



Original article

Temporal trends in disparities in COVID-19 seropositivity among Canadian blood donors

Yuan Yu,¹ Matthew J Knight,¹ Diana Gibson,¹ Sheila F O'Brien,^{2,3} David L Buckeridge ^{1,4} and W. Alton Russell ^{1,4,*}

¹School of Population and Global Health, McGill University, Montreal, Canada, ²Canadian Blood Services, Ottawa, Canada, ³School of Epidemiology and Public Health, University of Ottawa, Ottawa, Canada and ⁴COVID-19 Immunity Task Force, Montreal, Canada

*Corresponding author. School of Population and Global Health, McGill University, Suite 1200, 2001 McGill College Avenue, Montreal, QC H3A 1G1, Canada. E-mail: alton.russell@mcgill.ca

Abstract

Background: In Canada's largest COVID-19 serological study, SARS-CoV-2 antibodies in blood donors have been monitored since 2020. No study has analysed changes in the association between anti-N seropositivity (a marker of recent infection) and geographic and sociodemographic characteristics over the pandemic.

Methods: Using Bayesian multi-level models with spatial effects at the census division level, we analysed changes in correlates of SARS-CoV-2 anti-N seropositivity across three periods in which different variants predominated (pre-Delta, Delta and Omicron). We analysed disparities by geographic area, individual traits (age, sex, race) and neighbourhood factors (urbanicity, material deprivation and social deprivation). Data were from 420 319 blood donations across four regions (Ontario, British Columbia [BC], the Prairies and the Atlantic region) from December 2020 to November 2022.

Results: Seropositivity was higher for racialized minorities, males and individuals in more materially deprived neighbourhoods in the pre-Delta and Delta waves. These subgroup differences dissipated in the Omicron wave as large swaths of the population became infected. Across all waves, seropositivity was higher in younger individuals and those with lower neighbourhood social deprivation. Rural residents had high seropositivity in the Prairies, but not other regions. Compared to generalized linear models, multi-level models with spatial effects had better fit and lower error when predicting SARS-CoV-2 anti-N seropositivity by geographic region.

Conclusions: Correlates of recent COVID-19 infection have evolved over the pandemic. Many disparities lessened during the Omicron wave, but public health intervention may be warranted to address persistently higher burden among young people and those with less social deprivation.

Keywords: Serology, COVID-19, surveillance, Bayesian multi-level regression, health disparities, blood donors.

Key Messages

- Prior studies have found inequalities in COVID-19 infection by sociodemographic and geographic subgroups, but limited information is available on how these may have changed as the pandemic has evolved.
- While many disparities in markers of recent infection across sociodemographic subgroups of Canadian blood donors decreased in the recent Omicron period, seroprevalence was consistently higher among young people and rural residents of the Prairies region.
- Longitudinal serological surveillance in blood donors can be used to understand how disparities in population infection and immunity evolve over time.

Introduction

Disparities in COVID-19 infections and cases have been reported by geographic region, age, sex, racialized minority status and occupation.^{1–3} But limited information is available on how these disparities have evolved with time. A Barcelona analysis found that while low-income neighbourhoods had higher rates of laboratory confirmed cases in earlier waves, this trend reversed March to November 2022.⁴ Because case-based COVID-19 surveillance captures a

fraction of symptomatic cases and omits asymptomatic cases, many countries have invested in serological surveillance.^{5,6} Many countries have begun to do SARS-COV-2 surveillance in blood donations, strengthening partnerships between blood collectors and public health.⁷ Blood donors are well-suited for population serosurveillance due to access to repeated samples from healthy individuals, low cost and established high throughput testing infrastructure.⁸ Serological

Received: 26 September 2023. Editorial Decision: 3 April 2024. Accepted: 30 May 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of the International Epidemiological Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

surveillance can reveal temporal patterns, infection prevalence and differences in disease dynamics across geographic and demographic subgroups.^{9–11}

In Canada's largest SARS-CoV-2 seroprevalence study, Canadian Blood Services has tested monthly cross-sectional samples from nine Canadian provinces since May 2020. To date, four publications have reported logistic regression analyses over different periods, most recently up until June 2022.^{12–15} Reports found that anti-N seropositivity was associated with younger age, belonging to a visible minority, being from a more materially deprived neighbourhood, or from a less socially deprived neighbourhood.^{12,13} These findings help to identify disparities in SARS-CoV-2 infection and can inform resource allocation. However, the evolution of correlates of seropositivity through Omicron has not been analysed, urban and rural differences in seropositivity have not been explored, and geographic variability in seropositivity at the sub-provincial level has not been published. Modelling geographic variability is useful because many hard-to-track factors that influence COVID-19 epidemiology differ between and within Canadian provinces: occupational and social patterns, ability to work remotely and compliance with social distancing guidelines. These unobserved factors can limit use of generalized models for understanding geographic differences in COVID-19 seropositivity. Multi-level models with spatial effects could account for spatial heterogeneity that is not explained by available covariates and potentially improving seropositivity estimation through the borrowing of information from neighbouring areas. To date, multi-level models with spatial effects have not been used for analysing Canadian SARS-CoV-2 seropositivity data.¹⁶

In this study, we used Bayesian multi-level modelling with spatial effects to analyse evolutions in the relationship between SARS-CoV-2 seropositivity and geographic, demographic and neighbourhood characteristics across pandemic periods, including the recent Omicron wave. Our analysis uncovered shifts in COVID-19 health disparities, illustrating the value of blood donor serosurveillance and multi-level modelling to understanding seropositivity across diverse populations.

Methods

Using a Bayesian generalized linear model (GLM) framework with logit link function, we analysed changes in individual- and neighbourhood-level correlates of SARS-CoV-2 anti-N positivity over time. Using GLM models with fixed effects as a comparator, we assessed multi-level (hierarchical) models (MLM) with unstructured and structured random effects. Across four regions and three pandemic waves, we compared model fit, assessed seropositivity by census division and analysed correlations between individual and neighbourhood characteristics and seropositivity.

Data

We analysed the Canadian Blood Services SARS-CoV-2 seroprevalence study data from December 2020 to December 2022 in nine provinces of Canada (Québec and the territories were excluded). Samples were randomly selected for serologic testing from approximately last two weeks of each month

from January 2021 to December 2021 and from all weeks from January 2022 to November 2022. As previously described,¹³ a straight random sample was applied until June 2021, after which stratification by age group was applied to enable a smaller monthly sample size. Samples were tested for anti-nucleocapsid antibodies using the Roche Elecsys[®] Anti-SARS-CoV-2 assay. We extracted individuals' age, sex, self-reported race/ethnicity dichotomized into white and racialized minority and postal code from donation records. We used postal code to classify donors as living in an urban or rural area and assign donors to a quintile of the Pampalon Material and Social Deprivation Indices.¹⁷ Material deprivation is a neighbourhood measure of socioeconomic deprivation that captures lack of access to essential material resources that can impact one's quality of life and well-being; social deprivation is a neighbourhood measure based on education, employment, income, housing, family structure and access to social services.

Modelling

We developed separate models for four regions: Ontario, British Columbia (BC), the Prairies (Alberta, Saskatchewan and Manitoba) and the Atlantic (New Brunswick, Newfoundland and Labrador, Nova Scotia and Prince Edward Island). Within each region, we developed separate models for three periods based on pandemic waves: pre-Delta (14 December 2020–31 July 2021), Delta (1 August 2021–14 December 2021) and Omicron (15 December 2021–30 November 2022). While our dataset contained each donor's forward sortation area, the first three digits of their postal code, this level of geographic aggregation was too fine for multi-level modelling. We instead classified donors into census divisions using the postal code conversion file based on forward sortation area. When individuals' forward sortation areas spanned more than one census division (17.28% of donations), we assigned a census division randomly in proportion to the population in each division. We excluded donations with missing demographic or geographic information and donors living in census divisions in which fewer than five donations were tested (often, this occurs when a donor works or studies in an urban area but retains a rural address in their donation record). Our Bayesian models regressed individual characteristics (age, sex, racialized minority status) and neighbourhood characteristics (urbanicity, social deprivation, material deprivation) on SARS-CoV-2 anti-N serological test outcome. We included sample month as a continuous covariate to account for the time trend of increasing positivity during the pandemic waves, adjusting for bias that could arise through temporal shifts in subgroup representation. For each region and wave, we developed three models: a GLM with a logit link function with fixed effects (GLM); a multi-level model with regular fixed effects and census division random effects (MLM random effects), which did not consider information on the spatial relationship between census divisions; and a multi-level model with structured random effects through conditional autoregressive (CAR) prior,¹⁸ which used a binary, first order adjacency matrix to denote the adjacency between census divisions (MLM spatial).^{19–22} Our model specifications were as follows:

GLM

$$y_{kj} \sim \text{Bernoulli}(\theta_{kj}), \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$\ln(\theta_{kj}/(1 - \theta_{kj})) = \mathbf{x}_{kj}^\top \boldsymbol{\beta}, \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$p(\beta_p) \propto 1, \quad -\infty < \beta_p < \infty, \quad p = 0, \dots, 7$$

MLM random effects

$$y_{kj} \sim \text{Bernoulli}(\theta_{kj}), \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$\ln(\theta_{kj}/(1 - \theta_{kj})) = \mathbf{x}_{kj}^\top \boldsymbol{\beta} + \alpha_k, \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$p(\beta_p) \propto 1, \quad -\infty < \beta_p < \infty, \quad p = 0, \dots, 7$$

$$\alpha_k \sim N(\alpha_0, \sigma_k), \quad k = 1, \dots, K,$$

$$\alpha_0 \sim t(3, 0, 2.5), \quad k = 1, \dots, K,$$

$$\sigma_k \sim t(3, 0, 2.5), \quad \sigma_k > 0, \quad k = 1, \dots, K$$

MLM spatial

$$y_{kj} \sim \text{Bernoulli}(\theta_{kj}), \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$\ln(\theta_{kj}/(1 - \theta_{kj})) = \mathbf{x}_{kj}^\top \boldsymbol{\beta} + \phi_k, \quad k = 1, \dots, K, \quad j = 1, \dots, n_k$$

$$p(\beta_p) \propto 1, \quad -\infty < \beta_p < \infty, \quad p = 0, \dots, 7$$

$$\phi_k | \phi_i, \tau_k \sim N\left(\frac{\sum_{k \sim i} \phi_k}{d_{k,k}}, \frac{1}{d_{k,k} \tau_k}\right), \quad k = 1, \dots, K, \quad i \neq k$$

$$(1/\tau_k)^{1/2} \sim t(3, 0, 2.5), \quad k = 1, \dots, K$$

where y_{kj} denotes a binary outcome of the anti-N serological test by individual j in census division k , which follows the Bernoulli distribution with probability θ_{kj} , and is linked to a linear combination of the fixed effect covariates $\boldsymbol{\beta}$ through a logit function. Coefficients β_0 to β_7 represent the intercept, age, sex, race, material and social deprivation index quantiles, urban and month. In the MLM setting, α_k has a weakly-informed flat prior denoting the census division random effect. In the MLM spatial model setting, spatial variable ϕ_k is normally distributed, and each region's value equals the average of its neighbours. Variable $d_{k,k}$ is the number of neighbours for region k . Its variance is defined as the reciprocal of $d_{k,k} \tau_k$, so that uncertainty decreases as the number of neighbours increases, following a standard method for spatial modelling.¹⁸ We used a weakly-informed prior $t(3, 0, 2.5)$ on the inverse square-root of the precision parameter τ_k . Analysis was performed using the brms package in R²³ and code are published.²⁴

Results

Blood donors population

After excluding 4357 (0.9%) donations for missing race and 60 781 (12.5%) for missing material and social deprivation, our dataset contained 420 319 donations from 239 346 donors: 186 745 (44.4%) donations in Ontario, 67 593 (16.1%) in BC, 121 984 (29.0%) in the Prairies and 43 998 (10.5%) in the Atlantic region (Table 1). Regions had similar proportions of female donors (42.1–44.5%) and age distributions (39.4–46.0% of donors older than 55). About 22.9% of the blood donor cohort was aged 55–64, compared to

16.5% of the Canadian population age 17 and older.²⁵ The Prairies had the largest proportion of donors from more materially deprived neighbourhoods. The Atlantic region had the lowest proportion of donations from members of a racialized minority group and the most rural donors; BC had the most donations from members of a racialized minority group and the fewest rural donors. Within regions, the distribution of covariates was similar across waves (Supplementary Tables S1–S3, available as Supplementary data at IJE online).

While the switch to stratified sampling during the pre-Delta period slightly changed the age distribution (Supplementary Table S4, available as Supplementary data at IJE online), it did not appear to have impacted trends in seropositivity when stratifying by age, sex and race (Supplementary Figures S1 and S2, available as Supplementary data at IJE online). Moran's I test showed strong spatial autocorrelation among census divisions' seropositivity during Omicron in all regions except the Atlantic; Moran's I was 0.21 ($P < 0.01$) for Ontario, 0.11 ($P < 0.01$) for BC, 0.10 ($P < 0.05$) for Prairies and -0.01 ($P = 0.44$) for the Atlantic. For earlier waves, Moran's I test was only significant for Ontario in the Delta wave (0.20 [$P < 0.05$]). For all period-region combinations the R -hat value was below 1.01, suggesting successful convergence, for all parameters.

Model comparison

The GLM model had the highest widely applicable information criterion (WAIC) score, indicating poorer balance of goodness-of-fit and complexity, for all region-wave combinations except for the Atlantic region during the Delta wave which had only 31 positive cases. In the 11 other settings, the MLM spatial model had a lower WAIC score than the MLM random effects model, suggesting that considering adjacency of census divisions improved estimation, although the difference was sometimes small (Table 2). Using the rule-of-thumb suggested by Spiegelhalter *et al.*,²⁶ the MLM spatial model fit was 'about the same' as MLM random effects in 4 of 11 settings (WAIC score improved by less than 2); MLM spatial model was 'different but similar' in 3 of 11 settings (WAIC score improved by 2–7) and the MLM spatial model was 'different' in 4 of 11 (WAIC score improved by more than 7). More informative priors may have led to a greater performance improvement with spatial effects. The posterior mean was closer to the crude seropositivity in most census divisions using the MLM random effects model compared to the GLM model, except in the Atlantic region before omicron (Figure 1, Supplementary Figures S3 and S4, available as Supplementary data at IJE online). The 95% highest posterior density interval included the raw proportion seropositive for every census division for each MLM spatial model. This was not always true for the GLM model, suggesting that including spatial effects improved sub-regional seropositivity estimation. Given similar performance between the MLM random effects and spatial models, we restrict our comparison to the MLM spatial models and GLM models for the remainder of the article.

Predictors of seropositivity by region and wave

Across regions and waves, coefficient point estimates and uncertainty ranges were similar whether the GLM or MLM spatial model was used (Figure 2). The coefficient on month of

Table 1. Distribution of individual and neighborhood characteristics by region

	Overall	Atlantic	BC	Ontario	Prairies
Number	420 319	43 998	67 592	186 745	121 984
Sex (%)					
Female	180 859 (43.0)	19 597 (44.5)	29 589 (43.8)	80 351 (43.0)	51 322 (42.1)
Male	239 460 (57.0)	24 401 (55.5)	38 003 (56.2)	106 394 (57.0)	70 662 (57.9)
Age group, years (%)					
16–24	31 983 (7.6)	2806 (6.4)	4639 (6.9)	14 487 (7.8)	10 051 (8.2)
25–34	73 564 (17.5)	6286 (14.3)	12 525 (18.5)	33 502 (17.9)	21 251 (17.4)
35–44	71 834 (17.1)	6675 (15.2)	11 993 (17.7)	31 312 (16.8)	21 854 (17.9)
45–54	73 694 (17.5)	7959 (18.1)	11 487 (17.0)	33 900 (18.2)	20 348 (16.7)
55–64	96 046 (22.9)	11 227 (25.5)	14 468 (21.4)	42 938 (23.0)	27 413 (22.5)
65–74	62 027 (14.8)	7616 (17.3)	10 294 (15.2)	26 090 (14.0)	18 027 (14.8)
75+	11 171 (2.7)	1429 (3.2)	2186 (3.2)	4516 (2.4)	3040 (2.5)
Race (%)					
Minority	70 935 (16.9)	2881 (6.5)	16 453 (24.3)	33 163 (17.8)	18 438 (15.1)
White	349 384 (83.1)	41 117 (93.5)	51 139 (75.7)	153 582 (82.2)	103 546 (84.9)
Social deprivation quintile (%)					
1	89 390 (21.3)	5466 (12.4)	12 859 (19.0)	43 248 (23.2)	27 817 (22.8)
2	89 168 (21.2)	10 818 (24.6)	14 473 (21.4)	41 603 (22.3)	22 274 (18.3)
3	84 219 (20.0)	11 380 (25.9)	13 149 (19.5)	36 019 (19.3)	23 671 (19.4)
4	78 222 (18.6)	9111 (20.7)	11 889 (17.6)	32 452 (17.4)	24 770 (20.3)
5	79 320 (18.9)	7223 (16.4)	15 222 (22.5)	33 423 (17.9)	23 452 (19.2)
Material deprivation quintile (%)					
1	126 094 (30.0)	7754 (17.6)	19 775 (29.3)	49 068 (26.3)	49 497 (40.6)
2	103 314 (24.6)	9821 (22.3)	17 959 (26.6)	45 895 (24.6)	29 639 (24.3)
3	85 695 (20.4)	9145 (20.8)	13 784 (20.4)	41 023 (22.0)	21 743 (17.8)
4	66 536 (15.8)	9930 (22.6)	10 251 (15.2)	32 068 (17.2)	14 287 (11.7)
5	38 680 (9.2)	7348 (16.7)	5823 (8.6)	18 691 (10.0)	6818 (5.6)
Urban (%)					
Rural	55 623 (13.2)	7805 (17.7)	3117 (4.6)	29 851 (16.0)	14 850 (12.2)
Urban	364 696 (86.8)	36 193 (82.3)	64 475 (95.4)	156 894 (84.0)	107 134 (87.8)

BC, British Columbia.

Distribution of individual and neighbourhood characteristics for donations tested across all three waves. For social and material deprivation quintile, 1 is the lowest level of neighbourhood deprivation and 5 is the highest level.

Table 2. WAIC (widely applicable information criterion) score from different models by type, region and wave

Region-Wave	GLM	MLM random effects	MLM spatial
Ontario-Pre-Delta	10 953.04	10 792.09	10 778.91**
Ontario-Delta	3 262.97	3 212.39	3 211.86
Ontario-Omicron	161 243.10	160 681.99	160 674.40**
BC-Pre-Delta	3 539.64	3 494.69	3 493.32
BC-Delta	1 816.80	1 793.20	1 786.7*
BC-Omicron	56 952.86	56 580.27	56 578.95
Prairies-PreDelta	9 967.58	9 885.03	9 881.98*
Prairies-Delta	5 432.21	5 357.38	5 354.64*
Prairies-Omicron	99 922.69	99 358.87	99 351.41**
Atlantic-PreDelta	501.87	500.22	500.19
Atlantic-Delta	354.82	356.90	356.11
Atlantic-Omicron	31 353.12	31 233.64	31 223.49**

BC, British Columbia.

WAIC score from different models per region/wave. Lower WAIC scores indicate a better model fit. The lowest score for each region-wave combination is shown in bold. Using the rule-of-thumb suggested by Spiegelhalter *et al.*,²⁶ improvements in WAIC scores of 2 or less indicate 'similar' model fit; improvements in WAIC scores between 2 and 7 indicate 'different but similar' model fit (*), and improvements in WAIC scores greater than 7 indicate 'different' model fit (**).

sampling had a positive value in all models, indicating increasing seropositivity over time during the wave. Each region's coefficient was higher during Omicron (0.272–0.35 monthly increase in seropositivity) compared to pre-Delta (0.144–0.199). Coefficients during Delta ranged from 0.007

(Ontario) to 0.398 (Atlantic) (Supplementary Figure S5, available as Supplementary data at *IJE* online).

Neighbourhood material deprivation was positively associated with seropositivity in all regions and waves except for the Atlantic. Social deprivation was often negatively associated with seropositivity, possibly due to the protective effect of smaller average household sizes. The coefficient for urbanicity was not statistically significant in most regions; however, we estimated a large negative coefficient for the Prairies, suggesting that seropositivity was higher for rural individuals. Analysing the same coefficients over shorter 3-month periods, we found that urbanites in Ontario had higher seropositivity, but only in the earliest 3-month period of March–August 2021 (Supplementary Figure S6, available as Supplementary data at *IJE* online).

Younger age, male sex and racialized minority status were associated with greater seropositivity in most regions across each wave. As an exception, we found no significant association with sex and race in the Prairies during the pre-Delta and Delta waves. Coefficients for material deprivation, sex and race were smaller in the Omicron wave, suggesting that subgroup differences became less pronounced as more individuals were infected. In contrast, the coefficients on age were similar across waves. Analysing the same coefficients over shorter 3-month periods further confirmed that subgroup differences shrunk as the pandemic unfolded for urbanicity (Supplementary Figure S6, available as Supplementary data at *IJE* online), material deprivation (Supplementary Figure S7,

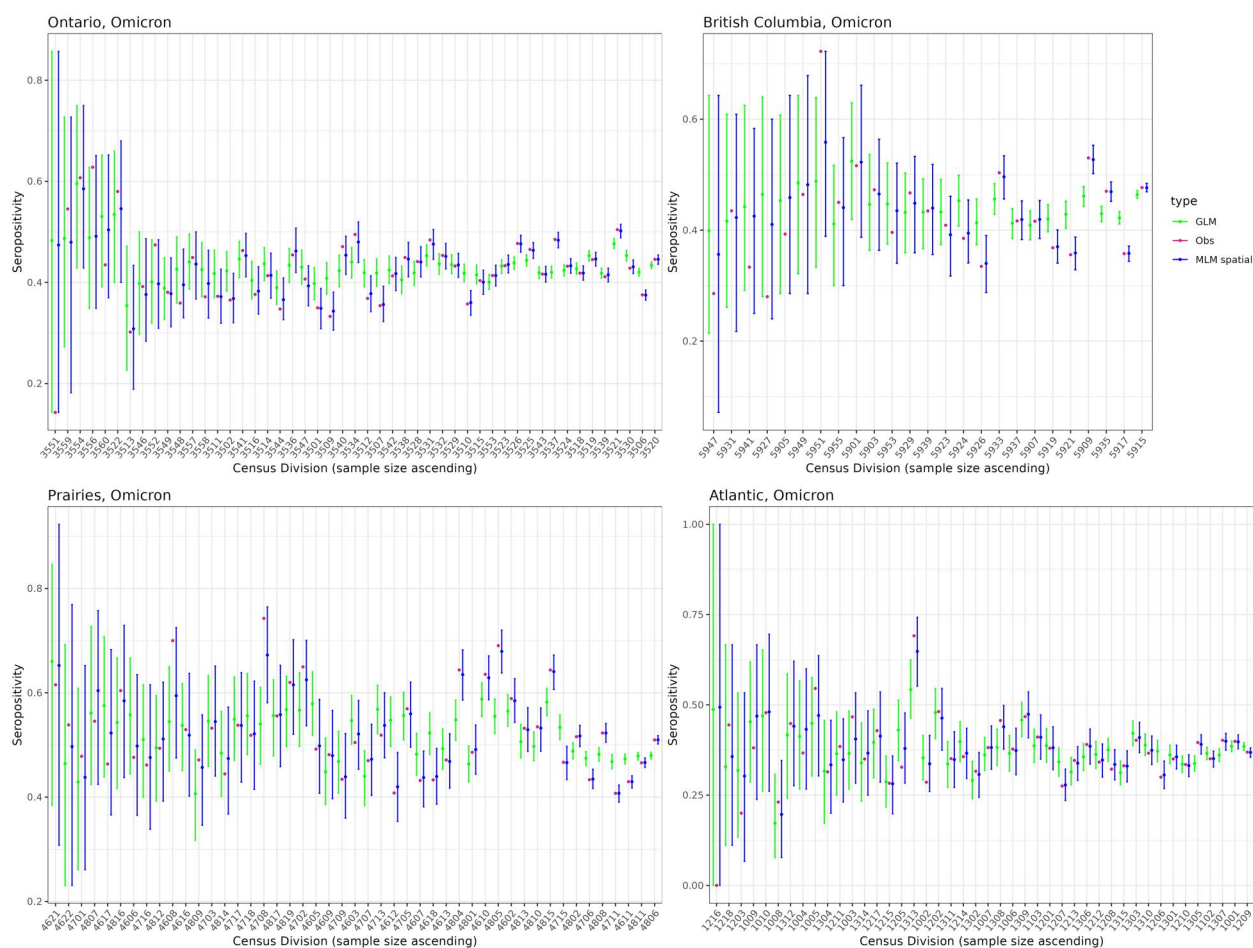


Figure 1. Posterior predicted seropositivity estimation during Omicron. Percent seropositivity by census division estimated using a multi-level model with spatial effects (MLM spatial), a generalized linear model (GLM) and the crude observed seropositivity (Obs). In 43 of 49 census divisions in Ontario, 23 of 25 in British Columbia, 43 of 47 in Prairies and 36 of 44 in Atlantic, seropositivity was more accurate using MLM spatial compared to GLM

available as [Supplementary data](#) at *IJE* online) and sex ([Supplementary Figure S8](#), available as [Supplementary data](#) at *IJE* online), but not for age ([Supplementary Figure S9](#), available as [Supplementary data](#) at *IJE* online) or social deprivation ([Supplementary Figure S10](#), available as [Supplementary data](#) at *IJE* online). Subgroup differences by race appeared to shrink in BC and Ontario ([Supplementary Figure S11](#), available as [Supplementary data](#) at *IJE* online). Analysing crude percent seropositivity by strata led to similar insights regarding correlates of seropositivity for both individual characteristics ([Supplementary Figure S12](#), available as [Supplementary data](#) at *IJE* online) and neighbourhood characteristics ([Supplementary Figure S13](#), available as [Supplementary data](#) at *IJE* online).

Discussion

Our study is the first to analyse shifts in seropositivity by sociodemographic subgroup across Canada through the Omicron period. Building on earlier analyses of the Canadian Blood Services COVID-19 serological surveillance study, we identified similar associations between individual covariates (age, race, sex) and seropositivity in the pre-Delta and Delta periods.²⁷ During Omicron, we found that male sex and racialized minority status were still associated with higher seropositivity in many regions, but with a much smaller effect

size. This suggests that disparities in COVID-19 infection by race and sex have lessened. In contrast, we found that young age remained highly associated with seropositivity through the omicron wave; coefficients in the Delta wave suggest ~ 0.02 lowered log odds of seropositivity for every one-year increase in age; this increased to ~ 0.03 in the Omicron wave.

Regarding neighbourhood characteristics, material deprivation was positively associated with anti-N seropositivity, while social deprivation was negatively associated. This may be because essential workers who cannot work remotely are more likely to live in materially deprived neighbourhoods,²⁸ and socially deprived neighbourhoods are protected by smaller household sizes. Interestingly, the association between seropositivity and high material deprivation lessened during the more recent Omicron wave, while the association with low social deprivation persisted. Living in a rural area was strongly associated with seropositivity in only one region: the Prairies. Calculating crude seropositivity by subgroup confirmed exceptionally high seropositivity for rural residents in the Prairies ([Supplementary Figure S12](#), available as [Supplementary data](#) at *IJE* online). High seropositivity for rural residents was also observed when analysing each province in the Prairies region separately (data not shown). Rural Prairies residents may have had greater occupational exposure; for instance, outbreaks were reported in meat-packing facilities in the pre-Delta wave.²⁹ During the pre-donation

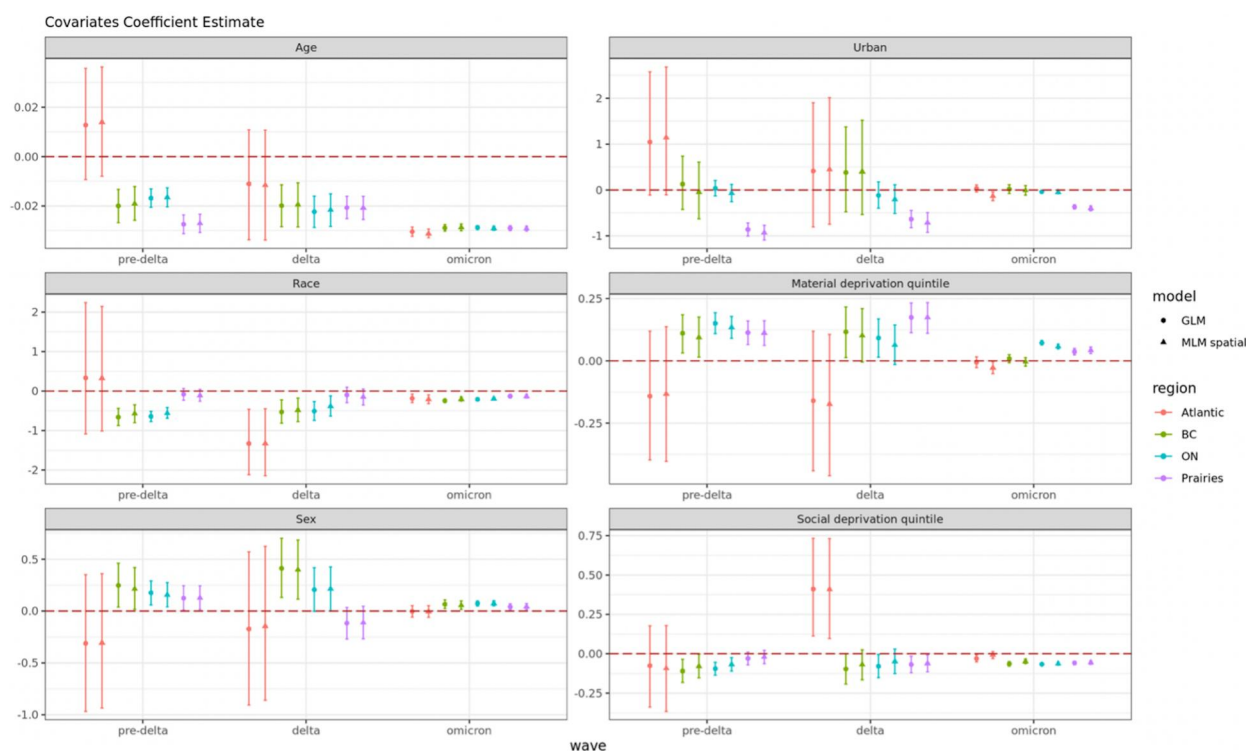


Figure 2. Covariate coefficient estimates by regions and wave. Coefficient estimate of individual- (age, race and sex) and neighbourhood-level (urban, material & social deprivation quintile) covariates across three waves from two models: multi-level models with spatial effect (MLM spatial) and generalized linear models (GLM) for Ontario (ON), British Columbia (BC), the Prairies and the Atlantic region. Coefficients greater than zero indicate a positive association with seropositivity

questionnaire, rural blood donors in the Prairies reported recent COVID-19 vaccination at a far lower rate than any other subgroup (Supplementary Figures S14 and S15, available as Supplementary data at *IJE* online). Low vaccination rates among rural Prairies residents, also reported elsewhere,³⁰ may partly explain disproportionately high seropositivity during the Delta and Omicron waves and could suggest less concern for risk mitigation, which could have contributed to higher seropositivity before vaccines were available. Research in populations other than blood donors may improve understanding of seropositivity for rural Prairies residents.

Comparing models, we found that the spatial MLM model, which incorporates correlated random effects between adjacent census divisions, had the best fit. This is consistent with previous research on spatial modelling in other contexts,²² though we found that the posterior predicted seropositivity was similar in our MLM with regular random effects. Comparing the spatial MLM and GLM models, we found that the coefficients on individual and neighbourhood covariates were generally similar. That said, the significance of coefficients for material and social deprivation sometimes differed between the GLM and spatial MLM model, suggesting neighbourhood covariates may be more sensitive to choice of a spatial model.

The blood donor cohort is large and includes many geographic areas, but donor eligibility rules, self-selection and lack of coverage for remote areas (e.g. the Canadian territories) can limit generalizability.^{31–33} A so-called ‘healthy donor effect’ has been documented wherein blood donors have better self-rated health and lower incidence of cardiovascular disease and cancer.^{31,34} Generalizability of donors differs by

health condition³⁵ and is not well-characterized for respiratory illnesses. Previous studies have found that population seroprevalence is sensitive to sampling frame, location, occupation and other demographic characteristics.^{27,36} Despite these selection biases, our findings for the earlier pandemic period are consistent with findings from other cohorts. General population meta-analyses found similar subgroup differences, including lower seropositivity among older adults and white racial groups in Europe and the Americas in the pre-Omicron period.^{35,37} Still, statistical adjustment and integrating data from other surveillance populations may improve representativeness and generalizability.³⁸

Our analysis has other limitations. Firstly, the MLM spatial model was at the level of census division because smaller areas were too fine for consistent model convergence. Because our neighbourhood covariates were defined at a more granular geographic area, we modelled them as covariates at the individual level. With sufficient data and computation power, future analyses could use finer geographic modelling. Secondly, census division-level seropositivity estimation may be improved through mapping directly from postal code to census division instead of from forward sortation area. Thirdly, while the spatial MLM model had the best fit to the data, incorporating spatial correlation can make model convergence challenging when there are few positive cases. Fourthly, our modelling framework did not account for repeated measures from the same individual, though a sensitivity analysis where we randomly selected one sample per donor in each wave was consistent with our primary analysis (Supplementary Figure S16, available as Supplementary data at *IJE* online). Due in part to limits on frequency of blood

donation, the number of samples included per donor were fairly low (1.20 for pre-Delta, 1.16 for Delta, 1.54 for Omicron).

In sum, we found that serological surveillance in blood donors can produce valuable insights into population infection and immunity, and Bayesian MLM can improve estimated seropositivity by geographic area. Extending prior analyses into the Omicron wave, we found that disparities in seropositivity waned for most individual and neighbourhood characteristics as larger swaths of the population became infected. As an exception, heightened seropositivity persisted through the Omicron wave for younger individuals, individuals in socially deprived neighbourhoods and rural residents in the Prairies region. Our analyses demonstrate the value of serosurveillance for monitoring impacts of an epidemic across demographic subgroups. Further research is warranted to refine methods for analysing blood donor surveillance data, strengthen use of these data for public health decision-making and explore blood donor serosurveillance for other pathogens.

Ethics approval

The SARS-CoV-2 seroprevalence study was approved by the Canadian Blood Services Research Ethics Board, and this secondary analysis was approved by the McGill University Research Ethics Board (study number 22-03-077).

Data availability

Data may be made available upon request from Canadian Blood Services (contact person: Sheila O'Brien), subject to internal review, privacy legislation, data sharing agreements and research ethics approval.

Code availability

All code is publicly available at www.doi.org/10.5281/zenodo.10680580.

Supplementary data

[Supplementary data](#) are available at *IJE* online.

Author contributions

WAR and YY conceived and designed the analysis with contributions from SO and DB. YY, MK and DG conducted data analysis and statistical modelling. YY, MK and WAR drafted the manuscript. All authors contributed to interpretations of the results and edited the final manuscript.

Funding

This work is funded by the Government of Canada through the COVID-19 Immunity Task Force and by Canadian Blood Services (CBS). Ce projet a été soutenu par un financement du Gouvernement du Canada, par le biais du Secrétariat du groupe de travail sur l'immunité COVID-19 et par le Société Canadienne du Sang. Dr. Russell was supported by a salary award (Chercheur-boursier) from the Fonds de Recherche du Québec – Santé.

Conflict of interest

None declared.

References

- Martinez-Beneito MA, Mari-Dell'Olmo M, Sánchez-Valdivia N *et al.* Socioeconomic inequalities in COVID-19 incidence during the first six waves in Barcelona. *Int J Epidemiol* 2023; 52:1687–95.
- Sannigrahi S, Pilla F, Basu B, Basu AS, Molter A. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustain Cities Soc* 2020;62:102418.
- van Ingen T, Brown KA, Buchan SA *et al.* Neighbourhood-level socio-demographic characteristics and risk of COVID-19 incidence and mortality in Ontario, Canada: a population-based study. *PLoS One* 2022;17:e0276507.
- Mariné Barjoan E, Chaarana A, Festraëts J *et al.* Impact of social and demographic factors on the spread of the SARS-CoV-2 epidemic in the town of Nice. *BMC Public Health* 2023;23:1098.
- Manuel DG, Delatolla R, Fisman DN *et al.* The Role of Wastewater Testing for SARS-CoV-2 Surveillance. Ontario COVID-19 Science Advisory Table. 2021. <https://doi.org/10.47326/ocsat.2021.02.40.1.0> (23 April 2024, date last accessed).
- Ward H, Whitaker M, Flower B *et al.* Population antibody responses following COVID-19 vaccination in 212,102 individuals. *Nat Commun* 2022;13:907.
- Lewin A, Osiowy C, Erikstrup C *et al.* Research partnerships between blood services and public health authorities: an international, cross-sectional survey. *Vox Sang* 2022; 117:1368–74.
- O'Brien SF, Drews SJ, Lewin A, Osiowy C, Drebot MA, Renaud C. Canadian blood suppliers: an expanding role in public health surveillance? *Can Commun Dis Rep* 2022;48:124–30.
- Saeed S, Drews SJ, Pambrun C, Yi Q-L, Osmond L, O'Brien SF. SARS-CoV-2 seroprevalence among blood donors after the first COVID-19 wave in Canada. *Transfusion* 2021;61:862–72.
- Charlton CL, Nguyen LT, Bailey A *et al.* Pre-vaccine positivity of SARS-CoV-2 antibodies in Alberta, Canada during the first two waves of the COVID-19 pandemic. *Microbiol Spectr* 2021; 9:e00291-21.
- O'Brien SF, Goldman M, Drews SJ. An expanded role for blood donor emerging pathogens surveillance. *CMAJ* 2023;195:E16.
- Reedman CN, Drews SJ, Yi Q-L, Pambrun C, O'Brien SF. Changing patterns of SARS-CoV-2 seroprevalence among Canadian blood donors during the vaccine era. *Microbiol Spectr* 2022;10:e00339-22.
- O'Brien SF, Caffrey N, Yi Q-L *et al.* Cross-Canada variability in blood donor SARS-CoV-2 seroprevalence by social determinants of health. *Microbiol Spectr* 2023;11:e03356-22.
- O'Brien S, Reedman C, Yi Q-L, Drews S. SARS-CoV-2 seroprevalence in the vaccine era: the Canadian Blood Services Serosurvey January to November 2021. *Transfus Med Rev* 2023;37:42.
- O'Brien SF, Caffrey N, Yi Q-L, Pambrun C, Drews SJ. SARS-CoV-2 seroprevalence among Canadian blood donors: the advance of Omicron. *Viruses* 2022;14:2336.
- Murphy TJ, Swail H, Jain J *et al.* The evolution of SARS-CoV-2 seroprevalence in Canada: a time-series study, 2020–2023. *CMAJ* 2023;195:E1030–37.
- Pampalon R, Hamel D, Gamache P, Philibert MD, Raymond G, Simpson A. An area-based material and social deprivation index for public health in Québec and Canada. *Can J Public Health* 2012;103:S17–22. <https://www.jstor.org/stable/41995684>. (21 July 2023, date last accessed).
- Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. *Ann Inst Stat Math* 1991; 43:1–20.
- Bryan ML, Jenkins SP. Multilevel modelling of country effects: a cautionary tale. *Eur Sociol Rev* 2016;32:3–22.

20. Earnest A, Morgan G, Mengersen K, Ryan L, Summerhayes R, Beard J. Evaluating the effect of neighbourhood weight matrices on smoothing properties of conditional autoregressive (CAR) models. *Int J Health Geogr* 2007;6:54.
21. Duncan EW, White NM, Mengersen K. Spatial smoothing in Bayesian models: a comparison of weights matrix specifications and their impact on inference. *Int J Health Geogr* 2017;16:47.
22. Xu H. Comparing spatial and multilevel regression models for binary outcomes in neighborhood studies. *Sociol Methodol* 2014;44:229–72.
23. Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. *J Stat Soft* 2017;80:1–28.
24. Yu Y, Russell WA. *altonrus/covid_donors_3waves: Code Repository for COVID-19 Seropositivity Across Sociodemographic and Geographic Subgroups of Canadian Blood Donors through the Omicron Period*. Published Online 19 February 2024. <https://zenodo.org/records/10680581> (23 April 2024, date last accessed).
25. Statistics Canada. *Table 17–10–0005–01 population estimates on July 1st, by age and sex*. 2018. <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501> (23 April 2024, date last accessed).
26. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol* 2002;64:583–639.
27. Skowronski DM, Kaweski SE, Irvine MA *et al*. Serial cross-sectional estimation of vaccine-and infection-induced SARS-CoV-2 seroprevalence in British Columbia, Canada. *CMAJ* 2022;194:E1599–609.
28. Lefebvre G, Haddad S, Moncion-Groulx D, Saint-Onge M, Dontigny A. Socioeconomic disparities and concentration of the spread of the COVID-19 pandemic in the province of Quebec, Canada. *BMC Public Health* 2023;23:1096.
29. Canadian Medical Association. *Containing COVID-19: How a Calgary Health Team Helped Suppress an Outbreak at a Meat-Packing Plant*. 2021. <https://www.cma.ca/clinical-blog/contain-ing-covid-19-how-calgary-health-team-helped-suppress-outbreak-meat-packing> (4 July 2023, date last accessed).
30. Paudel YR, Du C, MacDonald SE. The influence of place on COVID-19 vaccine coverage in Alberta: a multilevel analysis. *PLoS One* 2022;17:e0276160.
31. Atsma F, Veldhuizen I, Verbeek A, Kort W D, Vegt F D Healthy donor effect: its magnitude in health research among blood donors. *Transfusion* 2011;51:1820–28.
32. Tuite AR, Fisman D, Abe KT *et al*. Estimating SARS-CoV-2 seroprevalence in Canadian Blood Donors, April 2020 to March 2021: improving accuracy with multiple assays. *Microbiol Spectr* 2022;10:e02563-21.
33. Goldfarb DM, Mâsse LC, Watts AW *et al*. SARS-CoV-2 seroprevalence among Vancouver public school staff in British Columbia, Canada: a cross-sectional study. *BMJ Open* 2022;12:e057846.
34. van den Hurk K, Zalpuri S, Prinsze FJ, Merz EM, de Kort WLAM. Associations of health status with subsequent blood donor behavior—an alternative perspective on the Healthy Donor Effect from Donor InSight. *PLoS One* 2017;12:e0186662.
35. Bergeri I, Whelan MG, Ware H *et al*. Global SARS-CoV-2 seroprevalence from January 2020 to April 2022: a systematic review and meta-analysis of standardized population-based studies. *PLoS Med* 2022;19:e1004107.
36. Bobrovitz N, Arora RK, Cao C *et al*. Global seroprevalence of SARS-CoV-2 antibodies: a systematic review and meta-analysis. *PLoS One* 2021;16:e0252617.
37. Chen X, Chen Z, Azman AS *et al*. Serological evidence of human infection with SARS-CoV-2: a systematic review and meta-analysis. *Lancet Glob Health* 2021;9:e598–609.
38. O'Brien SF, Drews SJ, Lewin A, *et al*. How do we decide how representative our donors are for public health surveillance? *Transfusion* 2022;62:2431–37.