Statistical Methods and Modeling In Response to COVID-19 at NIAID

Dean Follmann

National Institute of Allergy and Infectious Diseases

5 May 2020

NIAID Biostatistics Research Branch


Dean Follmann


Erica Brittain


Sally Hunsberger


Michael Fay


Lori Dodd


Micheal Proschan


Jing Qin


Bruce Swihart


Misrak Gezmu


Zonghui Hu

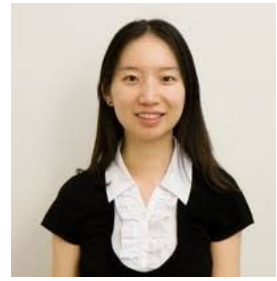

Jonathan Fintzi

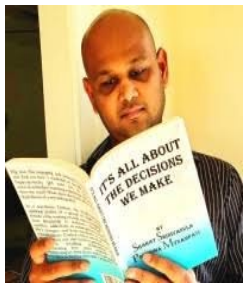

Ana Villa -Ortega


Michele Di Mascio


Martha Nason


C. Jason Liang


Jing Wang



Sharat Srinivasula

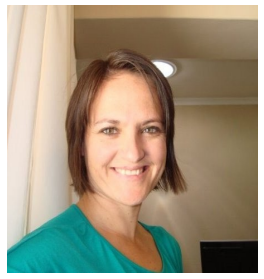

Keith Lumbard


Allyson Mateja


Shannon Gallagher


Alyhana Childs


Michael Duvenhage


Shera Weyers


Iman Gulati


Victoria Bera


Tyler Bonnett


Alyson Francis


Yi Chi


Claudine Casimir


Aurélie Gouel

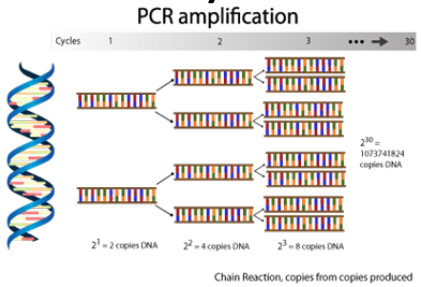

Erin Gabriel

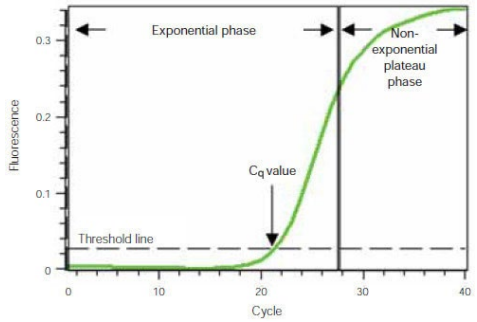
How we fight Infectious Diseases

- Identify
  - Diagnostic Tests
- Describe
  - Time from infection to symptoms (incubation)
  - Who has been infected (sero-prevalence)
  - Within household transmission
- Treatment
  - Adaptive Treatment Trial
- Prevention
  - Vaccines
  - Antibodies

Identify: Diagnostics

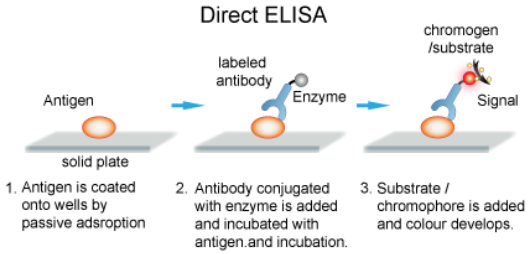
- Are you infected? (PCR tests)

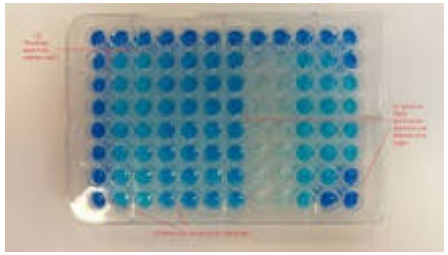




Y = abundance of SARS-CoV2 RNA

- Were you infected? (Serology)





Y = abundance of antibodies that fought SARS-CoV2

Exponential growth of testing?

- Ten Samples, Ten individual tests, Ten hours of work.

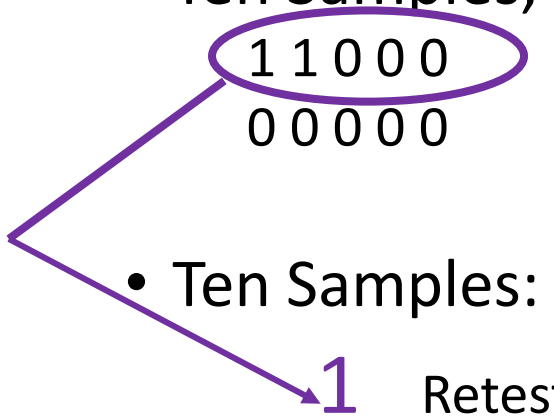
  1 1 0 0 0

  0 0 0 0 0

- 10, 20, 40, 80, 160, 320 test, 320 hours of work…

Exponential growth of testing?

- Ten Samples, Ten individual tests, Ten hours of work.

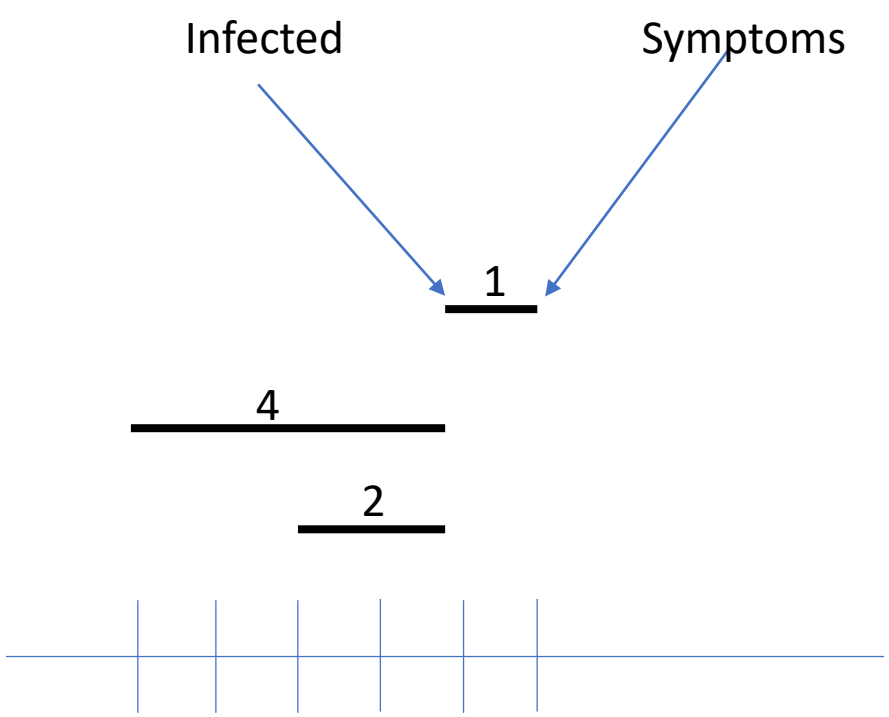


- Ten Samples: Two **pooled** tests. 5 individual tests. 7 hours of work.
  - 1 Retest individual samples 1 1 0 0 0
  - 0 *Infer individual samples* 0 0 0 0 0
- Concern
  - 1 positive in a big pool of negatives may be ‘drowned’ out
- Remedy
  - Do experiments

Current Work

- Collaborating with Asian Researchers to pool PCR tests
- Collaborating within NIAID to pool serologies
- Determine the pool size that still identifies one positive in the pool
- Labs choose number based on their current positivity rate
  - If rare maybe pool 10 or 20
  - If common don’t pool

Describe Incubation Distribution



$P(I = 2) = \Theta_2 = \#(I = 2) / \text{Total} = 1/3$

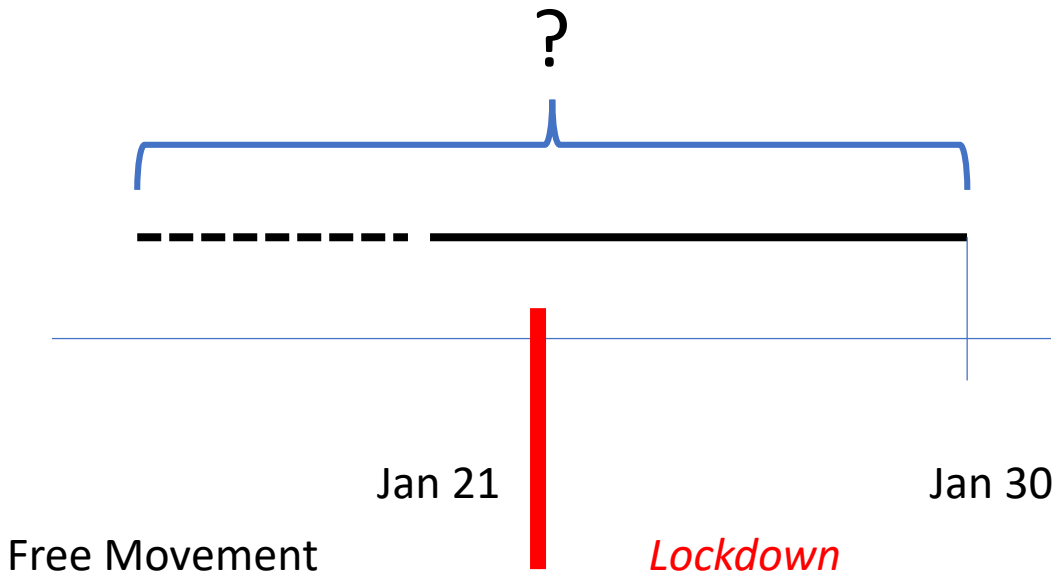
True Incubation Distribution

I	P(I)
1	Θ_1
2	Θ_2
3	Θ_3
4	Θ_4

If always $I < 14$ days, then quarantine for 14 days

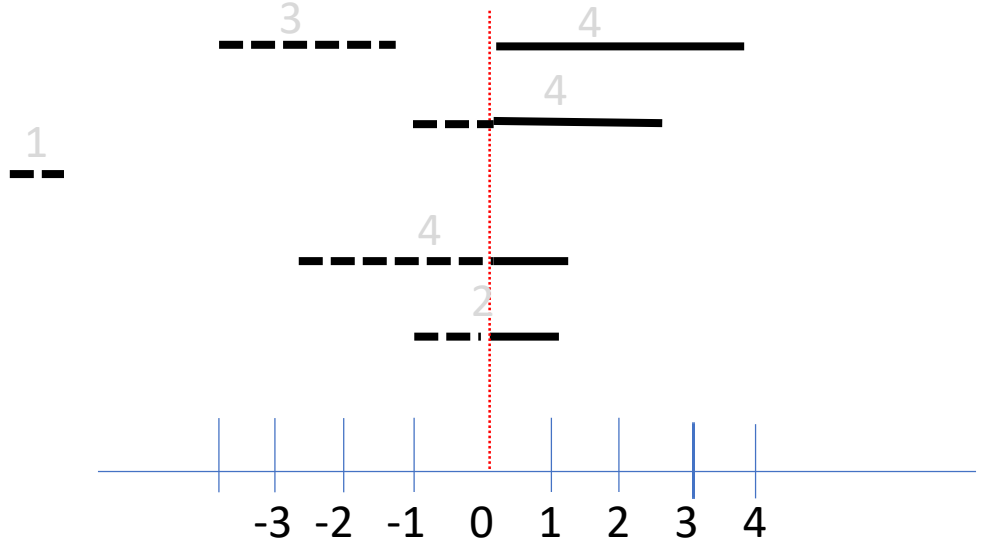
A natural experiment

- In January, Epidemic was mostly in Wuhan
- On Jan 23, China imposed a countrywide lockdown
- Suppose Zonghui leaves Wuhan on Jan 21 goes to Beijing. On Jan 23 she is stuck in her apartment
- Zonghui tests positive on Jan 30
  - Must’ve got it in Wuhan
  - Incubation must be at least 9 days
  - Can we do better?



Two issues, but a solution

- Incubation period as least what we see
- Wuhan emigres tend to nave longer incubations



True Incubation Distribution

I	P(I)
1	Θ_1
2	Θ_2
3	Θ_3
4	Θ_4

P(Symptoms day 4) = P(I=4 | Infected on Day 0) P(Infected on Day 0)/C
= Θ_4 (1/4)/C

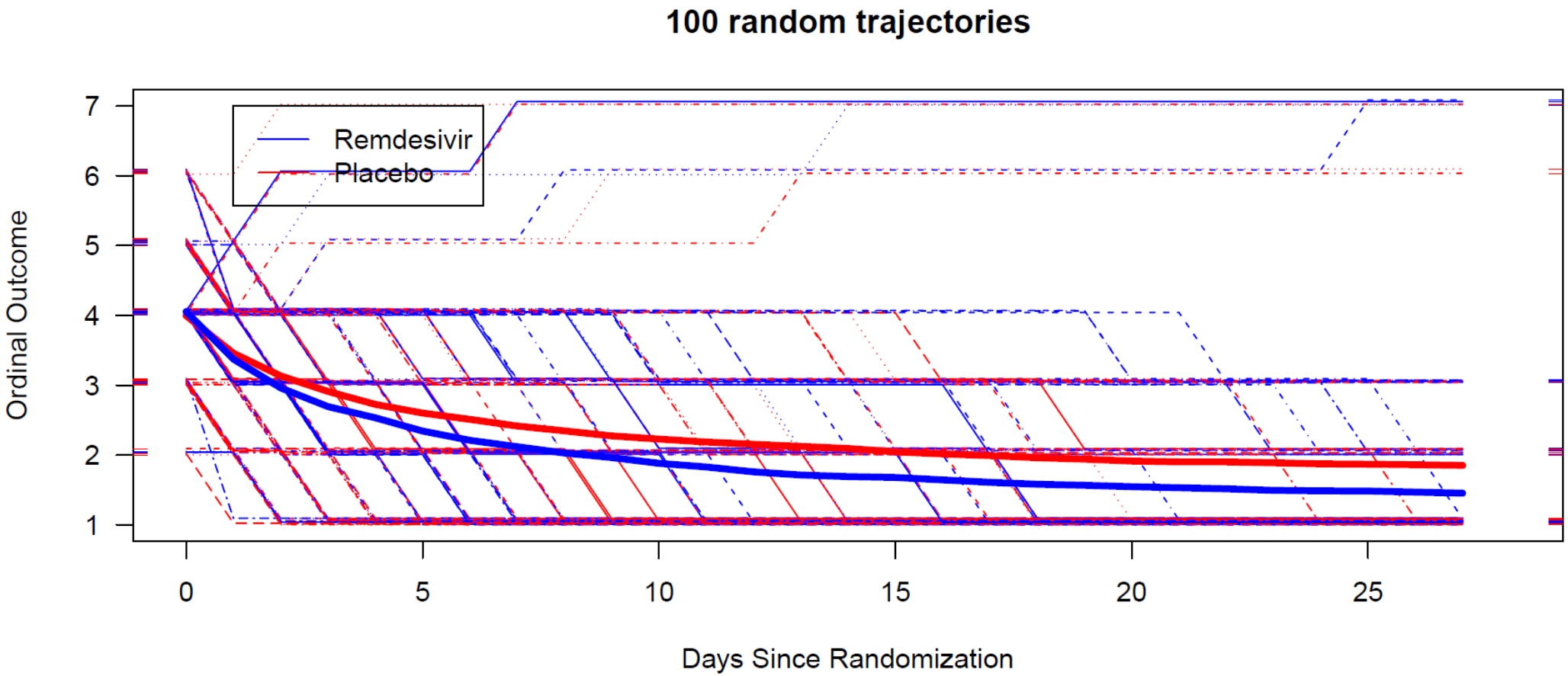
P(Symptoms day 3) = P(I=4 | Infected on Day -1) P(Infected on Day 0) + P(I=3 | Infected on Day 0) P(Infected on Day 0)
= {Θ_4 (1/4) + Θ_3 (1/4)}/C

Treatment: ACTT-1



- Double blind, adaptive, randomized trial of remdesivir vs placebo in mild-severe COVID-19 disease
- Measure ordinal scale every day. *Feels, functions, survives, … & logistics*
  - 8 Death;
  - 7 Hospitalized, on invasive mechanical ventilation or ECMO;
  - 6 Hospitalized, on non-invasive ventilation or high flow oxygen devices;
  - 5 Hospitalized, requiring supplemental oxygen;
  - 4 Hospitalized, not requiring supplemental oxygen - requiring ongoing medical care
  - 3 Hospitalized, not requiring supplemental oxygen - no longer requires ongoing medical care;
  - 2 Not hospitalized, limitation on activities and/or requiring home oxygen;
  - 1 Not hospitalized, no limitations on activities.
- Primary endpoint ordinal outcome at day 14, . . . But
  - How firm are the categories?
  - What if treatment effects show up later.
  - Blinded adaptation after pilot of 100 doesn't help much

Simulate trajectories, determine power



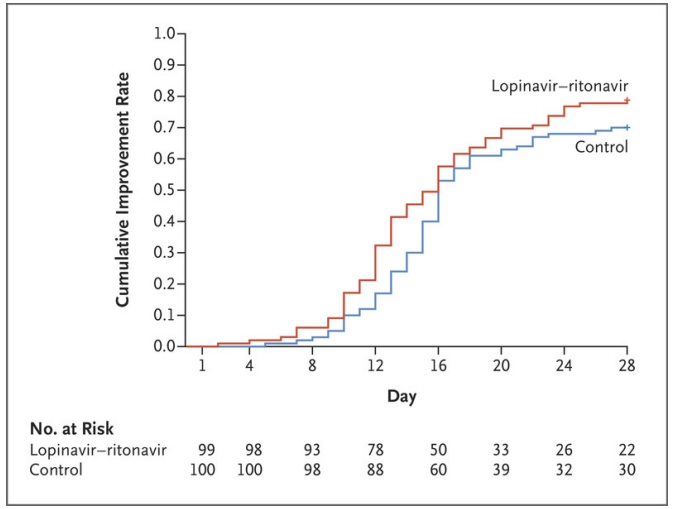
Test	P-Odds Day 1	P-Odds Day 7	P-Odds Day 14	P-Odds Day 28	Mean Score	Cox on 2 point	Cox on Recovery	Cox on Death	28 Day Mortality
Simple	.046	.755	.851	**.877**	.800	.808	.818	.626	.579
Adjusted					**.917**	.834	**.909**		

Power:
Proportion of times we conclude Remdesivir works for different tests

Endpoint: Time to Recovery

- How to treat deaths? People who die can never recover. Set their recovery times to infinity.
- Use log-rank test with time to recovery over day [0,28]
- Kaplan-Meier curve estimates the *cumulative incidence* of recoveries.
- Corresponds to Fine-Grey method for competing risks
- Designed to achieve 400 recoveries

Cao 2020 study of Kaletra



ACTT-1

- Study accrued extremely rapidly as epidemic exploded
  - Required quick flexible thinking/action
- First interim look = final look had more than 400 recoveries
- Study well powered and well run
  - 31% faster recovery p-value = 0.001
  - Median recovery 11 days Remdesivir vs 15 days Placebo
  - Mortality 8% Remdesivir 11% Placebo
- ACTT --- a powerful collaboration poised to keep getting answers
  - ACTT-2 being finalized given ACTT-1 results

Prevention: Planning for Vaccines

- Many vaccine candidates to be evaluated
- Possible Endpoints
  - Infection - PCR+ for virus
  - Disease - PCR+ for virus & symptomatic disease
- Which will be sensitive to vaccine effects? Help patient health?
- What if major effect of vaccine is to lessen disease severity?

Outcome	Placebo	Vaccine
Infection	1.0%	1.5%
Mild Disease	0.6%	0.4%
Hospitalized	0.3%	0.1%
Death	0.1%	0.0%

Vaccine "lift"

Prevention: Planning for Vaccines

- Many vaccine candidates to be evaluated
- Possible Endpoints
  - Infection - PCR+ for virus
  - Disease - PCR+ for virus & symptomatic disease
- Which will be sensitive to vaccine effects? Help patient health?
- What if major effect of vaccine is to lessen disease severity?

Outcome	Placebo	Vaccine
Infection	1.0%	1.5%
Mild Disease	0.6%	0.4%
Hospitalized	0.3%	0.1%
Death	0.1%	0.0%

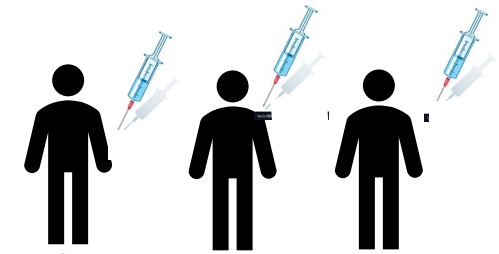
Potential Analyses

- Time to disease
  - Use proportional hazards model
  - % Reduction in the instantaneous of probability of disease
- Time to weighted disease: mild=1, hosp=2, death=3
  - Proportional means model
  - % Reduction in the mean severity of disease
- Simulations being conducted to inform choice
  - Usual Wilcoxon 66% power
  - Proportional Hazards 81% power
  - Proportional Means 93% power

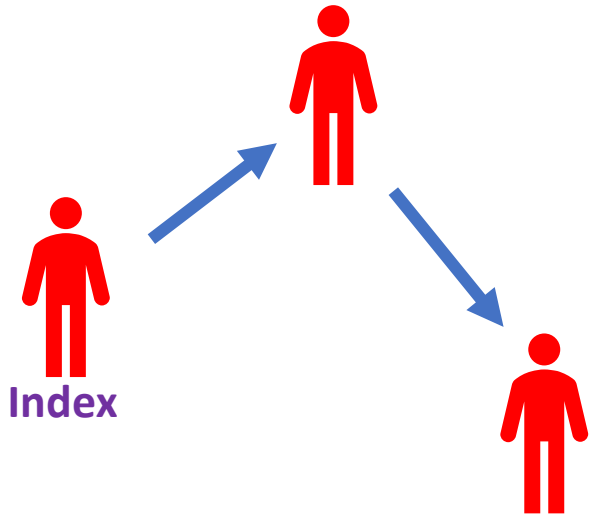
Outcome	Placebo	Vaccine
Infection	1.0%	1.5%
Mild Disease	0.6%	0.4%
Hospitalized	0.3%	0.1%
Death	0.1%	0.0%

Prevention: Antibodies

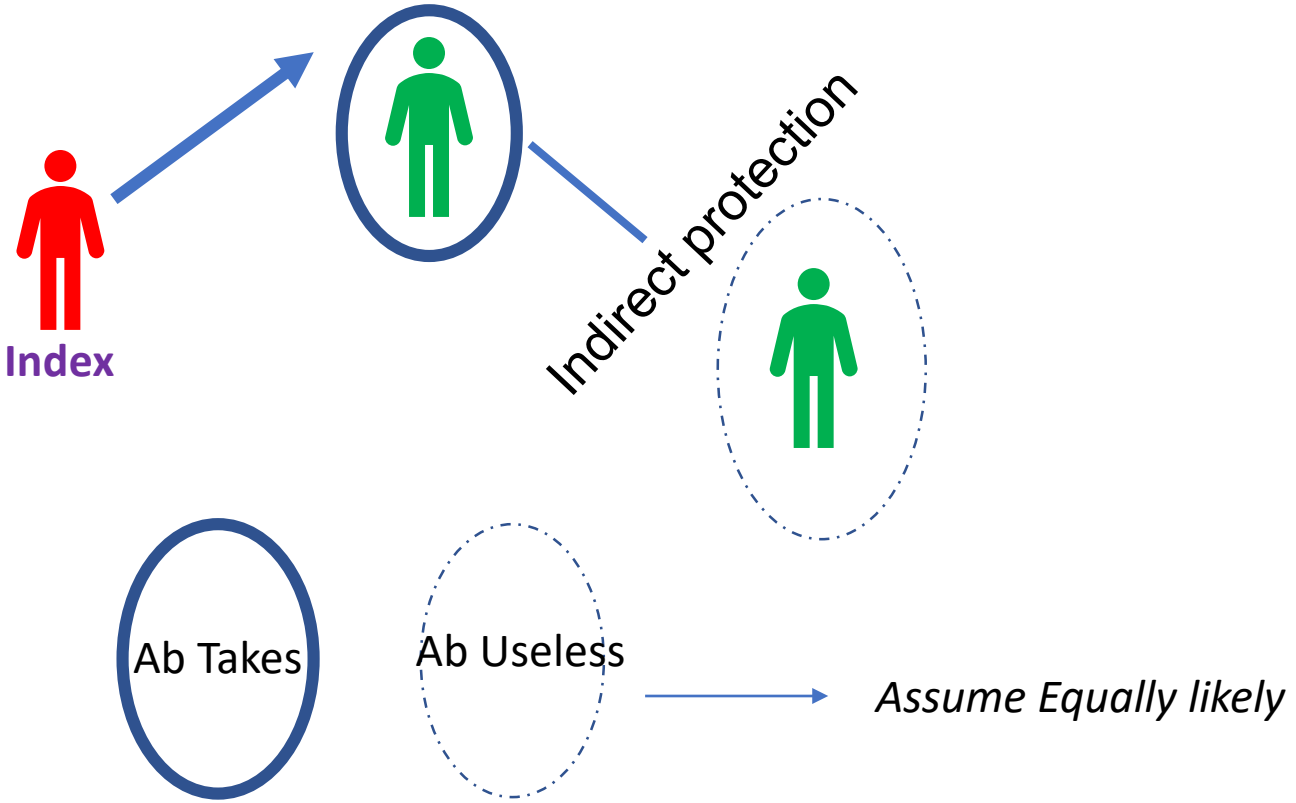
- Injected SARS-CoV2 antibodies might prevent infection
- Identify **someone** with COVID-19
- Enroll them and their family members
- Randomize the *entire family* to
  - Anti-CoV2 antibodies
  - Placebo
- Treatment trial for index case
- Prevention trial for family members
  - Abs reduce individual risk AND cuts down on the within family attack rate



Placebo Family



Counterfactual Antibody Family



Randomize families: 75% of the family members are protected
Randomize individuals: 50% of the family members are protected.

Effective family-ring interventions will help to control the virus

NIAID Biostatistics Research Branch


Dean Follmann


Erica Brittain


Sally Hunsberger


Michael Fay


Lori Dodd


Micheal Proschan


Jing Qin


Bruce Swihart


Misrak Gezmu


Zonghui Hu

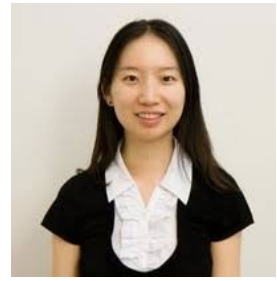

Jonathan Fintzi

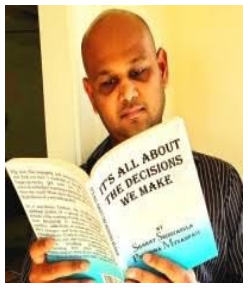

Ana Villa -Ortega


Michele Di Mascio


Martha Nason


C. Jason Liang


Jing Wang



Sharat Srinivasula

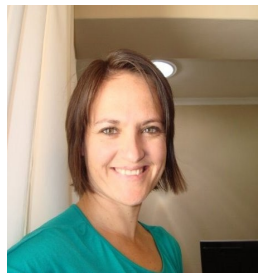

Keith Lumbard


Allyson Mateja


Shannon Gallagher


Alyhana Childs


Michael Duvenhage


Shera Weyers


Iman Gulati


Victoria Bera


Tyler Bonnett


Alyson Francis


Yi Chi


Claudine Casimir


Aurélie Gouel


Erin Gabriel

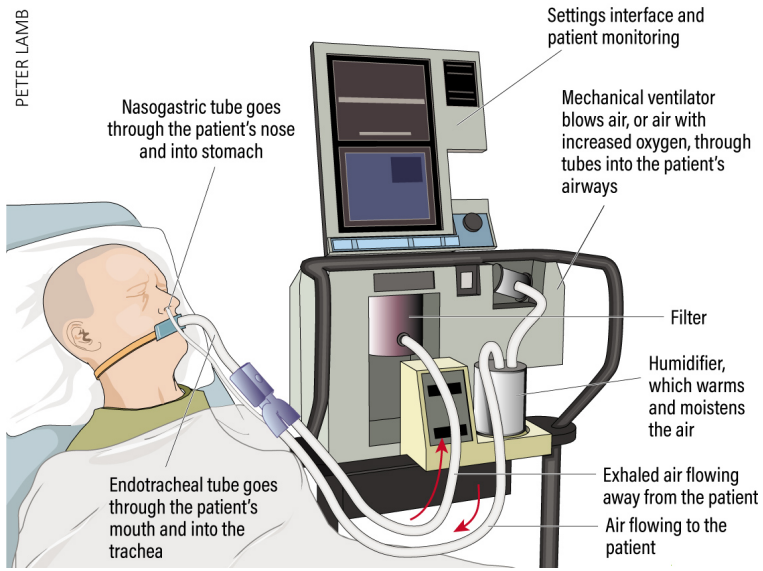
Thanks

- NIAID BRB
- NIH Statisticians
- Emmes Corporation
- Fred Hutch Statisticians

Treatment & Prevention

- Plasma from COVID-19 survivors is rich in SARS-CoV2 antibodies
- Extract it, check it, pool it, test it in clinical trials
- Huge logistical issues with tracking, cataloging, verifying etc.
- BRB-CTRS is essential in ensuring that survivor's donated antibodies can be rigorously evaluation for treatment and prevention

Figure I. Mechanical ventilator for positive pressure ventilation



Invasive Mechanical Ventilation
Intubated & Sedated

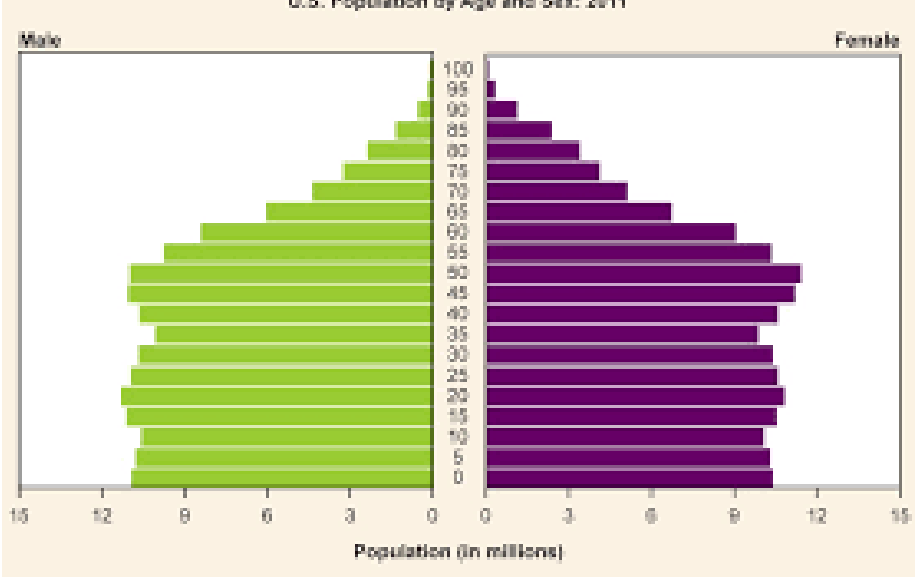
High Flow Oxygen

Low flow oxygen

Describe: Sero-prevalence

- Ideally, do a random sample of the US population
  - That would take a while, especially for us
- Encourage people to volunteer throughout the country
  - d
- Fix up this convenience sample so it represents the US population of sero-prevalence volunteers
  - Can’t really make it random

Sero-prevalence fixup



10 Male 30-40 y.o. donors in Elgin
3 seropositive

$$\text{Estimate} = \frac{3/10 \times w_{\_Elgin} + \ldots}{w_{\_Elgin} + \ldots}$$

$w_{\_Elgin} = 1/28$

Generalization

- Can correct for geographical location, age, gender, etc. Estimate of seroprevalence in Illinois

$$\frac{\sum_{i\ in\ Illinois}\ w_i\ Y_i}{\sum_{i\ in\ Illinois}\ w_i}$$

$w_i = \Pr(\text{ person } i\ \ would\ be\ \ selected\ in\ a\ random\ sample)$

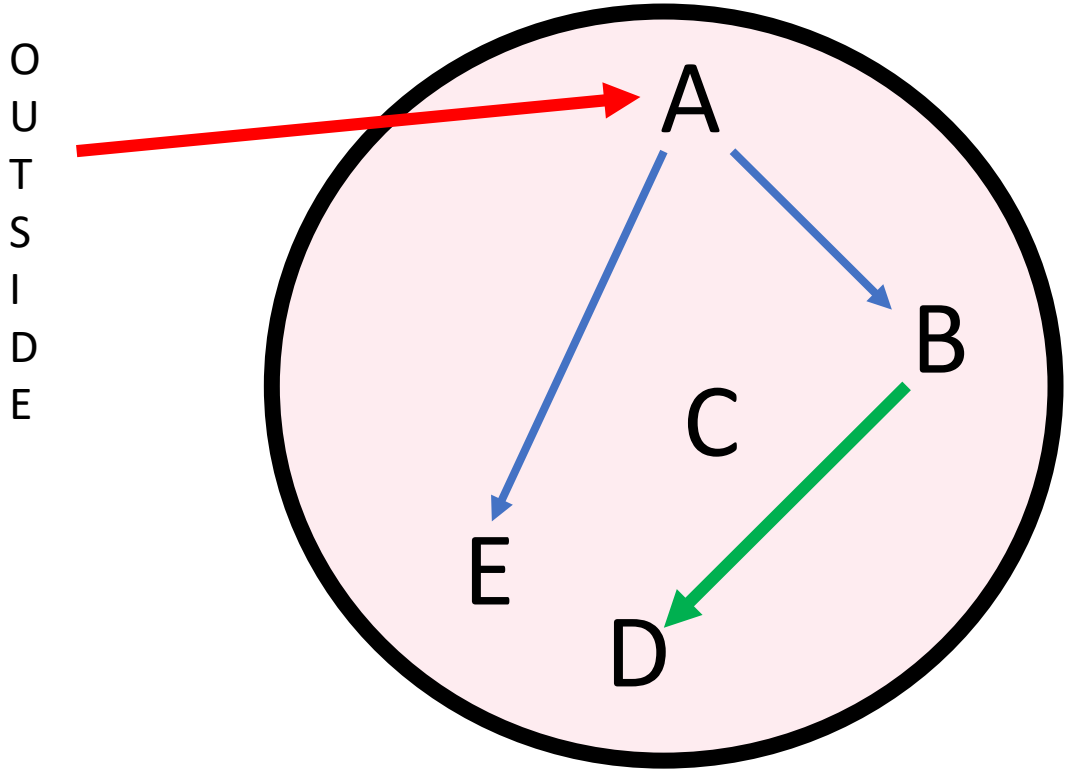
Y_i = 1 if person is is seropositive

Transmission

- NIH employees are getting COVID-19. Designed a protocol
- Identify contacts and family members for onward transmission.

Cluster	Members	Times of Detection	Covariates
1	A, B, C, D, E	0, 3, 7, --, 2, 4	X_A X_B X_C X_D X_E
2	A, B	0, ---	X_A X_B
3	A, B, C	0, 4, ---	X_A X_B X_C

Transmission Sequence Known

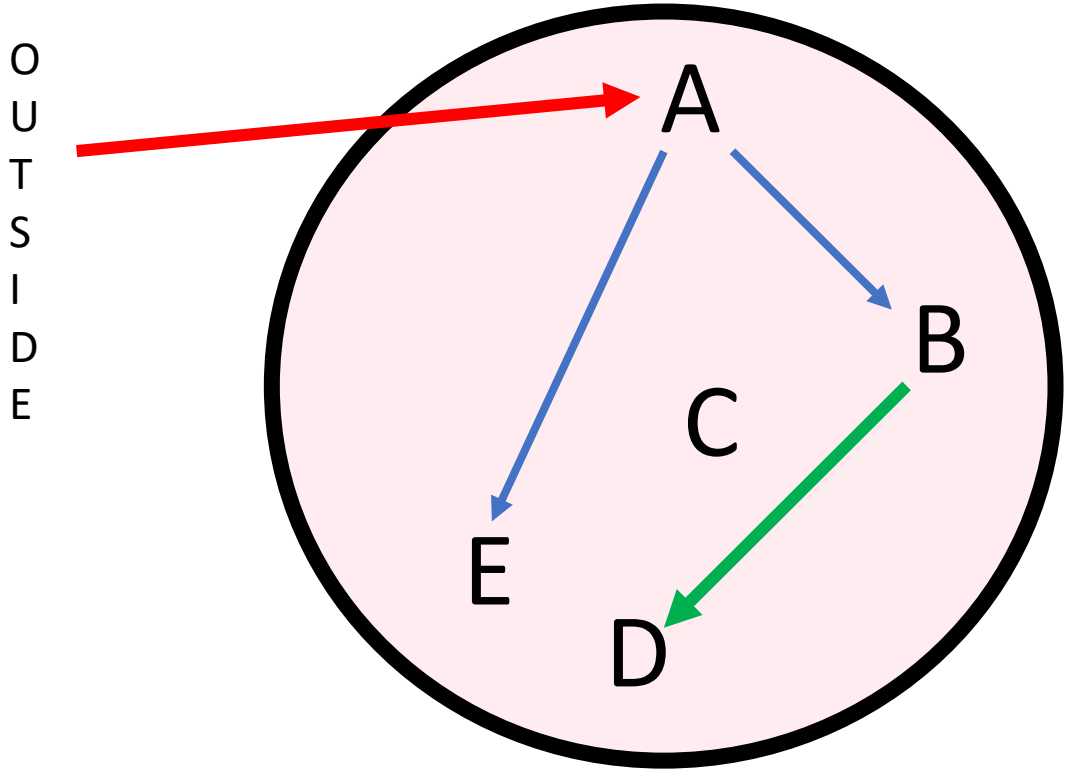


Logit

- Each person A, . . . , E flips a coin to see if they're infected
  - logit{P(Out->A)} = $\boldsymbol{\alpha}_0$+ $\boldsymbol{\alpha}_1$ $\mathbb{I}$(A works outside)

- Say A and B are infected from outside. A and B draw *avoidance scores* for everyone else.
  - If *S*(AC) < 1 then A infects C
  - *S*(AC) ~ Exponential{ exp($\boldsymbol{\beta}_0$ + $\boldsymbol{\beta}_1\mathbb{I}$(A,C share room) }
- Repeat with the newly infected.

Transmission Sequence Known



Outside: P(Out->A) (1-P(Out->B)) (1-P(Out->C)) (1-P(Out->D)) (1-P(Out->E))

1 – P(no one is infected)

Likelihood equals Product of :

Inside Gen 1: P(*S*(AB)<1) P(*S*(AE)<1) P(*S*(AC)>1) P(*S*(AD)>1)

Inside Gen 2: P(*S*(BD)<1 & *S*(BD) < *S*(ED)) P(*S*(BC)>1) P(*S*(EC)>1)

But we don’t know the sequence

- Missing data Likelihood contribution sums over possibilities
  - Suppose A,B infected, C not. Three possibilities
    - Out-> A,B {2}
    - Out-> A, A->B {1,1}
    - Out-> B, B->A {1,1}
- With bigger clusters # of possibilities explodes. Cluster of size 9 has many *partitions*
  - {9}
  - {8,1}, {7,2}, …. {1,8}
  - {1,1,7}, {1,2,6}, …. {7,1,1}
  - {1,1,1,6}, . . .
  - . . .

Evaluate

Agree Run 1 vs Run 2	Run 1 Individual	Run 2 Individual	Pool	Depooled	Agree Run 2 vs Pool
Yes	0 0 0 0 0	0 0 0 0 0	0		Yes
Yes	0 0 0 0 0	0 0 0 0 0	0		Yes
No	1 0 0 0 0	0 0 0 0 0	0		Yes
Yes	1 0 0 0 0	1 0 0 0 0	0		No
Yes	1 0 0 0 0	1 0 0 0 0	1	1 0 0 0 0	Yes
Yes	0 0 0 0 0	0 0 0 0 0	0		Yes

5/6 Pooling and re-running have similar reproducibility