Cross-Sectional Studies

Madhukar Pai, McGill University, Montreal

COVID-19 Antibody Seroprevalence in Santa Clara County, California

Eran Bendavid[1], Bianca Mulaney[2], Neeraj Sood[3], Soleil Shah[2], Emilia Ling[2], Rebecca Bromley-Dulfano[2],
Cara Lai[2], Zoe Weissberg[2], Rodrigo Saavedra-Walker[4], Jim Tedrow[5], Dona Tversky[6], Andrew Bogan[7],
Thomas Kupiec[8], Daniel Eichner[9], Ribhav Gupta[10], John P.A. Ioannidis[1,10], Jay Bhattacharya[1]

Methods

On April 3-4, 2020, we tested county residents for antibodies to SARS-CoV-2 using a lateral flow immunoassay. Participants were recruited using Facebook ads targeting a sample of individuals living within the county by demographic and geographic characteristics. We estimate weights to adjust our sample to match the zip code, sex, and race/ethnicity distribution within the county. We report both the weighted and unweighted prevalence of antibodies to SARS-CoV-2. We also adjust for test performance characteristics by combining data from 16 independent samples obtained from manufacturer's data, regulatory submissions, and independent evaluations: 13 samples for specificity (3,324 specimens) and 3 samples for sensitivity (157 specimens).

Results

The raw prevalence of antibodies to SARS-CoV-2 in our sample was 1.5% (exact binomial 95CI 1.1-2.0%). Test performance specificity in our data was 99.5% (95CI 99.2-99.7%) and sensitivity was 82.8% (95CI 76.0-88.4%). The unweighted prevalence adjusted for test performance characteristics was 1.2% (95CI 0.7-1.8%). After weighting for population demographics of Santa Clara County, the prevalence was 2.8% (95CI 1.3-4.7%), using bootstrap to estimate confidence bounds. These prevalence point estimates imply that 54,000 (95CI 25,000 to 91,000 using weighted prevalence; 23,000 with 95CI 14,000-35,000 using unweighted prevalence) people were infected in Santa Clara County by early April, many more than the approximately 1,000 confirmed cases at the time of the survey.
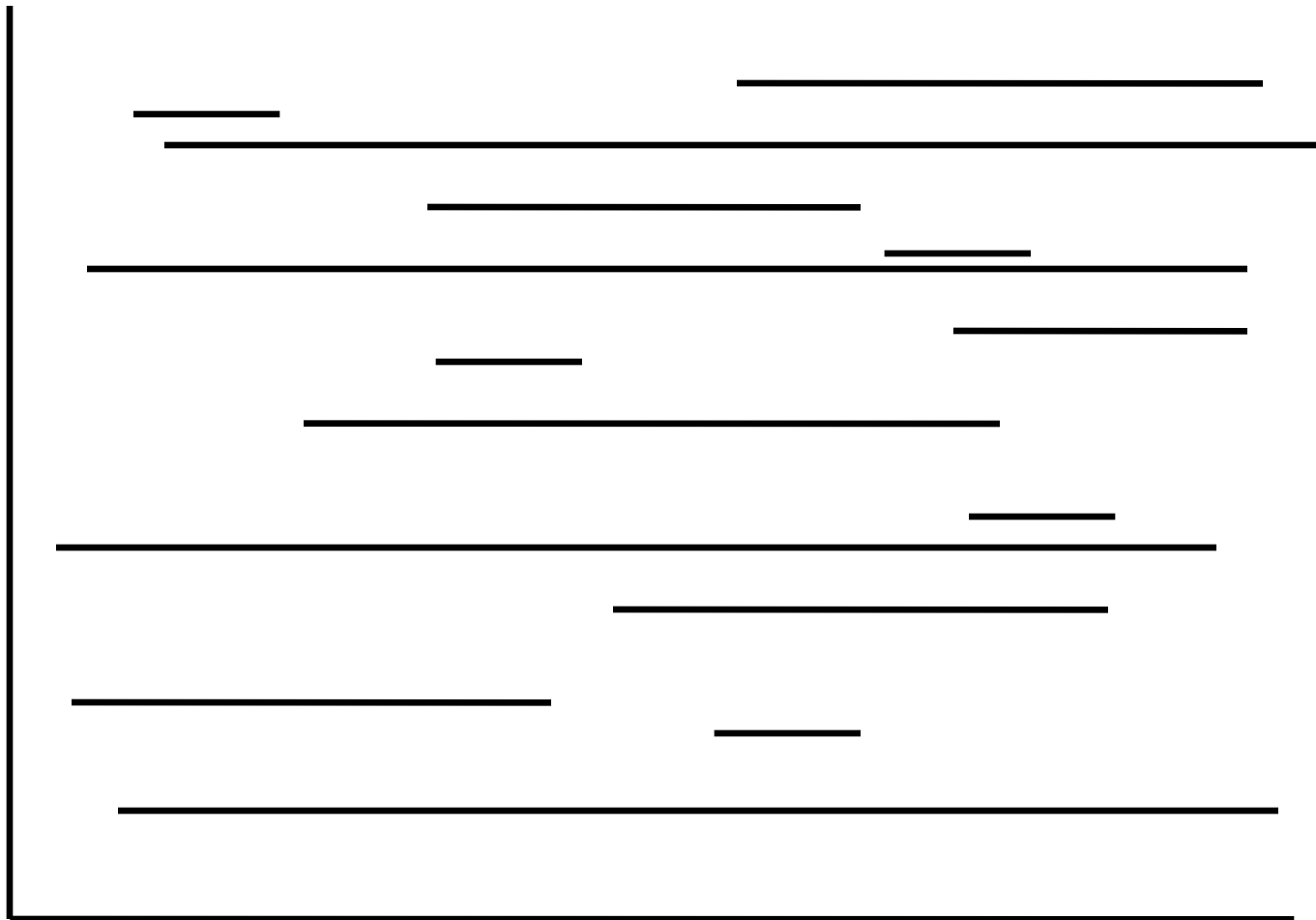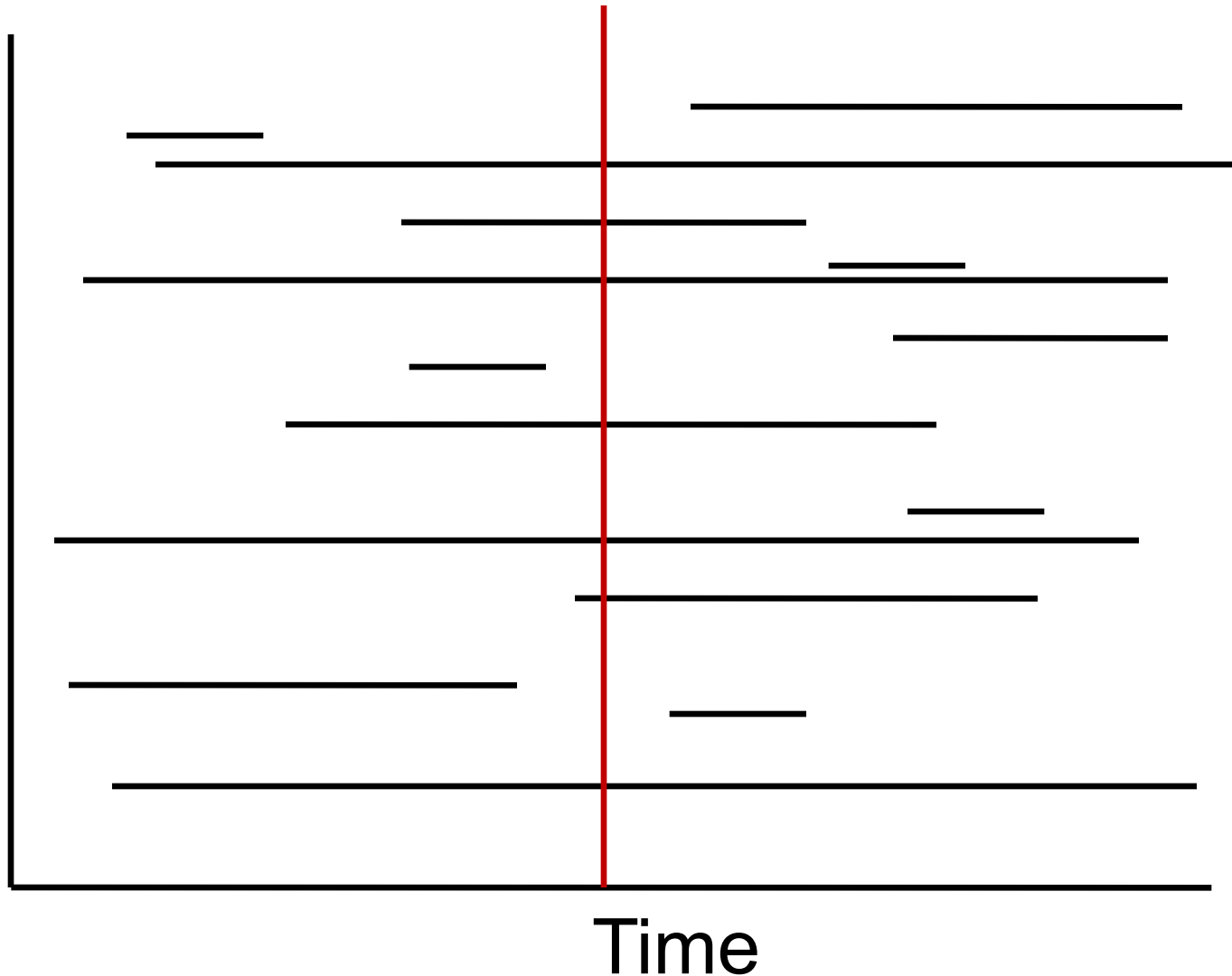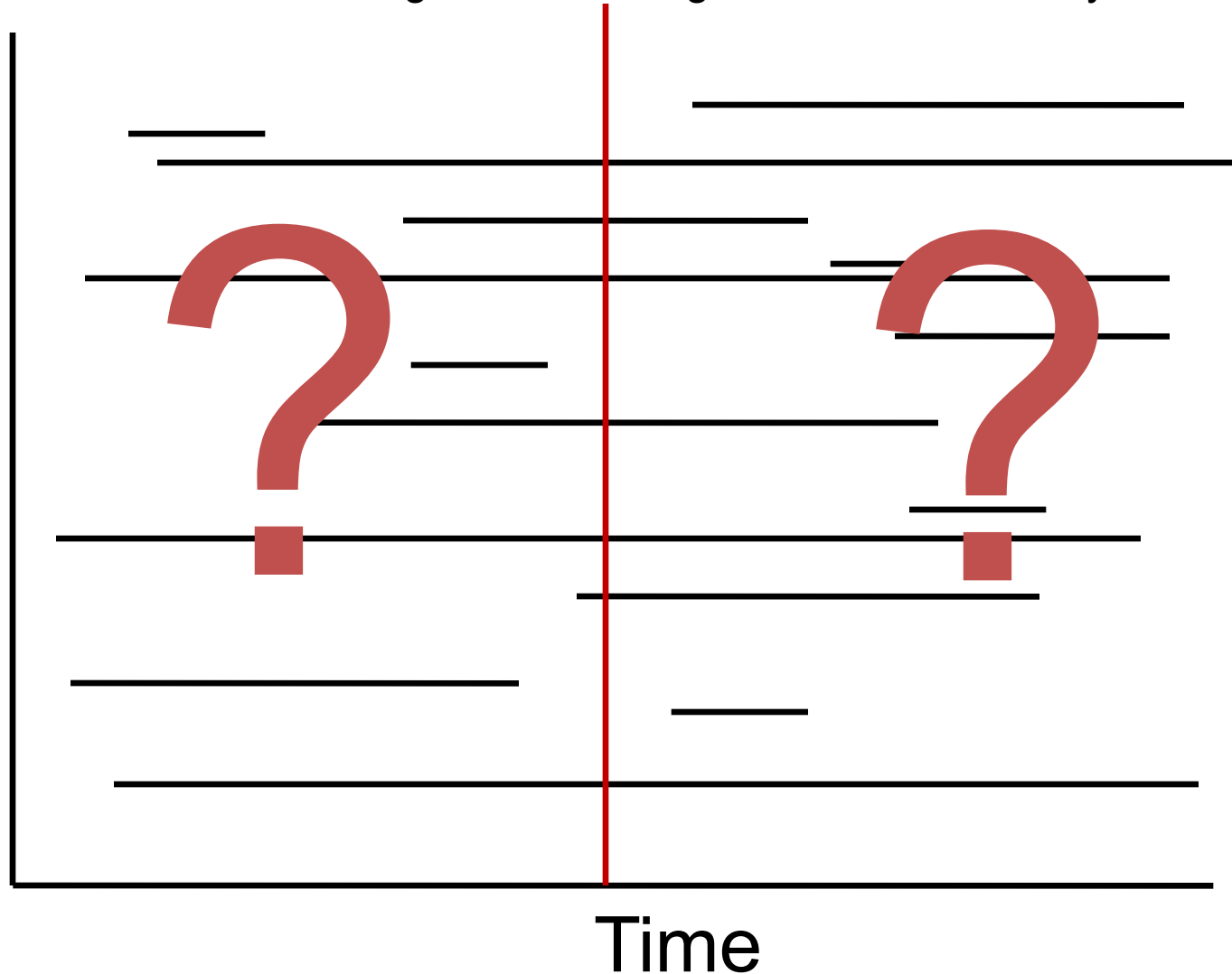
Time

Sampling at one point in time
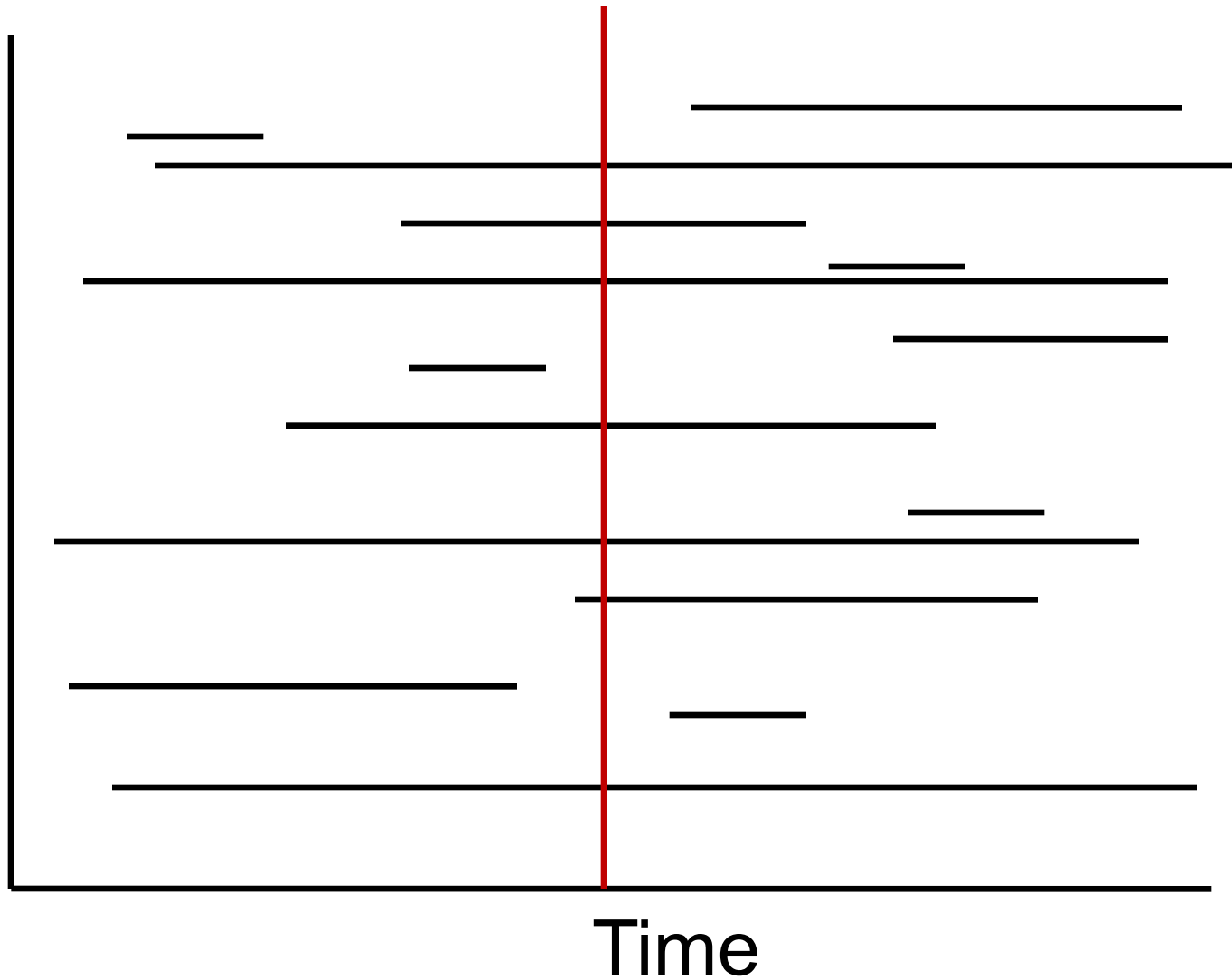
Time

Cross-sectional sampling

At one point in time info on both exposure and outcomes are collected
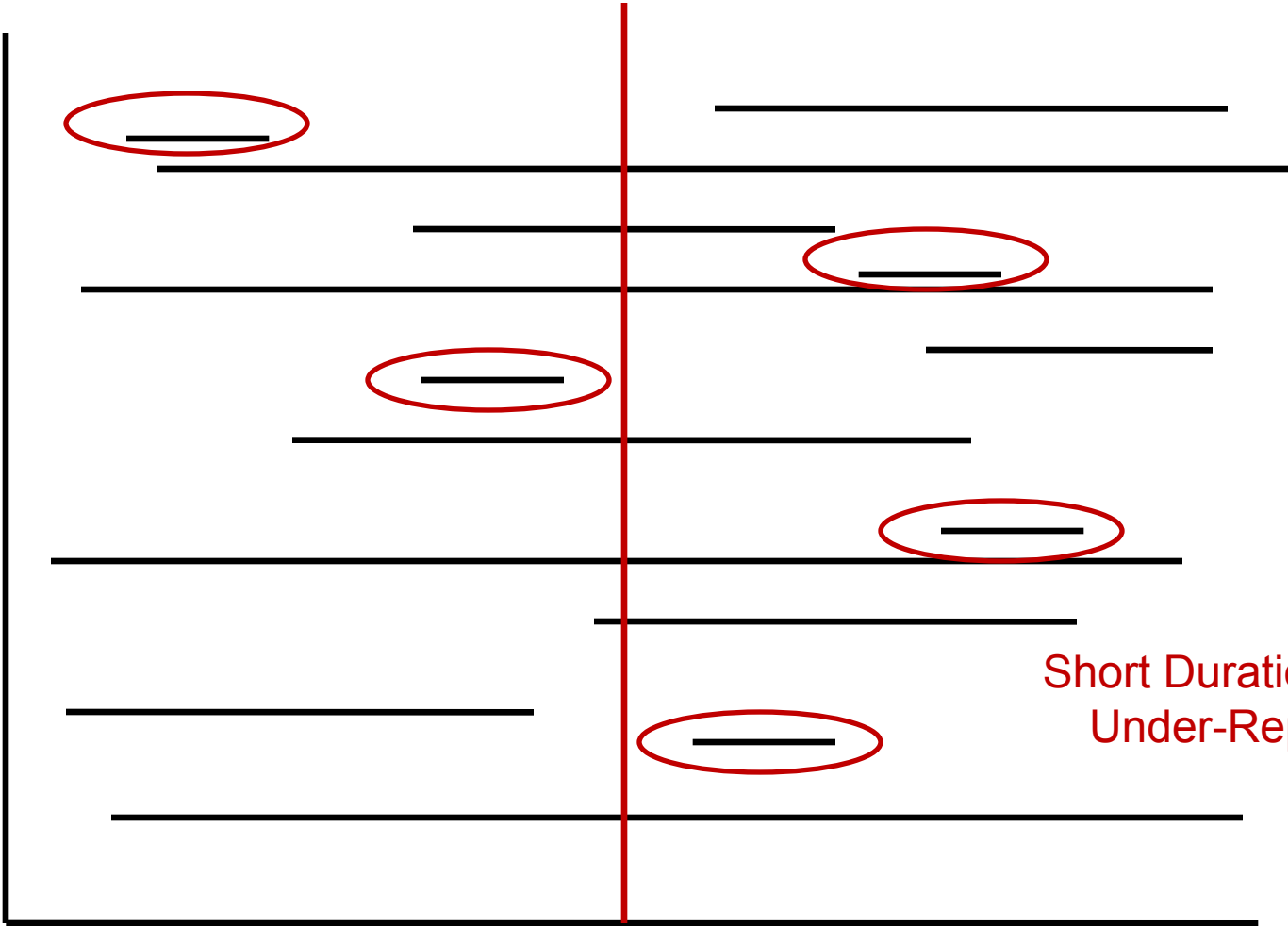

Time

Does not provide info on when cases occurred or what the true incidence is

Existing Cases at Single Point in Time Only
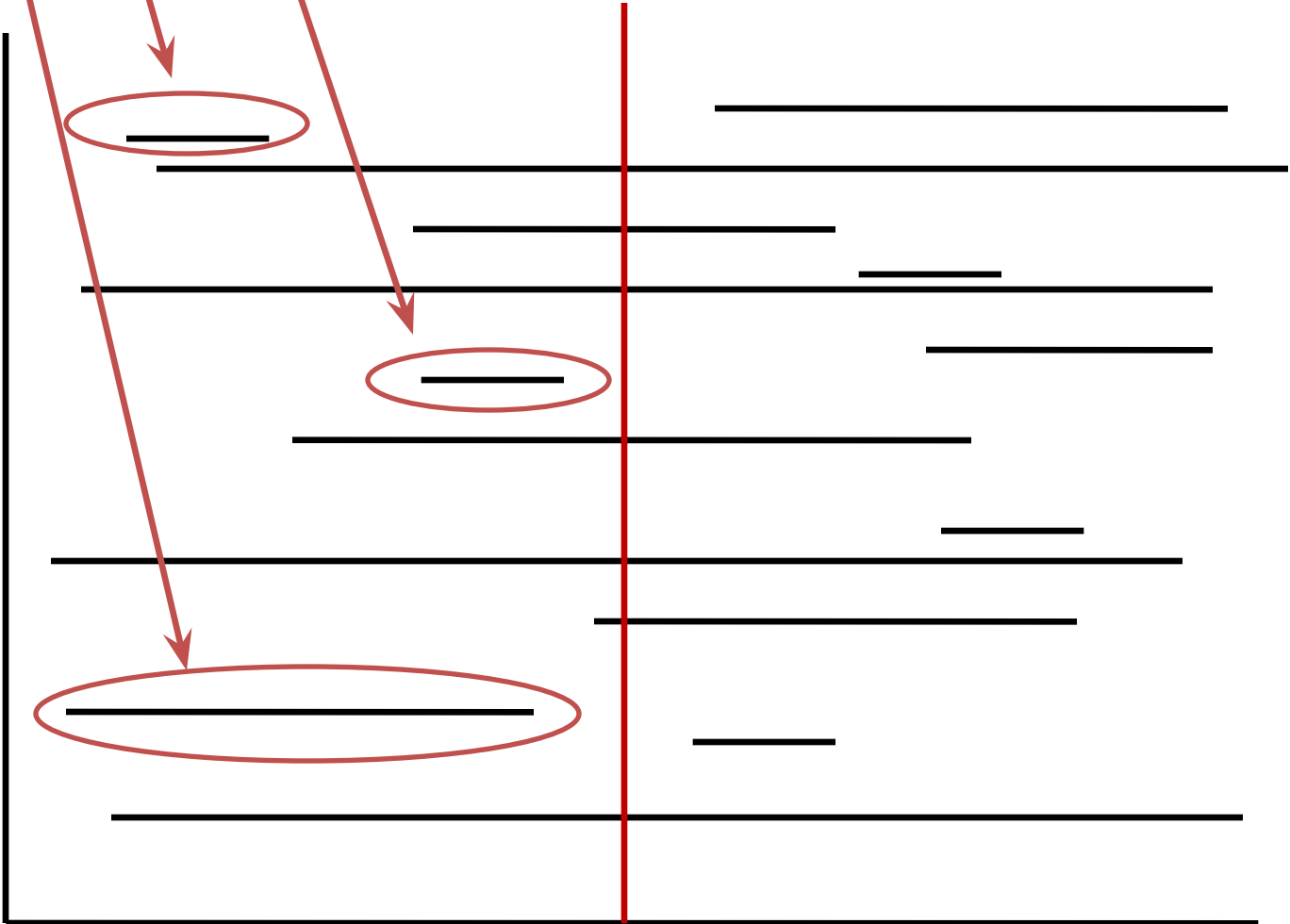

Time

Prevalent Cases

Sampling

Short Duration Diseases
Under-Represented

Time

Prevalent Cases

Non-Survivors Excluded

Sampling

Time

8

Cross-Sectional Study

- Also known as prevalence study or survey
- A snapshot or picture of health experience at one time
- Sampling independent of both exposure and outcome
- Commonly used in public policy and public health
- Used to provide measure of disease burden

Point Prevalence

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

C = # of observed cases at time t

N = Population size at time t

Measures the frequency of disease at a given point in time

Point Prevalence Example

COVID-19 Antibody Seroprevalence in Santa Clara County, California

Eran Bendavid[1], Bianca Mulaney[2], Neeraj Sood[3], Soleil Shah[2], Emilia Ling[2], Rebecca Bromley-Dulfano[2], Cara Lai[2], Zoe Weissberg[2], Rodrigo Saavedra-Walker[4], Jim Tedrow[5], Dona Tversky[6], Andrew Bogan[7], Thomas Kupiec[8], Daniel Eichner[9], Ribhav Gupta[10], John P.A. Ioannidis[1,10], Jay Bhattacharya[1]

Version 2, April 27, 2020
(revised in response to comments received. This remains a preliminary report of the work.)

- April 3rd and 4th, 2020, researchers did serologic testing for SARS-CoV-2 antibodies in 3,330 adults and children in Santa Clara County

- Total number of positive cases by either IgG or IgM = 50

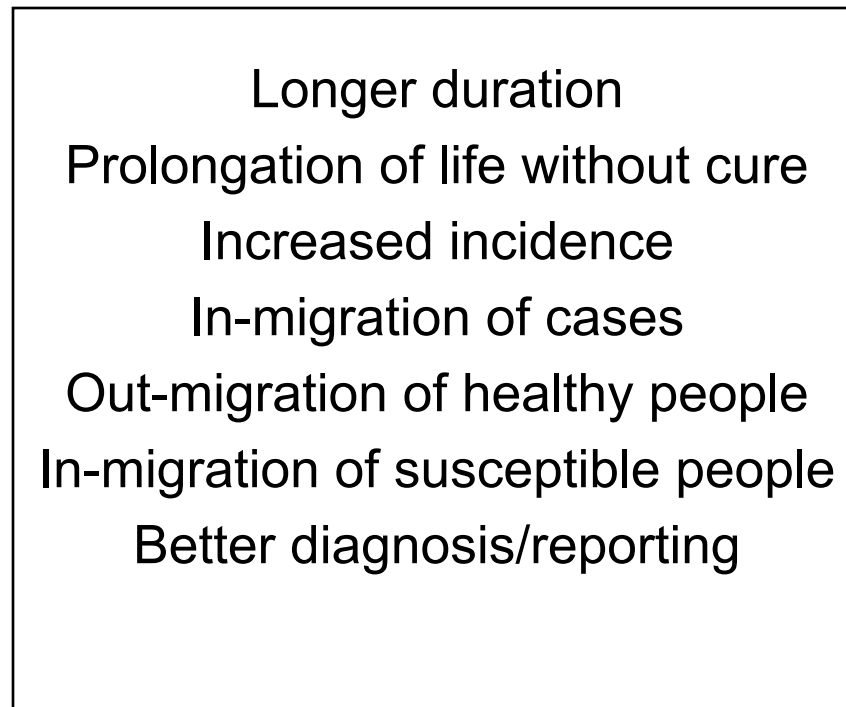- Crude point prevalence = 50/3330 = 1.5% (95 CI 1.1-2.0%)

Recall: Prevalence vs. Incidence
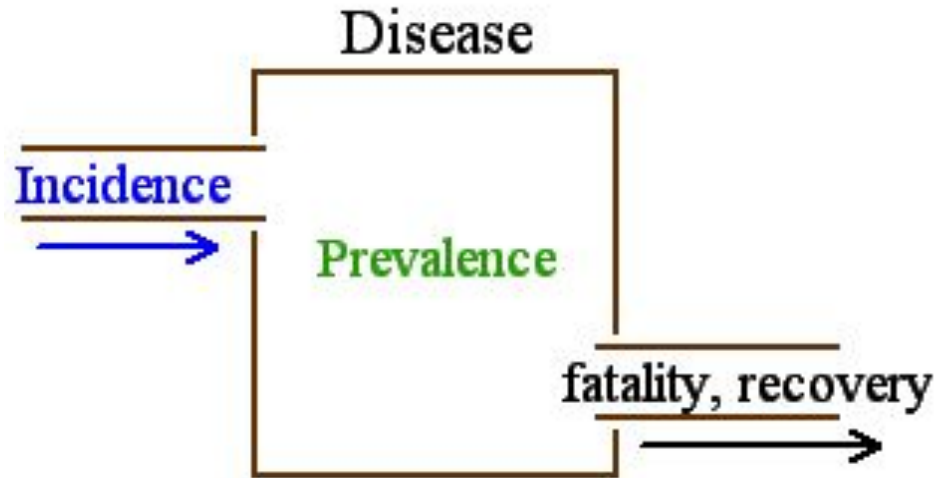
- Prevalence can be viewed as describing a pool of disease in a population.

- Incidence describes the input flow of new cases into the pool.

- Fatality and recovery reflects the output flow from the pool.

What factors can increase prevalence?
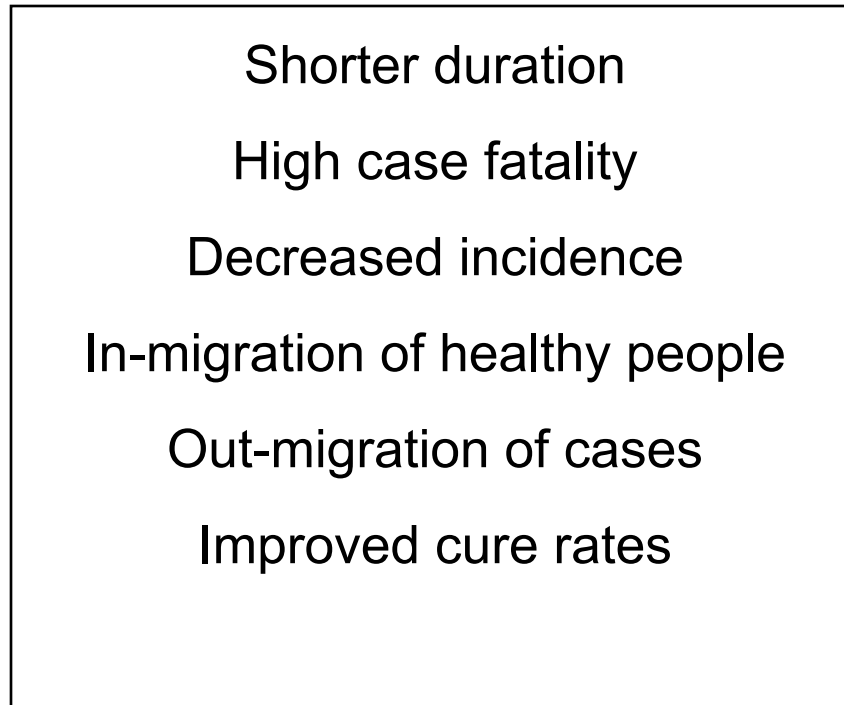
Prevalence

Longer duration

Prolongation of life without cure

Increased incidence

In-migration of cases

Out-migration of healthy people

In-migration of susceptible people

Better diagnosis/reporting

Source: Beaglehole, 1993

What factors can decrease prevalence?

Prevalence

Shorter duration

High case fatality

Decreased incidence

In-migration of healthy people

Out-migration of cases

Improved cure rates

Source: Beaglehole, 1993

Advantages

- Representative of general population
- Convenient, inexpensive, and fast
- Can assess several exposures and several outcomes
- Common diseases with long duration: generate hypotheses for future studies

Disadvantages

- Examines existing cases at a single point in time (not new cases over follow-up period)
- Includes only cases that survive to be available for study (i.e., prevalent cases)
- Short duration diseases may be under-represented
- Cannot establish temporality and directionality
  - Does E $\rightarrow$ D? Does D $\rightarrow$ E?

Selection bias in cross-sectional studies

- Sources:
  - Bias due to sampling
    - Selection of "survivors" or "prevalent" cases
    - Non-random sampling schemes
    - Volunteer bias
    - Membership bias
  - Bias due to non-participation
    - Non-response bias

Selection bias in sample surveys
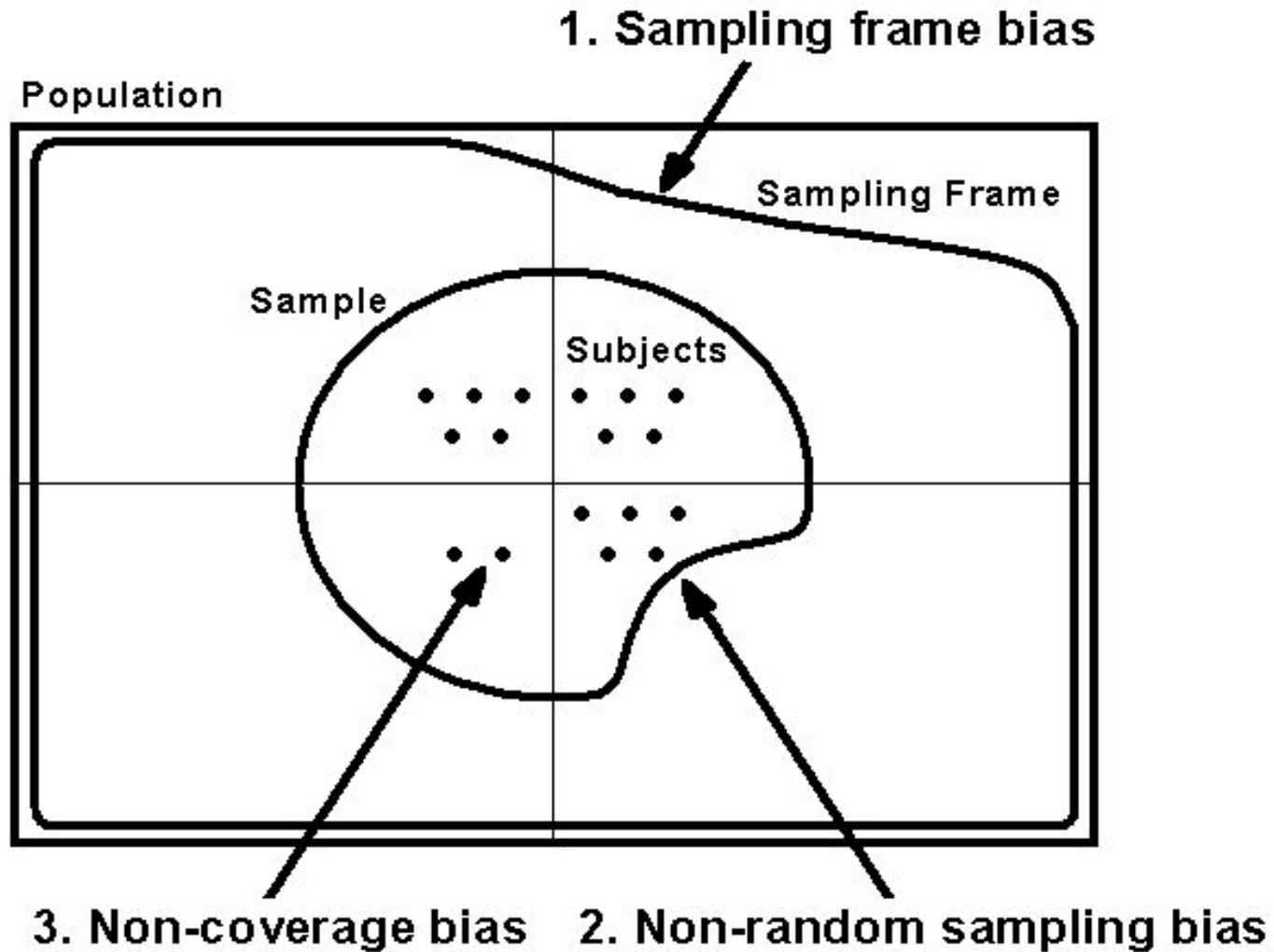
Information bias in cross-sectional studies

- Sources:
  - Quality of the tool, instrument or test used to collect data
    - Questionnaires
    - Tests, etc.
- Example: If you did a sero-prevalence survey on Covid-19, and used a test with low specificity, most positive people will be false-positives.
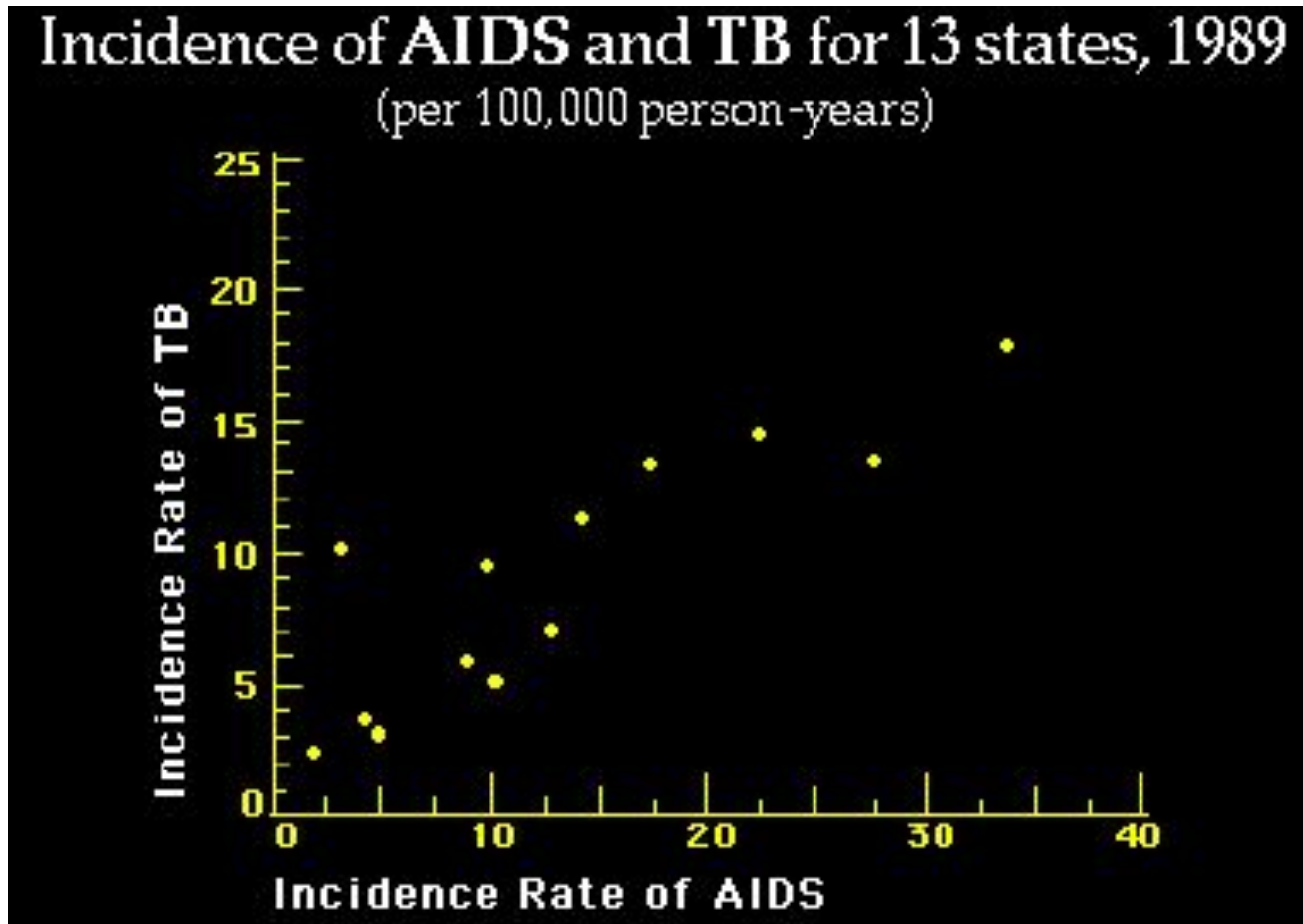
Ecologic Studies

Madhukar Pai, MD, PhD

McGill University

madhukar.pai@mcgill.ca

Ecologic Studies

- "A study in which units of analysis are populations or groups of people than individuals." – Dictionary [Porta 2008]

# Example: Correlation between TB and AIDS

Incidence of AIDS and TB for 13 states, 1989
(per 100,000 person-years)

Source: ActivEpi

Research

Air pollution and case fatality of SARS in the People's Republic of China: an ecologic study

Yan Cui[1], Zuo-Feng Zhang*[1], John Froines[2], Jinkou Zhao[3], Hua Wang[3], Shun-Zhang Yu[4] and Roger Detels[1]

SARS Case Fatality (%)

Case Fatality=-0.063+0.001*API
Correlation Coefficient=0.8568

Air Pollution Index
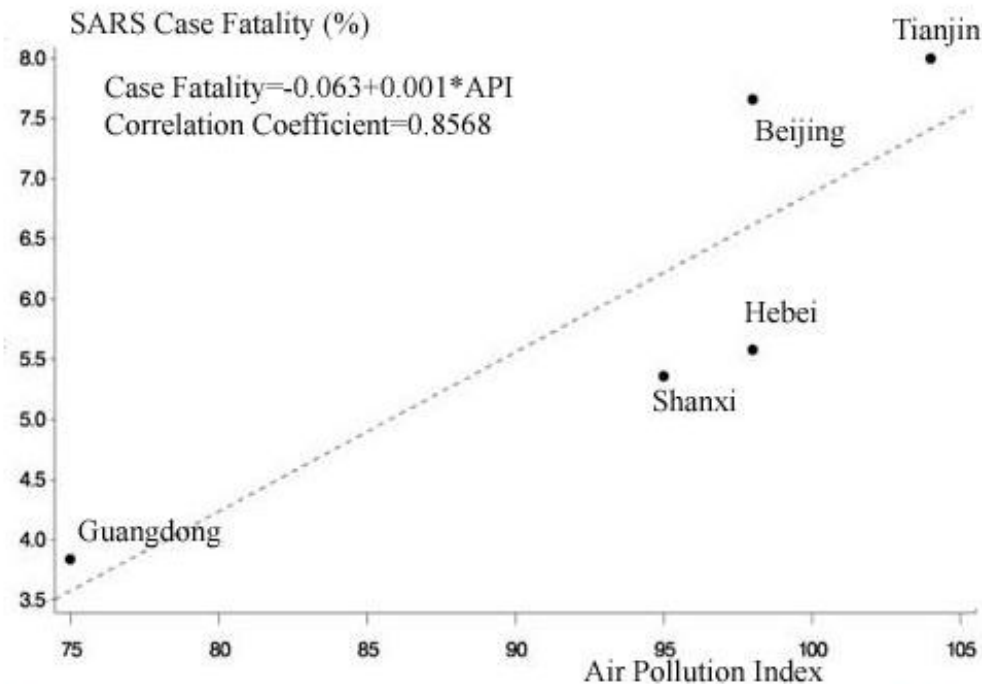
The Correlation and Association between Short-term Exposure to Ambient Air Pollution and Case Fatality of SARS in People's Republic of China.

💬 **Comments (145)**

Correlation between universal BCG vaccination policy and reduced morbidity and mortality for COVID-19: an epidemiological study

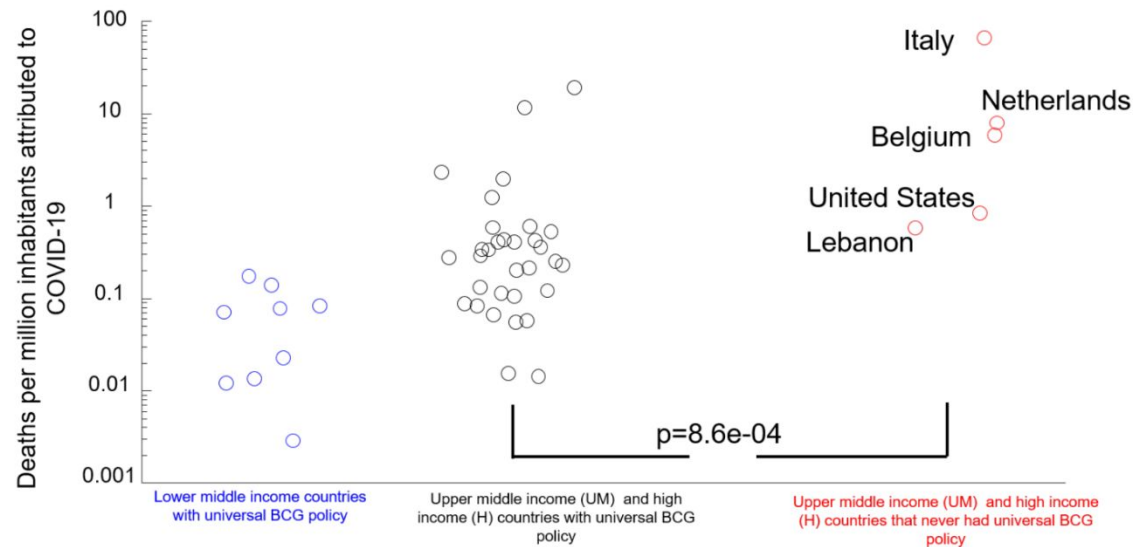Aaron Miller, Mac Josh Reandelar, Kimberly Fasciglione, Violeta Roumenova, Yan Li, Gonzalo H Otazu

doi: https://doi.org/10.1101/2020.03.24.20042937

Figure 1: Higher death rates were presented in countries that never implemented a universal BCG vaccination policy.

25

Dr C Vijayabaskar ✔
@Vijayabaskarofl

BCG vaccine shall be administered on trial basis for adults aged 60-95, ordered by Hon'ble @CMOTamilNadu. This is aimed at reducing the #Covid mortality rate in Senior people. National Institute for Research in Tuberculosis will start the pilot program soon. #ICMR #Vijayabaskar

4:48 AM · Jul 15, 2020 · Twitter for iPhone

MENU **MICROBIOLOGY**   Search Nature Research Microbiology Commun 🔍

CONTRIBUTOR

JOURNAL CLUB CORONAVIRUSES: PAST, PRESENT AND FUTURE

Universal BCG vaccination and protection against COVID-19: critique of an ecological study

This post was written by Lena Faust, Sophie Huddart, Emily MacLean and Anita Svadzian. The authors are PhD students in Epidemiology at the McGill International TB Centre, Montreal, Canada.

Emily MacLean
PhD candidate, McGill University

Published Apr 01, 2020

BCG Against Coronavirus: Less Hype And More Evidence, Please

Madhukar Pai Contributor ⓘ
Healthcare
I write about global health, infectious diseases, and equity

Close up of reaction of Bacillus Calmette Guerin or BCG vaccination infants. GETTY

A Skeptic's Guide To Ecologic Studies During A Pandemic

Madhukar Pai Contributor ⓘ
Healthcare
I write about global health, infectious diseases, and equity

24 January 2020, Bavaria, Munich: A face mask and protective goggles are displayed in front of a map ... [+] DPA/PICTURE ALLIANCE VIA GETTY IMAGES

Critique

- Ecologic correlations: do not have to be true at the individual level
- Low testing rates in many low-income countries
- Epidemic is not on the same time scale across countries
  - Many countries that give BCG now seeing huge increases in Covid19
- Confounding by age structure, other vaccines
- Confounding by indication – why is BCG still being given?
- Countries that give BCG also give many other vaccines or have other issues (e.g. malaria endemic)
- Inconsistencies: India, Brazil, Peru, Russia, Iran, Pakistan, Mexico, Chile – all give BCG & have huge Covid-19 epidemics
- Why should BCG at birth protect adults? Biology?

The role of vitamin D in the prevention of coronavirus disease 2019 infection and mortality

Petre Cristian Ilie[1] · Simina Stefanescu[2] · Lee Smith[3]
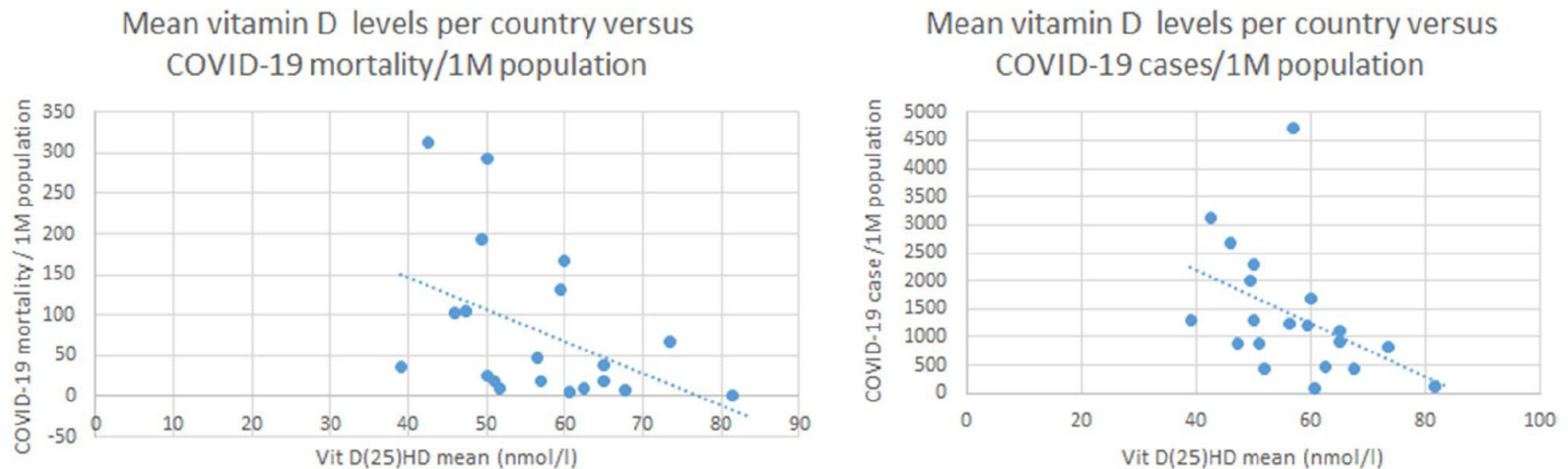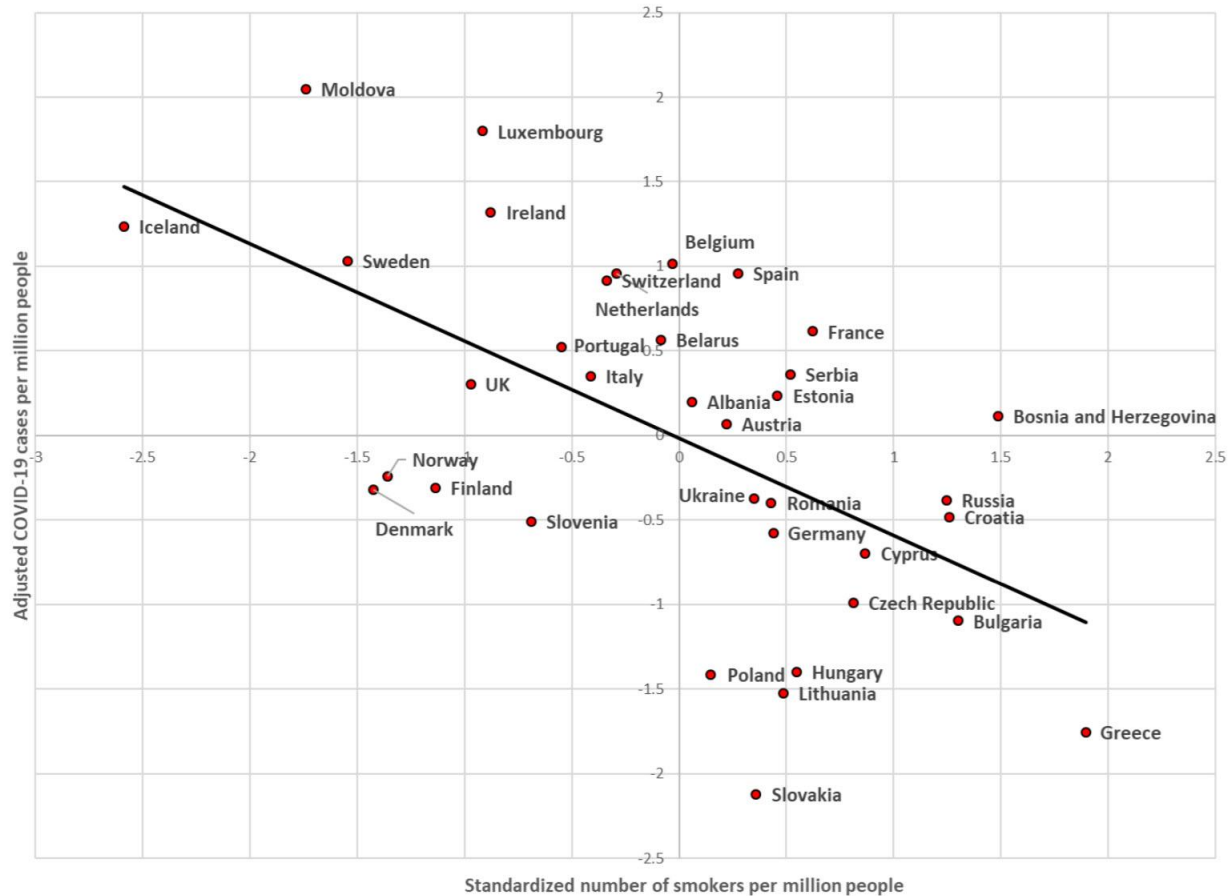
Fig. 1 Mean vitamin D levels per courtry versus COVID-19 cases and mortality/1M population
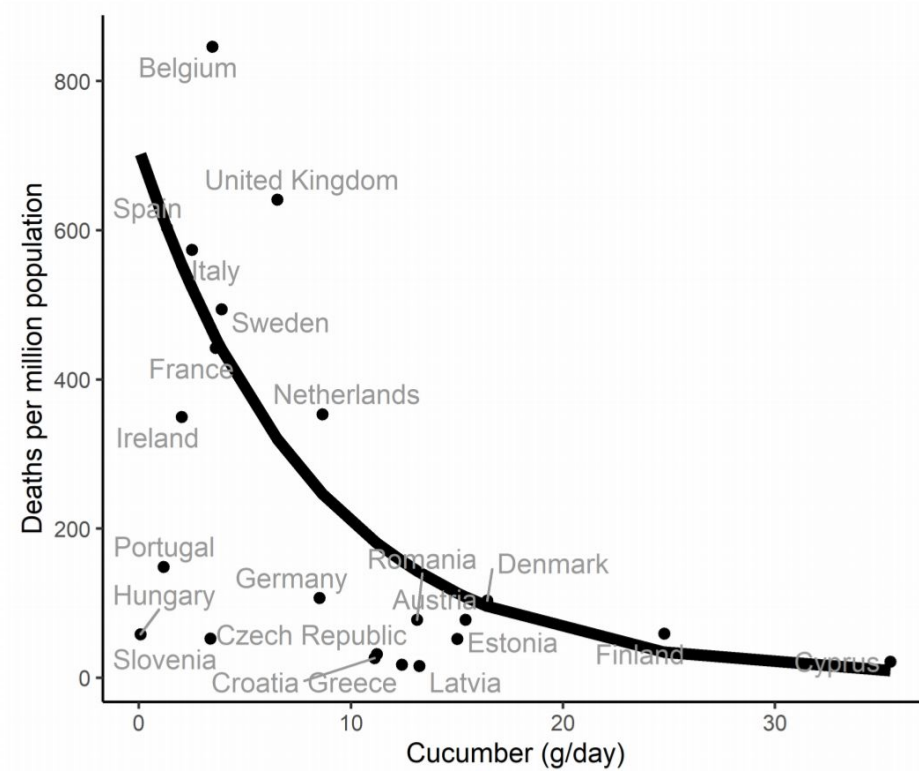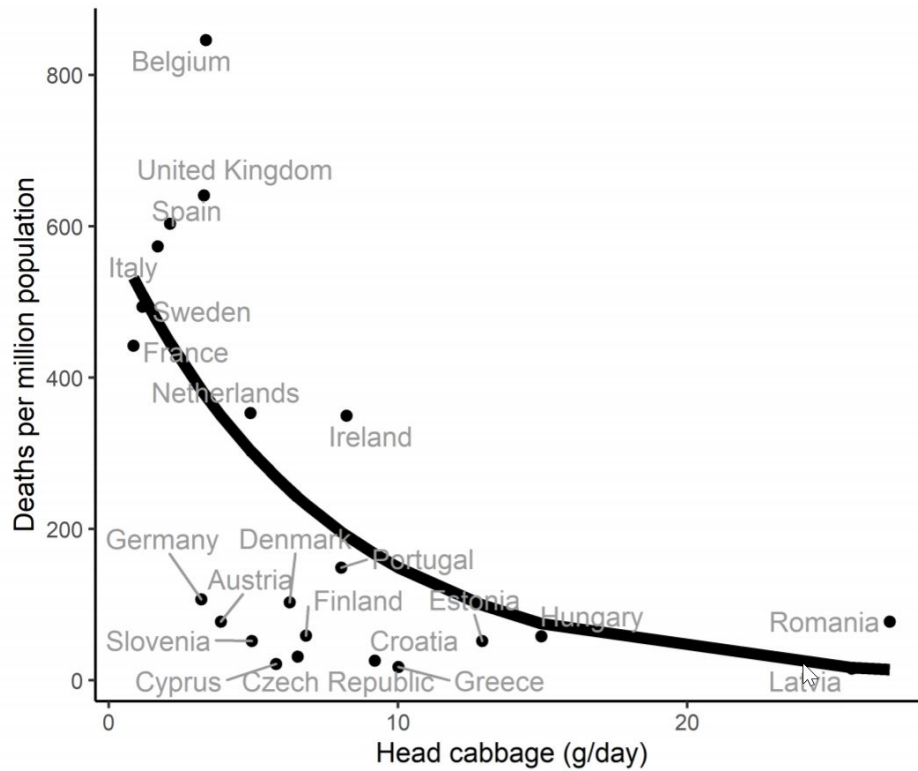
Smoking prevalence and COVID-19 in Europe FREE

Panagiotis Tsigaris, Jaime A Teixeira da Silva ✉

Figure: Scatter plot of Standardized number of smokers per million people (x-axis) vs Adjusted COVID-19 cases per million people (y-axis), with a downward trend line. Countries plotted include Moldova, Luxembourg, Ireland, Iceland, Belgium, Sweden, Switzerland, Spain, Netherlands, France, Portugal, Belarus, Serbia, Italy, UK, Estonia, Albania, Austria, Bosnia and Herzegovina, Norway, Finland, Denmark, Ukraine, Romania, Russia, Croatia, Slovenia, Germany, Cyprus, Czech Republic, Bulgaria, Poland, Hungary, Lithuania, Greece, Slovakia.

https://academic.oup.com/ntr/article/doi/10.1093/ntr/ntaa121/5866017

Association between consumption of vegetables and COVID-19 mortality at a country level in Europe

Why do ecologic studies?

- Low cost and convenience
- Some measurements cannot be made on individuals
- Ecologic effects are the main interest (at the population level)
- Simplicity of analyses and presentation
- Often helpful for generating new hypotheses for further research

Limitations of Ecological Studies

- Usually Hypothesis Generating
- Lack of adequate data and missing data
    - May not be recorded a group level
- Confounding is hard to adjust for
- Collinearity
- Temporal Ambiguity
- Ecological Fallacy:
    - Ecological fallacy arises from thinking that relationships observed for groups necessarily hold for individuals.

Rothman and Greenland 1998
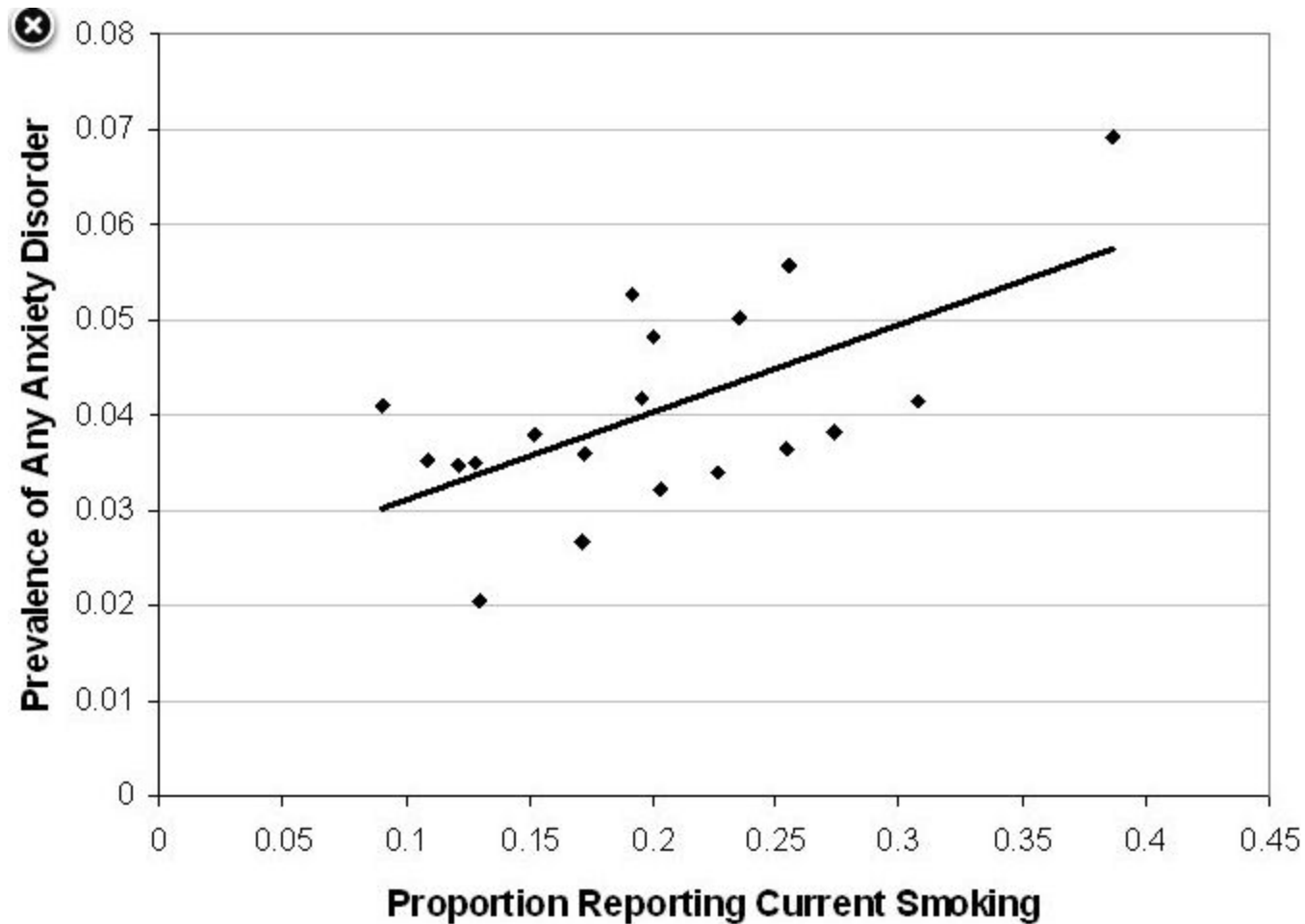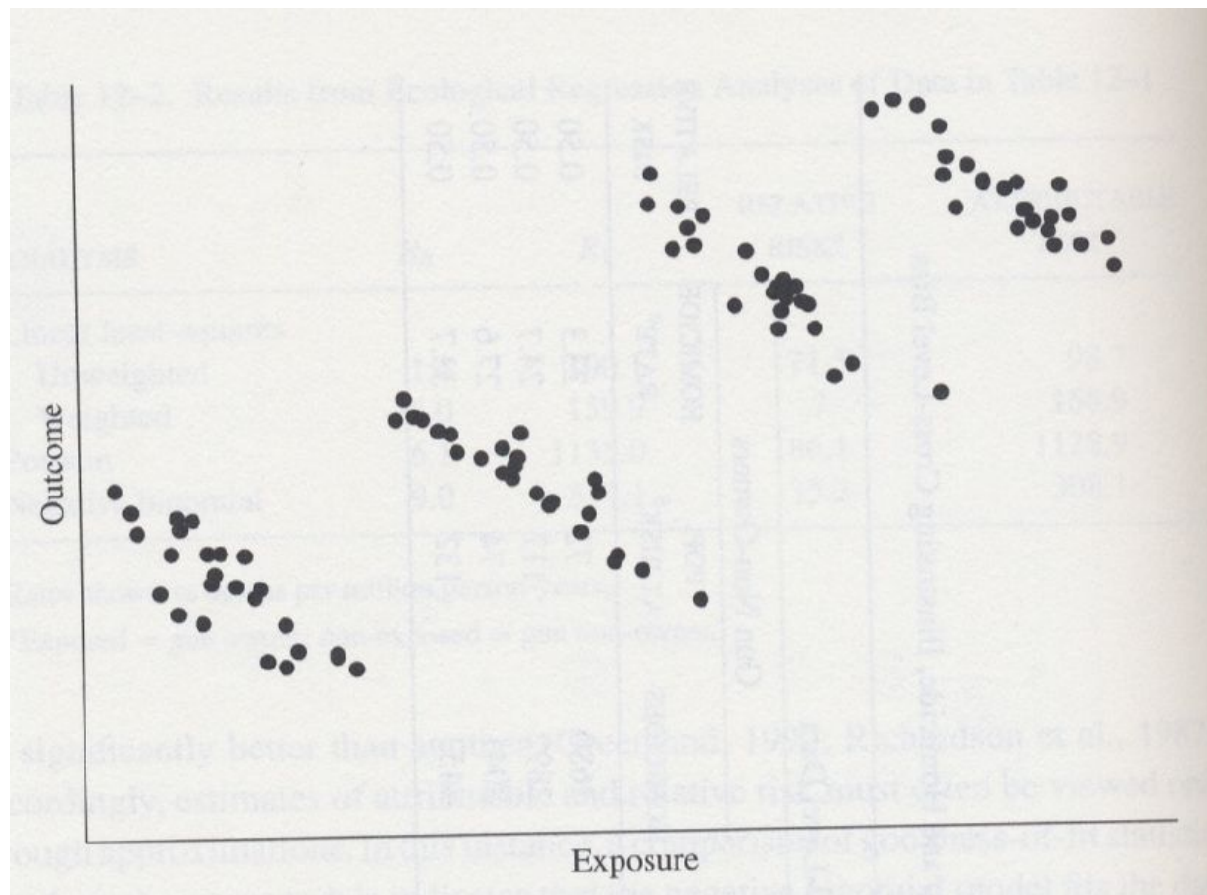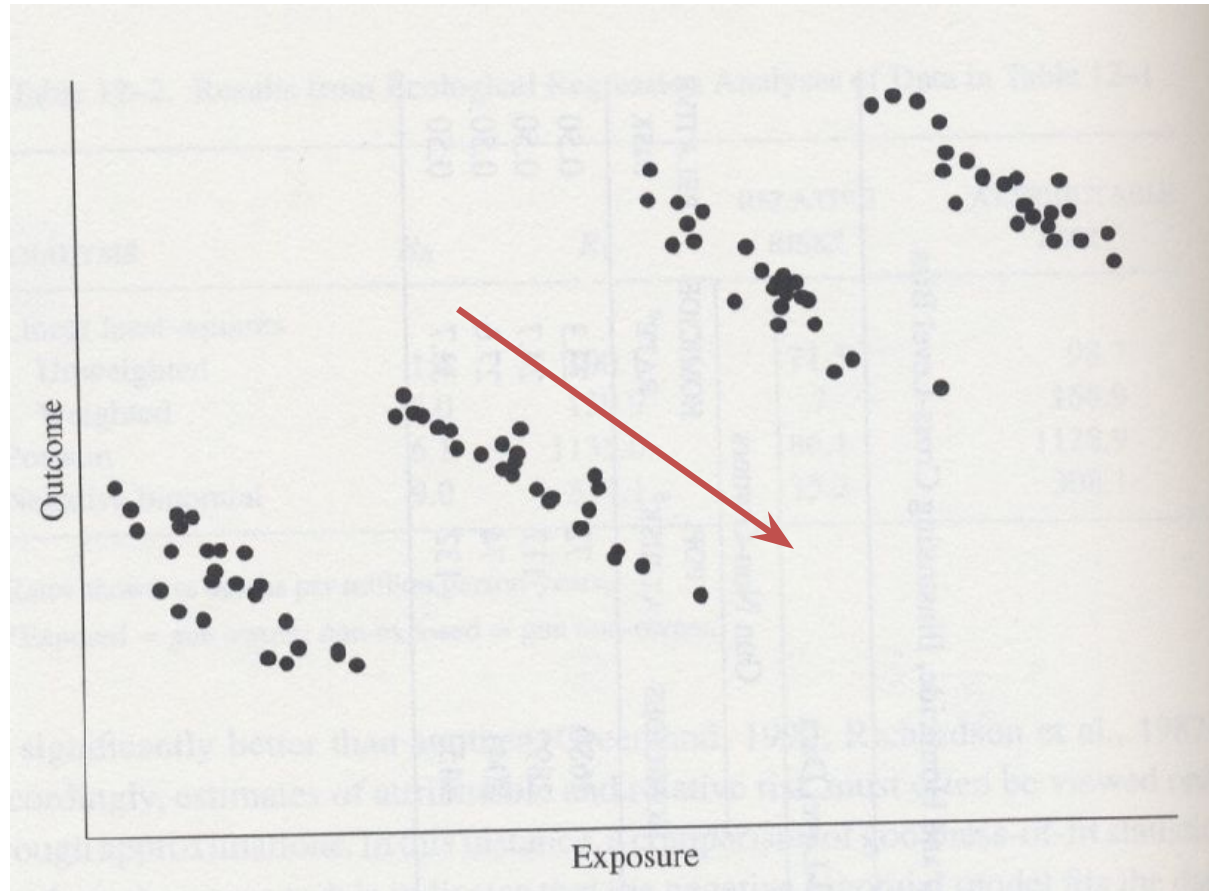
Temporal sequence is not clear

Figure 4: Ecological Analysis: Current Smoking and Anxiety Disorder Prevalence

Aggregation Bias

Koepsell and Weiss Epidemiologic Methods 2003

Aggregation Bias

Correlation negative at individual level

Koepsell and Weiss Epidemiologic Methods 2003

Aggregation Bias
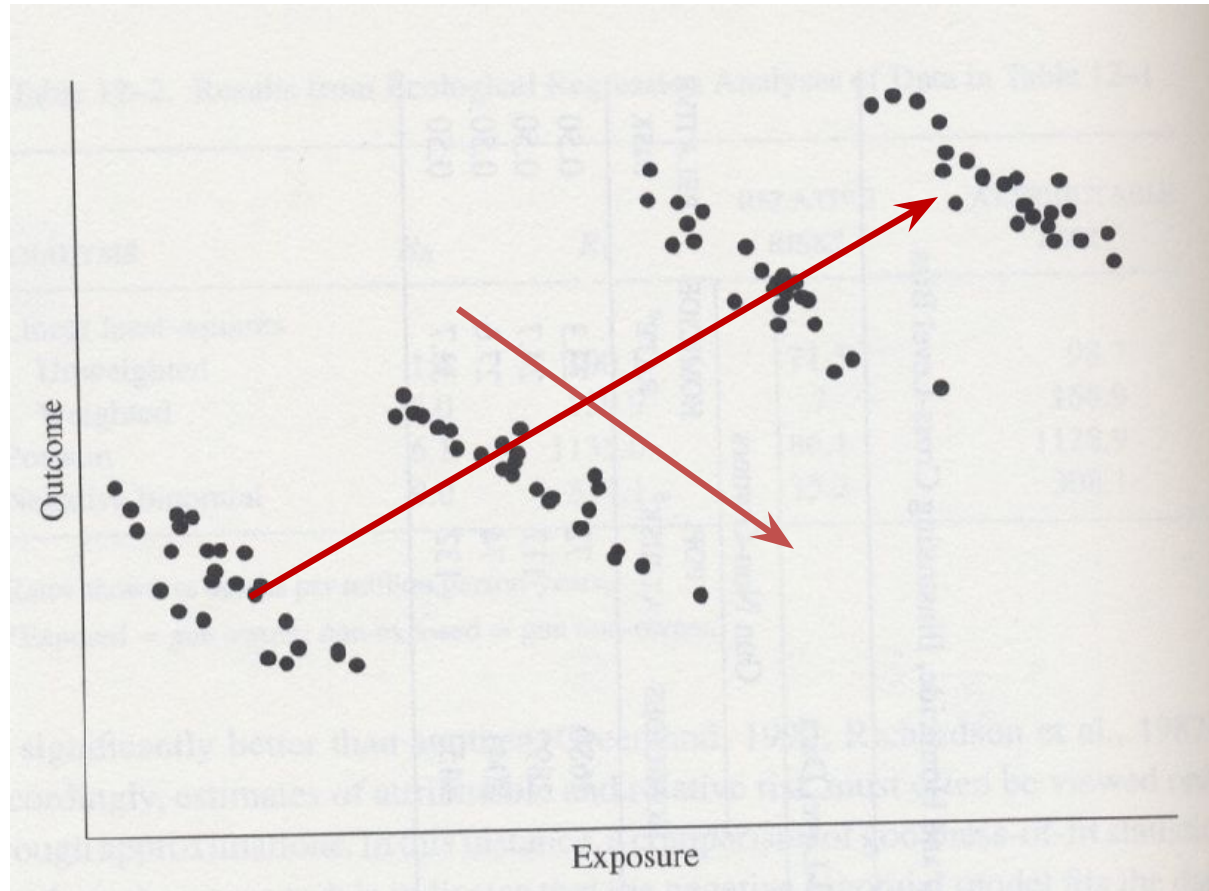
Correlation negative at individual level but positive at group level

Control of confounding

- Always a concern
  - E.g. BCG and Covid-19
    - Age structure across countries
    - Other vaccines given
    - Testing rate variability
- We don't often have data on confounders
- Even if we adjust for confounding at the group level, it can still be an issue at the individual level

Key issues with ecologic Studies

- Explores correlations between aggregate (group level) exposure and outcomes
- Unit of analysis: usually not individual, but clusters (e.g. countries, counties, schools)
- Useful for generating hypothesis
- Prone to "ecological fallacy"
- Cannot adjust well for confounding due to lack of comparability (due to lack of data on all potential covariates)
- Missing data is another concern