HOW ASSAY DIFFERENCES AND STATISTICAL ADJUSTMENT IMPACT INFERENCES FROM SARS-COV-2 SEROLOGICAL SURVEILLANCE STUDIES

Jiacheng Chen1,2, Yuan Yu1, Sheila O’Brien3,4, Carmen L. Charlton5,6,7, Steven Drews5,6, Yu Nakagama8, Yasutoshi Kido8, David L. Buckeridge1,2, W. Alton Russell1,2

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

2COVID-19 Immunity Task Force, Montreal, Canada

3Epidemiology and Surveillance, Canadian Blood Services, Ottawa, Ontario, Canada

4School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada

5Microbiology, Canadian Blood Services, Edmonton, Alberta, Canada

6Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada

7Alberta Precision Laboratories (ProvLab), University of Alberta Hospital, Edmonton, Alberta, Canada

8Graduate School of Medicine, Osaka Metropolitan University, Osaka, Japan

**Corresponding author:** W. Alton Russell, McGill University, alton.russell@mcgill.ca

**Key words:** population health surveillance, serology, assay adjustment

**Running title:** Serologic assay adjustment

# Abstract

Previous research shows substantial variability in SARS-CoV-2 seroprevalence estimates using different serological assays. The Roche Elecsys® anti-nucleocapsid (anti-N) assay demonstrates less waning sensitivity over time compared to the Abbott anti-N assay. We analyzed two Alberta serosurveillance datasets: 124,008 blood donor samples by Canadian Blood Services (CBS) and 214,780 outpatient samples by Alberta Precision Labs (APL). For anti-N analysis, we estimated seroprevalence using Rogen-Gladen (RG) adjustments and Bayesian intervals, with and without regression models. We also estimated rolling incidence using multinomial regression models. Early pandemic anti-N seropositivity was similar, but diverged after May 2022 due to greater waning sensitivity in the Abbott assay. RG adjustments had minimal impact, while regression improved concordance. For anti-S assays, 95% of blood donors and 90% of APL outpatients had seroconverted by the end of 2021, with 90% and 80% respectively achieving high antibody levels. Serosurveillance estimates are sensitive to the choice of anti-N assay, with common adjustments having limited impact. Regression-based approaches enhance comparability, and rolling incidence is a more population health-relevant measurement. Anti-S assays performed similarly between two datasets, with minor differences likely due to vaccination rates. These methods help to integrate and compare serosurveillance data across populations. (194/200 words)

# Introduction

[Be concise in this section]

[The Covid-19 pandemic overview]: Brief, general intro on Covid-19. Serosurveillance programs in Canada, and its importance. Use previous serosurveillence studies as examples.

[Two types of antibodies to be detected by assays]: Anti-spike and anti-nucleocapsid. Why we treated them differently in data analysis, one for quantitative and anothor one for qualitative test results.

* Anti-S (humoral response to infection or vaccination) and anti-N (marker of natural infection)
* Anti-S level elevates after infection and can be boosted by vax, while anti-N level elevates with infection (even drops) but didn’t change so much with vax.
* Anti-S level wanes slower than anti-N level after infection

[Ab level and immunity]: Previous studies have found an association between seropositivity/ab levels and intensity of immunity (humoral + cellular). Declining ab level suggesting waning of immunity against SARS-COV-2 after infection and immunization.

[Our measurement and adjustment]: Seropravalence and rolling incidence, why we need to employ RG equation or Bayesian model to estimated seropravalence, and why we also need probabilistic models to predict past infection using early 2020 and pre-Covid neg ref data.

# Methods

## Dataset

To estimate the SARS-CoV-2 seroprevalence and rolling incidence, we analyzed two residual blood serosurveillance datasets in Alberta, Canada: 124,008 blood donor plasma samples tested with Roche Elecsys® Anti-SARS-CoV-2 (anti-N) and Anti-SARS-CoV-2 S (anti-S) by Canadian Blood Services (CBS) and 214,780 outpatient lab plasma or sera samples tested with the Abbott SARS-CoV-2 IgG Assay (anti-N) and SARS-CoV-2 IgG QUANT (anti-S) by Alberta Precision Labs (APL).

Two datasets have different data collection time frames. The CBS dataset has the most extended period, ranging from December 2020 to February 2023. Meanwhile, APL’s data collection spans from June 2020 to October 2022. However, the anti-N assay results are only available starting from December 2021 in the APL data.

## **Seroprevalence adjusted by RG and Bayesian models**

To evaluate the severity of Covid-19 pandemic overtime in each dataset, we estimated crude seroprevalence, and seroprevalence with the Rogen-Gladen (RG) adjustment for previously reported sensitivity and specificity using Bayesian credible intervals.

Rogan Gladen equation:

Where “Prevalence” indicate RG adjusted seroprevalence estimates. “Test Positivity” indicates the proportion of test samples with positive results. “Spec” and “Sens” indicate the specificity and sensitivity of the anti-N assay, respectively.

Bayesian models adjusting for sensitivity:

[Bayesian models adjusting for waning sensitivity]: Equation (above) and what each factor indicates (What model input and output. What covariates we’ve adjusted for…).

## **Regression-based assay transformation**

Using data from samples tested with both assays early in the pandemic (in 2020), we developed linear regression models to transform CBS Roche qualitative results to Abbott and applied the Abbott cutoff index and made the opposite transformation to convert APL to Roche. Then we estimated seroprevalence in each dataset based on raw and transformed data.

[Add our transformation equation here?]

## **Rolling Incidence**

Following seroprevalence estimation, we further estimated monthly rolling incidence proportion of SARS-CoV-2 infection. We developed multinomial logistic regression models using data collected in the early pandemic to estimate the probability a tested sample was tested 0-3, 4-6, 7-9, or >9 months from infection based on the qualitative anti-N assay results. Then we applied this model to the CBS and APL datasets to estimate rolling incidence from 2021 to 2023.

## **PRNT50 and corresponding antibody level**

In evaluating whether participants possess anti-S antibody levels potentially providing protection against SARS-CoV-2 infection, first we estimated the positivity of anti-S assays among CBS blood donors and APL outpatients by applying manufacturers’ cutoffs related to the presence of antibody. Then we applied the cutoffs used by previous researchers to estimate the proportion of population with relatively high antibody titers. These cutoffs are associated with a 50% reduction at plague reduction neutralization test (PRNT) at 1:120 dilution. Additionally, we conducted stratified analysis by immunization status.

# Results

The datasets had similar anti-N seropositivity earlier in the pandemic but diverged after May 2022 due to greater waning sensitivity in the Abbott anti-N assay used by APL (Figure 1). RG and Bayesian adjustments had minimal impact, increasing seroprevalence estimates by 2.04% for the CBS Alberta subset and 7.76% for APL (Panel A), according to the estimates in October 2022. Our regression-based approaches increased concordance between estimated seropositivity (Panel B) and led to more similar estimates of rolling incidence (Panel C). The gap between CBS and APL seropositivity estimates for October 2022 was 48% using raw seropositivity, 15% after transforming CBS to Abbott, and 14% after transforming APL to Roche.

[Should we use paneled fig our separate figures?]

Regarding the results of the anti-S assay, among CBS blood donors 95% had seroconversion by the end of 2021, with over 90% achieving high antibody levels that potentially protect them from the infection of SARS-CoV-2 (Figure 2). In contrast, among APL outpatients, 90% had seroconversion at the same time, yet only 80% reached high antibody titers, leaving a wider gap (10%) between assay positivity and the proportion of population gained strong immunity. In the stratified analysis, 75% of the unvaccinated APL outpatients gained anti-S seropositivity by the end of 2021, with over 60% having high antibody titers. Among outpatients who received at least one dose of vaccine, 98% of them are anti-S positive at the same time, and over 95% achieved high antibody titers.

[we need to adjust this part to match Iris’ imputation on vax status]

# Discussion

## **RG and Bayesian models**

After applying RG and Bayesian models seroprevalence estimating, the RG adjusted seroprevalence increased 2.04% among CBS blood donors and 7.76% among APL outpatients. The wider gap in APL can be explained by lower assay sensitivity we applied. The test performance figures, including sensitivity and specificity, were estimated by previous researchers [citation: Public Health England] in England. Where sensitivities are 0.927 for Abbott anti-N assay and 0.972 for Roche anti-N assay, and specificities are mostly 1.0 (0.999 and 0.998 for Abbott and Roche anti-N assay respectively). Therefore the RG equation can be simplified as:

Where “Prevalence” indicate simplified RG adjusted seroprevalence estimates. “Test Positivity” indicates the proportion of test samples with positive results. “Sens” indicate the sensitivity of the anti-N assay.

According to **Eq**.3, the worse the sensitivity (smaller in value), the larger the gap between the adjusted and unadjusted seroprevalence curves in the figure. We found larger gaps among the Abbott assay results, as the sensitivity we used for Abbott assays are smaller, when compared to sensitivity we used for Roche assays.

## Comparison among studies

[Difference/Consistence among adjusting methods, regions in Canada] Is our finding consistent with previous studies?

## **Strength and Limitation**

[Strength and Limitation] In our sensitivity analysis, we found the regression models used for Roche–>Abbott transformation is sensitive to the ratio of positive vs negative cases in the training data.

[Public health implications]

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.      Serosurveillance estimates are sensitive to choice of assay, and the most common adjustment methods (RG) do little to make estimates more concordant, particularly later in the pandemic when differences in waning sensitivity become important.

3.

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 some differences among all those different quantitative measurements, and also they have different public health inference. Future researchers should be cautious when making references to those different measurements.

Future studies should… When using matched data to train regression models for assay results transformation, negative samples should be down-weighted if the proportion of negative cases is unbalanced high.

# Declarations

**Funding:** A

**Conflicts:** A

**Ethics/Consent:** A

**Data and materials:** A

**Code availability:** A

**Authors’ contributions: A**

# References

# 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: Figure 1 presents the raw and Rogen Gladen-adjusted seroprevalence estimated from the APL and CBS data with 95% credible intervals. The calculations of seropositivity, seroprevalence are based on the qualitative results of the anti-N assay. |

|  |
| --- |
| **Figure** 3: Figure 2 depicts raw seroprevalence and seroprevalence transformed from Roche to Abbott (APL) and Abbott to Roche (APL) using a log-linear regression model. |

|  |
| --- |
| **Figure** 4: Figure 3 presents regression-based estimate of rolling incidence (proportion of population infected in the last ever or in the past 3, 6, or 9 months). |

|  |
| --- |
| **Figure** 5: 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

[A short paragraph about weighting] Our analysis is based on individual level data. As we are calculated measurements at a monthly scale, samples sampled at different months were unweighted as they contributed independently to each month’s point prevalence estimates. While in all of those three dataset, it’s possible to have one individual being sampled multiple times within the same month, in this case we applied weights to each individual according to the number of samples they have in the same month. In regression model training, individuals who provided multiple samples were also down-weighted according to the number of samples they provided.

[A paragraph about the training data] Cite the table summarizing the features of training data (Table S1). Clarify which data has matched assay results, and what data we used for each part of analysis.

Regarding the data we used to train the regression models, we identified several studies conducted in the early phases of the Covid-19 pandemic (in 2020). Blood samples are collected from France, Jamaica, Japan and UK [**add citation**], with known sample positivity at the time of collection. The data from France and UK includes anti-N assay data for convalescent patients, the data provided by Osaka Metropolitan University in Japan included anti-N assay results for both convalescent patients and negative controls, and the data from Jamaica are used for negative controls only. Among all four datasets, the samples in Jamaica, Japan and UK were tested both for Abbott anti-N and Roche anti-N assay, which makes them suitable for transformation model training (figure).

[A short paragraph about anti-N sensitivity analysis]

[A short paragraph about anti-S sensitivity analysis]

# Supplemental tables

Add a table here summarizing the sample size, data source of the training data we used.

**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.

