Novel Coronavirus Seroprevalence

A Stanford Study

David M. Rocke

April 29, 2021

"COVID-19 antibody seroprevalence in Santa Clara County, California," Eran Bendavid, Bianca Mulaney, Neeraj Sood, Soleil Shah, Rebecca Bromley-Dulfano, Cara Lai, Zoe Weissberg, Rodrigo Saavedra-Walker, Jim Tedrow, Andrew Bogan, Thomas Kupiec, Daniel Eichner, Ribhav Gupta, **John P A Ioannidis, Jay Bhattacharya**, *International Journal of Epidemiology*, 22 February 2021, dyab010.

- This was the first of a series of articles, interviews, and posts by Ioannidis and Bhattacharya (among later others) denying the severity of COVID-19, and the high infection fatality rate (see pretentiously labeled Great Barrington Declaration).
- Often, this used estimates of seroprevalence that were artificially biased up, so that the death rate would seem to have an higher denominator and thus seem to be lower than it was.
- The cause of this campaign is not clear, but may stem first from financial incentives and second from stubborn defense of what eventually became an indefensible position.

Issues in Data for COVID-19

- We need to know a number of things to understand the issues in COVID-19.
- On any given day in any given location
  - How many cases of infection have there been?
    - Still infectious?
    - Recovered and serologically positive?
  - How many new cases are there?
  - How many have died from or with the virus?
- If there are more recovered cases, including identified cases, then the (infection) fatality rate is lower and "opening" is safer.

Measuring Seroprevalence

- Antibody response to a virus includes specific immunoglobulins such as IgG and IgM.
- The test used in the Stanford study measured both and called a sample positive if either IgG or IgM tested positive.
- The test used was applied by the manufacturer to 371 samples from pre-COVID blood draws.
  - 369/371 tested negative for IgG
  - 368/ 371 tested negative for IgM
  - So the FPR for "tested positive on at least one" is estimated to be somewhere between $3/371 = 0.81\%$ and $5/371 = 1.35\%$.

Study Results

- Target is Santa Clara County.
- The recruitment method (Facebook) is not an unbiased sampling method and poststratification was used to adjust. We will look only at the crude rate.
- 50 out of 3300 tested positive on one or the other or both of IgG and IgM. This is a crude rate of $50/3300 = 1.52\%$

Is There Evidence of any True Positives?

Source	Positives	Negatives
Stanford Study	50	3250
Manufacturer	3	368

- Fisher's exact test gives $p = 0.36$. In other words, all of the Stanford positives could be false positives.
- Optimistic estimate of true positives fraction is $50/3300 - (3/371) = 0.71\%$. Compare to the reported 1.50% which is not corrected for false positives.
- The estimated number of seropositives corrected for false positive is $(3300)(0.71\%) = 23.3$

Notes

- There is an undocumented adjustment for false positives in the paper, performed after poststratification, but of much smaller magnitude that what we used.
- The adjustment for false positives should be (3/371)(3250/3300), but the latter factor is very close to 1 and its inclusion makes the confidence interval slightly more difficult to compute. In general, it would need to be included.

Unreported Infections

- The population of Santa Clara County is given in the paper as 1,943,411.
- A crude estimate of seropositives at a rate of 0.71% positives is 13,731.
- Reported confirmed cases is 956, so undercounted by an estimated factor of 14.
- Paper estimates 23,000–82,000 seropositives with an alleged undercount of a factor of 24–80.

CI for seroprevalence

The estimated fraction of seroprevalent individuals is $50/3300 - 3/371 = 0.0071$ with 95% confidence interval $(-.0045, 0.0186)$ with the negative end of the interval corresponding to the true test-positive rate being less than the false positive rate, which can only happen if there is a large false negative rate. Practically speaking the lower end is at 0. The CI for seropositive individuals is $(0, 36119)$. This used prop.test() in R. Alternative and possibly better methods are available.

Infection Fatality Ratio

- Infection Fatality Ratio (IFR) is not the Case Fatality Ratio (CFR) which requires symptoms.
- Reported deaths is at four weeks was 106.
- This is an estimated IFR of $106/13731 = 0.77\%$ compared to an estimated influenza IFR of 0.1% with CI 0.28–10.4%. (Upper end based on true infected = reported infected.)
- The estimate in the paper is 0.12–0.20%.
- This would be very encouraging, but is seriously flawed, and inconsistent with known death rates in NYC, Italy, and other places.

Other Comments

- The reported deaths from COVID can be compared to excess deaths from previous years.
- This is often 2–5 times the reported COVID deaths.
- It is quite likely that there are undiscovered cases, since possibly 20–40% are asymptomatic. This percentage was very poorly estimated at that date.
- The IFR and CFR almost certainly depend on the community practices, which were comparatively excellent in Santa Clara County. No guarantees with lifting the restrictions.

Other Comments

- The IFR also depends on age, comorbidities, and other patient and community characteristics.
- The ratio between the number dying of COVID (sometimes badly reported and hard to define) and the number infected (poorly known) is therefore a ratio between "dont know" and "don't know either".
- We need better data, and more testing.
- This is both for the PCR test for current infection and antibody tests for past infection.

Other Comments

- A year later, the known deaths from COVID-19 exceed 600,000. The US population is 328 million.
- If everyone in the US had been infected, the IFR would be 0.2%, already double that of influenza.
- Since only a fraction of the US population has been infected, the IFR is likely more on the order of 0.5%–1%.
- Yet Ioannidis still cites 0.2% as a plausible IFR.