

# Biased sampling underestimates the transmissibility of COVID-19 causing weak public policies

Tianxiao Ma

09 December, 2020

## **Abstract**

People from all around the world are gradually getting back to work, however, the rising count of COVID-19 confirmed cases is warning the public that the pandemic is still far from over. Even though most of the individuals are following the public guidelines, COVID-19 is still spreading among communities. Looking at the problem, biased sampling provides distorted data causing the public and policymakers to underestimate the severity of COVID-19. The model in the report suggests the true infected population is substantially underestimated. . .

**Keywords:** Biased sampling; COVID-19; Simulation; Contact trace; Canada

# Introduction

People from all around the world are gradually getting back to work, however, the rising count of COVID-19 confirmed cases is warning the public that the pandemic is still far from over. Even though most of the individuals are following the public guidelines, COVID-19 is still spreading among communities. Looking at the problem, biased sampling provides distorted data causing the public and policymakers to underestimate the severity of COVID-19.

Many governors underestimate the pandemic at the beginning, delays in pandemic responding harms the world as a whole, as the nature of prevalence growth is exponential. Some policymakers can't wait to implement policies more flexibly in order to boost the economy when the number of diagnoses drops, without thinking about the accuracy of resource data.

Most reports of confirmed cases rely on polymerase chain reaction-based testing of symptomatic patients (Spychalski et al., 2020). Estimates of COVID cases based on PCR could give undercoverage biased sampling for missing asymptomatic patients and patients not been tested due to the scarce testing ability. In April, a community seroprevalence study in Los Angeles County shows the prevalence of antibodies to COVID was 4.65% (Sood et al., 2020), which is substantially greater than the 0.1% confirmed infections in the county on April 10 (Bendavid et al., 2020). The observational study indicates the existence of information bias leads to a substantial underestimation of COVID infections. Besides this, the study indicates using contact tracing methods to stop the spread is challenging.

With an underestimation of the number of cases as background, contact tracing has systematic bias resulting in an underestimate of the true effect again. Contact tracers choose previously ignorant patients they can find for a limited level of effort, and only a small proportion of confirmed cases are traced. The tiny non-random sample is highly unlikely to be representative. (Andrew Whitby, 2020)

To minimize the overall loss caused by COVID, governments execute relational policies such as restrictions on gatherings. These public policies vary between countries, states and regions. Policymaking highly relies on explicit theory and/or empirical evidence. Typical evidence that policymakers pay attention to is the COVID-19 curve. However, these policies could be too weak if the curve substantially underestimates the severity due to the impact of bias.

The report aims to show how biased sampling underestimates the COVID severity by conducting a simple model comparing actual infected and observed population in a small event by adjusting the probability of getting infected with consideration of multiple infections, then predict the actual infected population in Canada with simulation. Then discuss biased sampling and its effects on the cognition of COVID and how this affects public policy.

Data

## Model

## Results

### Model Result

**Discussion**

**Implication**

**Weakness and next step**

## References

1. Andrew Whitby (2020, November 25). Contact tracing can give a biased sample of COVID-19 cases. Retrieved from <https://andrewwhitby.com/2020/11/24/contact-tracing-biased/>
2. Bendavid, E., Mulaney, B., Sood, N., Shah, S., Ling, E., Bromley-Dulfano, R., . . . Bhattacharya, J. (2020). COVID-19 Antibody Seroprevalence in Santa Clara County, California. doi:10.1101/2020.04.14.20062463
3. Sood, N., Simon, P., Ebner, P., Eichner, D., Reynolds, J., Bendavid, E., & Bhattacharya, J. (2020). Seroprevalence of SARS-CoV-2-Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. *Jama*, 323(23), 2425. doi:10.1001/jama.2020.8279
4. Spychalski, P., Błażyńska-Spychalska, A., & Kobiela, J. (2020). Estimating case fatality rates of COVID-19. *The Lancet Infectious Diseases*, 20(7), 774-775. doi:10.1016/s1473-3099(20)30246-2

## Appendices

1. Github link which contains all the code, dataset, and report for the project: <https://github.com/rubytianxiaoma/Biased-sampling-underestimates-the-transmissibility-of-COVID-19-causing-weak-public-policies>