Assessing methods to adjust for assay differences across COVID-19 serological surveillance studies in Canada

Jiacheng Chen1,2, Yuan Yu1,…, W. Alton Russell1,2

1School of Population and Global Health, McGill University, Montreal, Canada

2COVID-19 Immunity Task Force

**Corresponding author:**

**Key words:**

**Running title:**

# Abstract

**Background:** A

**Methods:** A

**Results:** A

**Conclusions:** A

# Introduction

Citation [1]. See [Figure 1](#fig-sample), [Table S 1](#cor-sample), [Table 1](#thm-sample), [Table 2](#thm-data), and [Figure S 1](#lem-sample).

The Covid-19 pandemic overview, global and what it looks like in Canada.

Serosurveillance programs in Canada, and its importance

Two types of antibodies to be detected by assays: Anti-S and Anti-N. Why we treated them differently in data analysis. Quantitative and qualitative test results.

anti-S (humoral response to infection or vaccination) and anti-N (marker of natural infection)

Waning of immunity against SARS-COV-2 after infection and immunization, indicated by decreasing blood Ab level overtime.

[insert the figure of declining Ab with re-activation of immunization]

Seropravalence and attack rate, why we need to employ RG equation or Bayesian model to estimated seropravalence, and why we also need models to adjust for waning immunity.

# Methods

**Notes**: Unlike the DBS vs VBS manuscript, the methods section of this project is mostly unfinished. This section only shows the plan and the flow chart of the project in general, while the details of it needs constantly updating and polishing overtime, as the project goes. The potential obstacles were listed below:

* APL assay data before the Omicron wave was mostly missing, Carmen said they have already approved the data access but the data is still on its way. Updates Jan-16-2024: Carmen said more data will be available soon.
* In AB Outpatient Lab Data, they changed the assay (Diasorin S –> Abbott S) in the midway, this would affect the validity of analysis based on quantitative test results, as the results from different assay are not directly comparable. Hopefully this can be clarified in the updated version of the dataset. Our currently analysis assume the Anti-S used in omicron wave are all Abbott’s, based on the paper published by Carmen in the post-vax stage: https://doi.org/10.1080/23744235.2022.2080250

## Dataset

To estimate the SARS-CoV-2 seroprevalence and attack rate in Alberta, we used data from the Alberta Outpatient Lab (APL), Canadian Blood Services (CBS) and Canadian Partnership for Tomorrow’s Health (CANPATH). Alberta Health Services (AHS) tested >160,000 APL blood draws for SARS-CoV-2 anti-N and anti-S antibodies from June 2020 until October 2022. CBS has tested >450,000 blood donations for both antibodies levels since March 2020. CANPATH has provided data with sample size > XX. Besides of assay results, all three datasets include demographic variables such as donor’s sex, self-reported race/ethnicity, and age. From their address, neighborhood material/social deprivation and urban/rural indicator have been derived.

**Notes**: *The paragraph below describes briefly about the time frame of data collection in those datasets. CBS is still actively updating (last ver downloaded on Nov 30th) and APL is also actively updating, we are currently waiting for additional linked variables from Carmen.*

The three datasets feature different data collection time frames. The CBS dataset has the most extended period, ranging from December 2020 to February 2023. APL’s data collection spans from June 2020 to October 2022, but it’s important to note that anti-N assay results are only available starting from December 2021. Meanwhile, the CANPATH dataset is accessible for a shorter duration, from September 2020 to May 2021.

We did not conduct analysis using a merged data, but instead, we conducted analysis separately for each data and presented the results in a graph panel ([Figure 2](#fig-seroprevalence)). [Table 2](#thm-data) below summarized the differences among three data utilized in the study.

[Table 2](#thm-data) CBS, CANPATH and APL

[insert 2](#thm-data)

## **Seroprevalence and attack rate**

To evaluate the severity of Covid-19 pandemic overtime in Alberta province, first we calculated the anti-N assay test positivity. Then, to bridge the gap between test positivity and population-level seroprevalence, we employed the Rogen-Gladen (RG) equation and Bayesian models. Following seroprevalence estimation, we further estimated the attack rate while adjusting for waning immunity using Bayesian models. These estimations are dessignated to answer what proportion of the target population that has been infected by SARS-CoV-2. We presented both seroprevalence and attack rate in a unified graph ([Figure 2](#fig-seroprevalence)), displaying changes over time on a monthly scale.

Rogan Gladen equation:

Bayesian models adjusting for sensitivity:

$$

$$

Bayesian models adjusting for waning immunity: $$

$$

## **NT50 and corresponding antibody level:**

In evaluating whether participants possess humoral immunity levels sufficient to shield them from future infections, from either vaccination or infection, we applied a threshold of protection to the quantitative Anti-S assay results. The threshold is cited from the package insert of each assays [Roche and Abbott], indicating sufficient antibody level to neutralize over 50% virus, with certain times of dilution (1:20). We also presented the fluctuation of the proportion sufficiently protected over time in another figure ([Figure 3](#fig-immunity)), and we conducted sensitivity analyses in which we compared alternative threshold levels. The details of the sensitivity analysis is presented in the **Fig**. Y (haven’t scratched this one yet).

[Insert 3](#fig-immunity)

# Results

ANOVA test, p-values…

# Discussion

Notes: The gap between Rogen-Gladen adjusted sero-prevalence and unadjused sero-positivity.

In our case, the sensitivity is where most variance come from. Specificity is good over time, and we shall not only take the numbers published by the manufacturer (Roche & Abbott) into consideration. Since sens is imperfect (0.8–1.0) and specificity is mostly 100% trustworthy (~1.0), the R-G equation can be simplified as:

According to the equation above, the worse the sens (smaller in value), the larger the gap between the adjusted (blue) and unadjusted (red). The gap is larger among unadjusted lines (red).

**Notes**: *Variances among data (APL, CBS, CANPATH) is not the major focus of this study, Yuan’s paper will focus on those by evaluating the representativeness of the serosurveillance data. For this paper, the major focus should be the difference between adjusting methods. The potential readers of this papers includes those who are using serosurveillance data to monitor the progress of Covid pandemic, or other pathogens of interests which possess similar features as SARS-CoV-2.*

The take-home messages would be:

1. Without any adjustment, the public health inference on test positivity is very limited. It only estimates the proportion of test positive results among blood samples collected in the serosurveillance study.
2. Adjusting for certain factors will change the public health inference, meanwhile it will also improve its generalizability. When the imperfect nature of the seroassays is taken into account, the adjusted proportion of positive test results (seroprevalence) will be able to represent the target population in our target population, from which our sample has been collected. Upon the seroprevalence, after adjusting for seroreversion using Bayesian models, we get attack rates of SARS-CoV-2, which estimates the proportion of population infected by the virus since the beginning of pandemic.
3. There’s gap among all those different measurements, and also they have different public health inference. Future researchers should be cautious when making references to those different measurements.

# Declarations

**Funding:** A

**Conflicts:** A

**Ethics/Consent:** A

**Data and materials:** A

**Code availability:** A

**Authors’ contributions: A**

# References

1. Langham S, Wright A, Kenworthy J, Grieve R, Dunlop WCN. Cost-effectiveness of take-home naloxone for the prevention of overdose fatalities among heroin users in the United Kingdom. *Value in Health*. 2018;21(4):407-415. doi:[10.1016/j.jval.2017.07.014](https://doi.org/10.1016/j.jval.2017.07.014)

# Tables

**Table 1** This is a sample table

| **col1** | **col2** |
| --- | --- |
| **A** | |
| 1 | Yes |
| 2 | Yes |
| **B** | |
| 1 | No |
| 2 | No |

**Table 2** The table below displays the distribution of age groups and gender across all three datasets. It also details the provinces where the data was collected, along with the corresponding data collection periods.

| **...1** | **CBS** | **CANPATH** | **APL** |
| --- | --- | --- | --- |
| character | character | character | character |
|  | N (%) | N (%) | N (%) |
| **Age Groups** |  |  |  |
| 18-35 | 200 (33.33%) |  |  |
| 36-64 | 201 (33.33%) |  |  |
| 65+ | 202 (33.33%) |  |  |
| **Gender** |  |  |  |
| Male |  |  |  |
| Female |  |  |  |
| **Provinces** |  |  |  |
| Data Collecting Period |  |  |  |
| n: 10 | | | |

# Figures

|  |
| --- |
| **Figure** 1: Figure caption here. |

|  |
| --- |
| **Figure** 2: The figure panel above presents the temporal changes in seroprevalence and attack rate, grouped by province and data source. Each figure displays the unadjusted seropositivity, the seroprevalence adjusted using the Rogan-Gladen equation, the seroprevalence adjusted by Bayesian models, and the attack rates. The calculations of seropositivity, seroprevalence, and attack rate are based on the qualitative results of the anti-N assay. (Just a scratch, not the real fig) |

|  |
| --- |
| **Figure** 3: The figure panel above depicts the temporal fluctuations in quantitative anti-S assay results, grouped by province and data source. Two cutoff indices (COIs) are utilized for different public health interpretations. The COI of 0.8, recommended by the manufacturer (Roche), is optimal for identifying the presence of antibodies in the blood sample. Meanwhile, a COI of 15 is indicative of a threshold for sufficient humoral immunity against the SARS-CoV-2 virus. We also applied Rogen-Gladen equation and Bayesian models to adjust for test sensitivity. (Just one fig, will add more to build a panel) |

# Supplemental materials

# A. Supplement section

## Supplement Methods

Besides of serology test results and demographic data, CBS has linked the administrative data for the cohort, which provided additional information on their hospitalizations, clinic visits, diagnoses, COVID vaccinations, and long-term disability claims. Similarly, AHS and CANPATH have completed linkages and included additional vaccination history.

**Notes**: *Current ver of APL has completed very limited data linkage, which only includes PCR test results/time and vaccination time. The paragraph briefly describe the data linkage for three datasets, however for now we do not need those variables from linked administrative data.*

# Supplemental tables

**Table S 1** Table caption here.

| **col1** | **col2** |
| --- | --- |
| **A** | |
| 1 | Yes |
| 2 | Yes |
| **B** | |
| 1 | No |
| 2 | No |

# Supplemental figures

**Figure S 1** Figure caption here.

