Risk Analysis, and*
[d]*Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts;*
[e]*Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; and*
[f]*Yale School of Medicine, New Haven, Connecticut.*

KEYWORDS:
HIV/AIDS;
HIV EIA;
Testing;
Screening;
Cost-effectiveness

BACKGROUND: The Centers for Disease Control and Prevention guidelines recommend human immunodeficiency virus (HIV) counseling, testing, and referral for all patients in hospitals with an HIV prevalence of ≥1%. The 1% screening threshold has not been critically examined since HIV became effectively treatable in 1995. Our objective was to evaluate the clinical effect and cost-effectiveness of current guidelines and of alternate HIV prevalence thresholds.

METHODS: We performed a cost-effectiveness analysis using a computer simulation model of HIV screening and disease as applied to inpatients in U.S. hospitals.

RESULTS: At an undiagnosed inpatient HIV prevalence of 1% and an overall participation rate of 33%, HIV screening increased mean quality-adjusted life expectancy by 6.13 years per 1000 inpatients, with a cost-effectiveness ratio of $35 400 per quality-adjusted life-year (QALY) gained. Expansion of screening to settings with a prevalence as low as 0.1% increased the ratio to $64 500 per QALY gained. Increasing counseling and testing costs from $53 to $103 per person still yielded a cost-effectiveness ratio below $100 000 per QALY gained at a prevalence of undiagnosed infection of 0.1%.

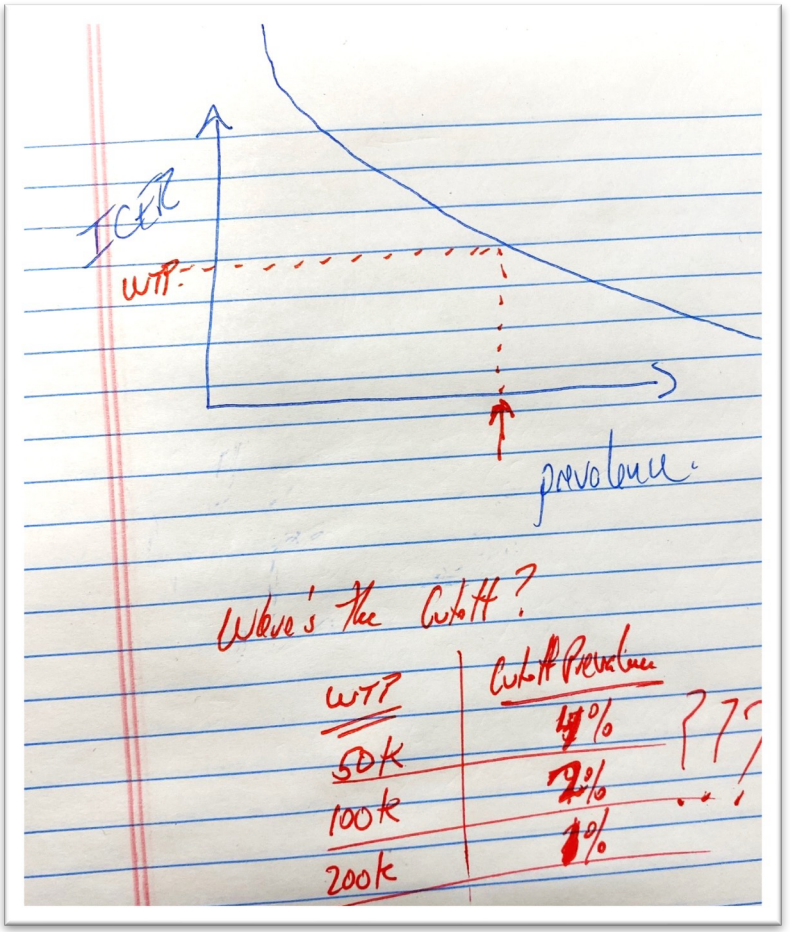
CONCLUSION: Routine inpatient HIV screening programs are not only cost-effective but would likely remain so at a prevalence of undiagnosed HIV infection 10 times lower than recommended thresholds. The current HIV counseling, testing, and referral guidelines should now be implemented nationwide as a way of linking infected patients to life-sustaining care.

This research was funded by the National Institute of Allergy and Infectious Diseases (K23AI01794, K24AI062476, K25AI50436, R01AI42006, Center for AIDS Research P30AI42851), the National Institute of Mental Health (R01MH65869), the National Institute on Drug Abuse (R01DA015612), and the Centers for Disease Control and Prevention (S1396-20/21).

Requests for reprints should be addressed to Rochelle P. Walensky, MD, MPH, Division of General Medicine, Massachusetts General Hospital, 50 Staniford Street, 9th Floor, Boston, Massachusetts 02114.
E-mail address: rwalensky@partners.org.

0002-9343/$ -see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjmed.2004.07.055

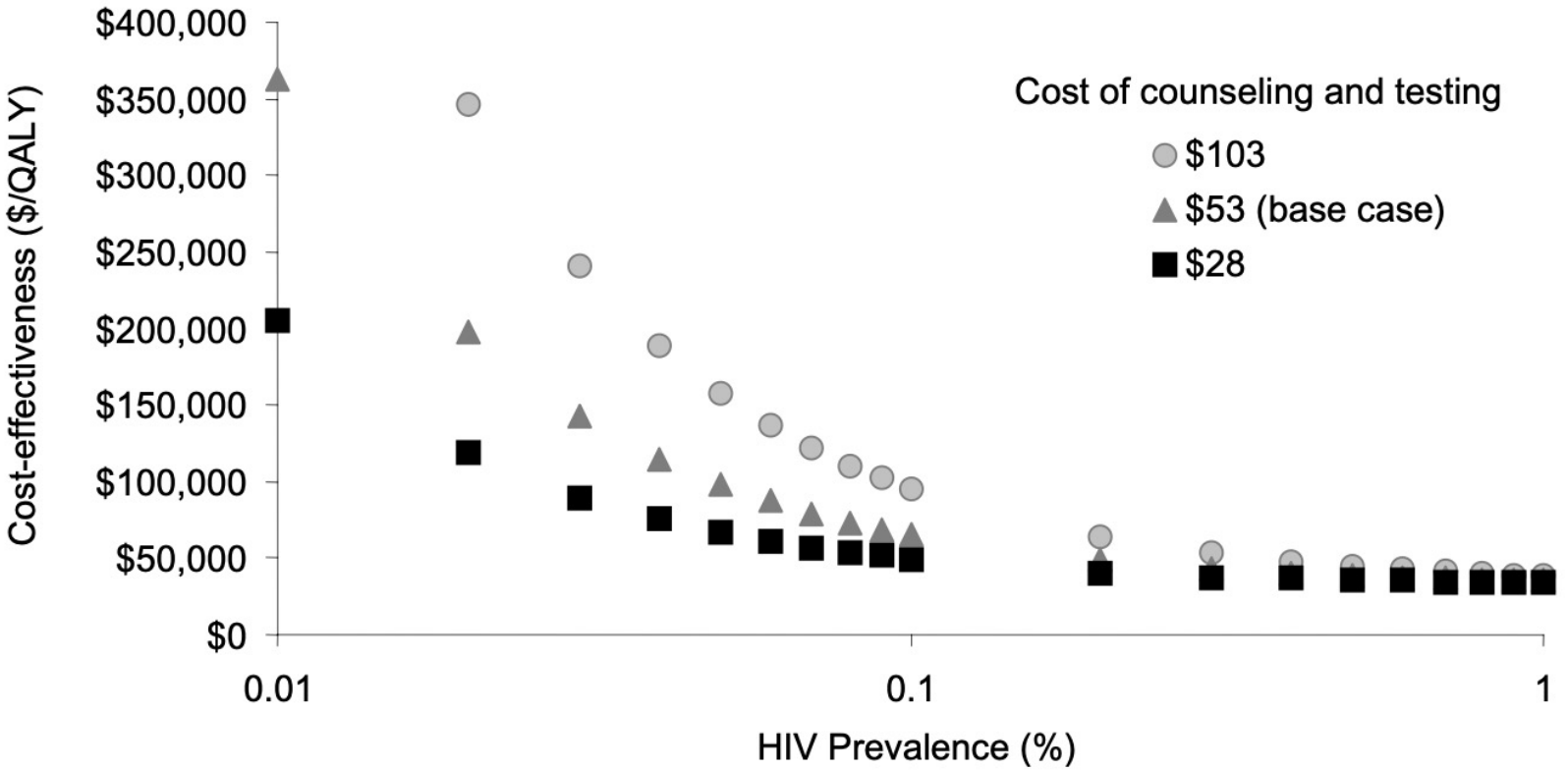
Before getting started



Hypotheses:

- As prevalence → 0
  ICER → ∞.
- As prevalence → 100%
  ICER → 0.
- In the "policy zone" ~linear returns, with cut-off prevalence in the range 1% to 4% depending on WTP threshold.

Study investigators were surprised by the results



… though they were consistent with other studies

Routine HIV Screening in France: Clinical Impact and Cost-Effectiveness

Yazdan Yazdanpanah[1,2,3*], Caroline E. Sloan[4], Cécile Charlois-Ou[6], Stéphane Le Vu[7], Caroline Semaille[3,7], Dominique Costagliola[8,9,10,11], Josiane Pillonel[7], Anne-Isabelle Poullié[12], Olivier Scemama[12], Sylvie Deuffic-Burban[13], Elena Losina[4,14,15], Rochelle P. Walensky[4,5,16,17], Kenneth A. Freedberg[4,5,14,17], A. David Paltiel[18]

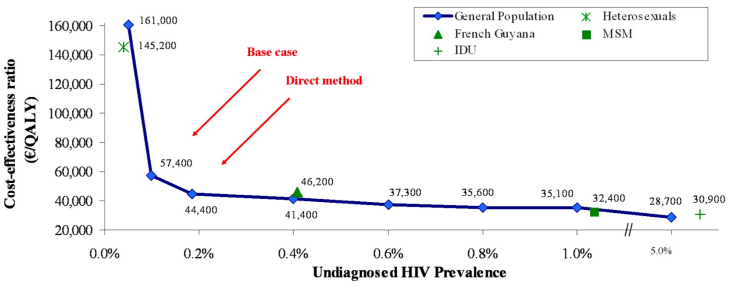
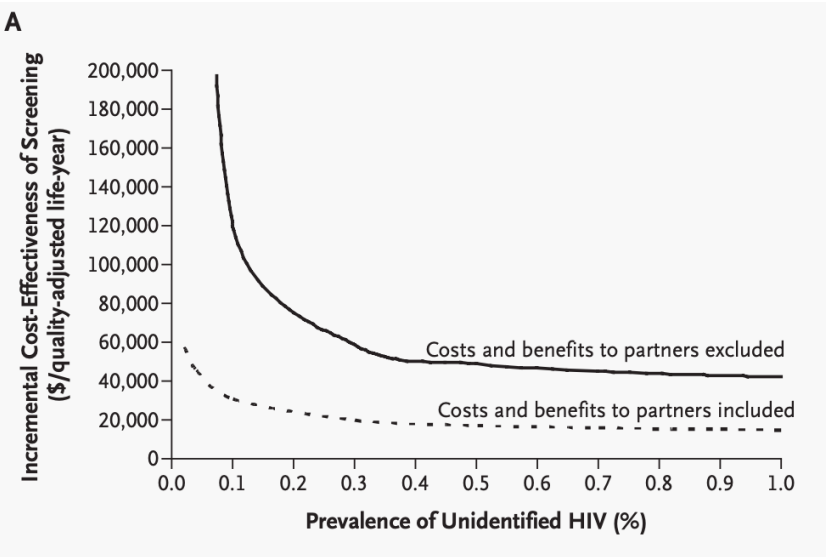


Figure 1. Effect of undiagnosed HIV prevalence on the cost effectiveness a one-time routine, voluntary HIV test vs. "current practice", with base case incidence. Incidence rates are as follows: general population, 0.01/100PY; heterosexuals, 0.01/100PY; French Guyana, 0.35/100PY; MSM, 0.99/100PY; and IDU, 0.17/100PY. MSM: men who have sex with men; IDU: injection drug users; PY: person-year.
doi:10.1371/journal.pone.0013132.g001

Cost-Effectiveness of Screening for HIV in the Era of Highly Active Antiretroviral Therapy

Gillian D. Sanders, Ph.D., Ahmed M. Bayoumi, M.D., Vandana Sundaram, M.P.H., S. Pinar Bilir, A.B., Christopher P. Neukermans, A.B., Chara E. Rydzak, B.A., Lena R. Douglass, B.S., Laura C. Lazzeroni, Ph.D., Mark Holodniy, M.D., and Douglas K. Owens, M.D.



Study investigators were surprised by the results