



RESEARCH ARTICLE

REVISED Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada [version 2; peer review: 1 approved, 2 approved with reservations]

Eliseos J. Mucaki^{1,2}, Ben C. Shirley², Peter K. Rogan ¹⁻³

¹Department of Biochemistry, University of Western Ontario, London, Ontario, N6A 5C1, Canada

²CytoGnomix Inc, London, Ontario, N5X 3X5, Canada

³Department of Oncology, University of Western Ontario, London, Ontario, N6A 5C1, Canada

V2 First published: 23 Dec 2021, 10:1312 <https://doi.org/10.12688/f1000research.75891.1>
 Latest published: 21 Jul 2022, 10:1312 <https://doi.org/10.12688/f1000research.75891.2>

Abstract

Introduction: This study aimed to produce community-level geospatial mapping of confirmed COVID-19 cases in Ontario Canada in near real-time to support decision-making. This was accomplished by area-to-area geostatistical analysis, space-time integration, and spatial interpolation of COVID-19 positive individuals.

Methods: COVID-19 cases and locations were curated for geostatistical analyses from March 2020 through June 2021, corresponding to the first, second, and third waves of infections. Daily cases were aggregated according to designated forward sortation area (FSA), and postal codes (PC) in municipal regions Hamilton, Kitchener/Waterloo, London, Ottawa, Toronto, and Windsor/Essex county. Hotspots were identified with area-to-area tests including Getis-Ord G_i^* , Global Moran's I spatial autocorrelation, and Local Moran's I asymmetric clustering and outlier analyses. Case counts were also interpolated across geographic regions by Empirical Bayesian Kriging, which localizes high concentrations of COVID-19 positive tests, independent of FSA or PC boundaries. The *Geostatistical Disease Epidemiology Toolbox*, which is freely-available software, automates the identification of these regions and produces digital maps for public health professionals to assist in pandemic management of contact tracing and distribution of other resources.

Results: This study provided indicators in real-time of likely, community-level disease transmission through innovative geospatial analyses of COVID-19 incidence data. Municipal and provincial results were validated by comparisons with known outbreaks at long-term care and other high density residences and on farms. PC-level analyses revealed hotspots at higher geospatial resolution than public reports of FSAs, and often sooner. Results of different tests and

Open Peer Review

Approval Status ✓ ? ?

	1	2	3
version 2 (revision) 21 Jul 2022	✓ view		
	↑		
version 1 23 Dec 2021	✓ view	? view	? view

1. **Lorant Foldvary**, Associate Professor, Budapest University of Technology and Economics, Budapest, Hungary
2. **Christine Tedijanto** , University of California, San Francisco, San Francisco, USA
3. **Ezra Gayawan** , Federal University of Technology, Akure, Nigeria

Any reports and responses or comments on the article can be found at the end of the article.

kriging were compared to determine consistency among hotspot assignments. Concurrent or consecutive hotspots in close proximity suggested potential community transmission of COVID-19 from cluster and outlier analysis of neighboring PCs and by kriging. Results were also stratified by population based-categories (sex, age, and presence/absence of comorbidities).

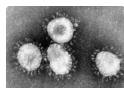
Conclusions: Earlier recognition of hotspots could reduce public health burdens of COVID-19 and expedite contact tracing.

Keywords

COVID-19, epidemiology, geostatistics, space-time analysis, kriging, infectious diseases



This article is included in the **Emerging Diseases and Outbreaks** gateway.



This article is included in the **Coronavirus** collection.

Corresponding author: Peter K. Rogan (progan@uwo.ca)

Author roles: **Mucaki EJ:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Shirley BC:** Methodology, Resources, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Rogan PK:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: PKR cofounded and BCS and EJM are employees of CytoGnomix Inc.

Grant information: Data acquisition and computing facilities used in this study was supported by ICES (Canadian Institute for Health Information), which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). Data and information compiled were also provided by MOH. This study was also supported by the Ontario Health Data Platform (OHDP), a Province of Ontario initiative to support Ontario's ongoing response to COVID-19 and its related impacts. The work described was funded and performed under a contribution agreement between the Innovation for Defence Excellence and Security (IDEaS) program at the Department of National Defence and CytoGnomix. The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources. No endorsement by ICES, OHDP, its partners, or the Province of Ontario is intended or should be inferred.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Mucaki EJ *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Mucaki EJ, Shirley BC and Rogan PK. **Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada [version 2; peer review: 1 approved, 2 approved with reservations]** F1000Research 2022, 10:1312 <https://doi.org/10.12688/f1000research.75891.2>

First published: 23 Dec 2021, 10:1312 <https://doi.org/10.12688/f1000research.75891.1>

REVISED Amendments from Version 1

We thank the reviewers for their insightful suggestions and have incorporated them where possible. The descriptions of the "Preparation of the Ontario COVID-19 case data" and "FSA and PC boundary files" now appear at the beginning of the Methods section. Selection of the Thin Plate Spline semivariogram for interpolation of case counts for Toronto is now justified in the Methods section (Empirical Bayesian Kriging analysis), including Chi-Square statistical testing comparing Power vs TPS with ground truth hotspot counts. Table 1 contains thumbnails of Postal Code maps for each municipality analyzed and of their locations in Ontario. The Zenodo archive now combines extended, underlying data and software. Table 2 has been moved to this archive (Extended data, Section 1). Results (Wave 1) and Figure S3 describes 2 additional examples detecting COVID-19 hotspots before it was reported in the media (line plots based on true and interpolated case counts, with G_i^* significance).

Any further responses from the reviewers can be found at the end of the article

Introduction

Many critical elements of disease epidemiology were unknown during the early stages of the COVID-19 pandemic. This included the case fatality rate, at what point during the natural history of the disease was it most transmissible (before, during or after the onset of symptoms), the frequencies of asymptomatic presentation of positive cases, and the duration of infectivity.¹ Quarantining, social distancing, and isolation of infected populations were quickly recognized as effective public health measures. Without implementation of these measures in Ontario, COVID-19 patients presenting with severe disease would have exceeded available hospital capacity.² Delayed testing and the sheer numbers of infected individuals often precluded comprehensive contact tracing.³ Indeed, COVID-19 testing was limited in many regions within Canada early in the pandemic, as was the availability of necessary resources (e.g., intensive care units, nursing and other clinical personnel, personal protective equipment).^{4,5} Limited resources must be adequately allocated to those locations where the rates of transmission of COVID-19 are high and where the highest numbers of infected individuals reside. This highlights the need for an effective, systematic way of tracking concentrations of this disease after community-level transmission. We address the persistence of geographic disease hotspots as a potential sentinel for efficient distribution of sparse resources.

The SARS-CoV-2 virus spreads by person-to-person contact, especially through respiration.⁶ Geographic epidemiology can be a key factor for the prevention of disease spread, as it can be used to monitor disease progression (or regression) and manage allocation of medical resources. Geostatistical methods have been utilized to map communicable diseases such as influenza-like diseases,⁷ dengue fever,⁸ cholera,⁹ legionellosis and shigellosis, among others.¹⁰ Various geostatistical analyses have been used in the surveillance of COVID-19 spread across the globe,¹¹⁻¹⁸ often with the geo-spatial software, ArcGIS (ESRI).¹⁹ Other available tools for geostatistical epidemiology include SaTScanTM,²⁰ and GeoDa,²¹ however these were not used in the present study because of their requirement to input disease background levels, which was not relevant for COVID-19. Available case data analyzed has been typically aggregated at the municipal^{11,16} or county level,^{13,14,22} however higher resolution spatial analyses has been limited as finer granularity for the distribution of COVID-19 cases has often been lacking.^{16,19} We performed geostatistical area-to-area analyses as well as interpolation of COVID-19 cases within six municipalities of the Canadian Province of Ontario (Hamilton, Kitchener/Waterloo, London, Ottawa, Toronto and Windsor/Essex region) distributed at the level of individual forward sortation areas (FSA) and postal codes (PCs; a smaller geographic segment within an FSA). Area-to-area analyses were performed on COVID-19 case data at the provincial level by FSA, a geographical unit similar to zip codes in the United States which can encompass hundreds of PCs. Each PC comprises an average of 19 households (Statistics Canada, 2006). Therefore, by evaluating COVID-19 case data at the PC level, we will be able to evaluate disease distributions at a high level of granularity.

The success of short-term COVID-19 geostatistical studies at the county level early in the pandemic²² suggested that hotspots and localized concentrations of infected individuals could be delineated more precisely with epidemiological data collected with higher granularity. The present analysis evaluates the three major waves of COVID-19 infection to date in Ontario, allowing for the observation of trends and commonalities between waves. We utilized ArcGIS to carry out multiple area-to-area geostatistical tests, including Getis-Ord G_i^* hotspot analysis (Figure 1),^{23,24} Cluster and Outlier analysis (Figures 2A and 2B; Anselin Local Moran's I^{25}), and Spatial Autocorrelation (Figures 2C and 2D; Global Moran's I^{26}). Hotspots were also inferred from spatially-interpolated counts by kriging and represented as topographic contour maps^{27,28} utilizing the ArcGIS geographic information system. Area-to-Area G_i^* analysis finds regions with a high disease burden (i.e., model for community-driven spread), but does not inform how these cases are distributed throughout the region. PC-level point-to-point analyses (e.g., kriging) can be employed to provide context of how cases cluster and identify localized hotspots, where cases are concentrated in a single PC and independent of the disease burden of the FSA it is within (Figure 3). The Getis-Ord G_i^* local statistic identifies regions with high disease burden with high, spatially clustered COVID-19 case counts (a local sum) relative to the rest of the Province,^{23,24} and is formulated as:

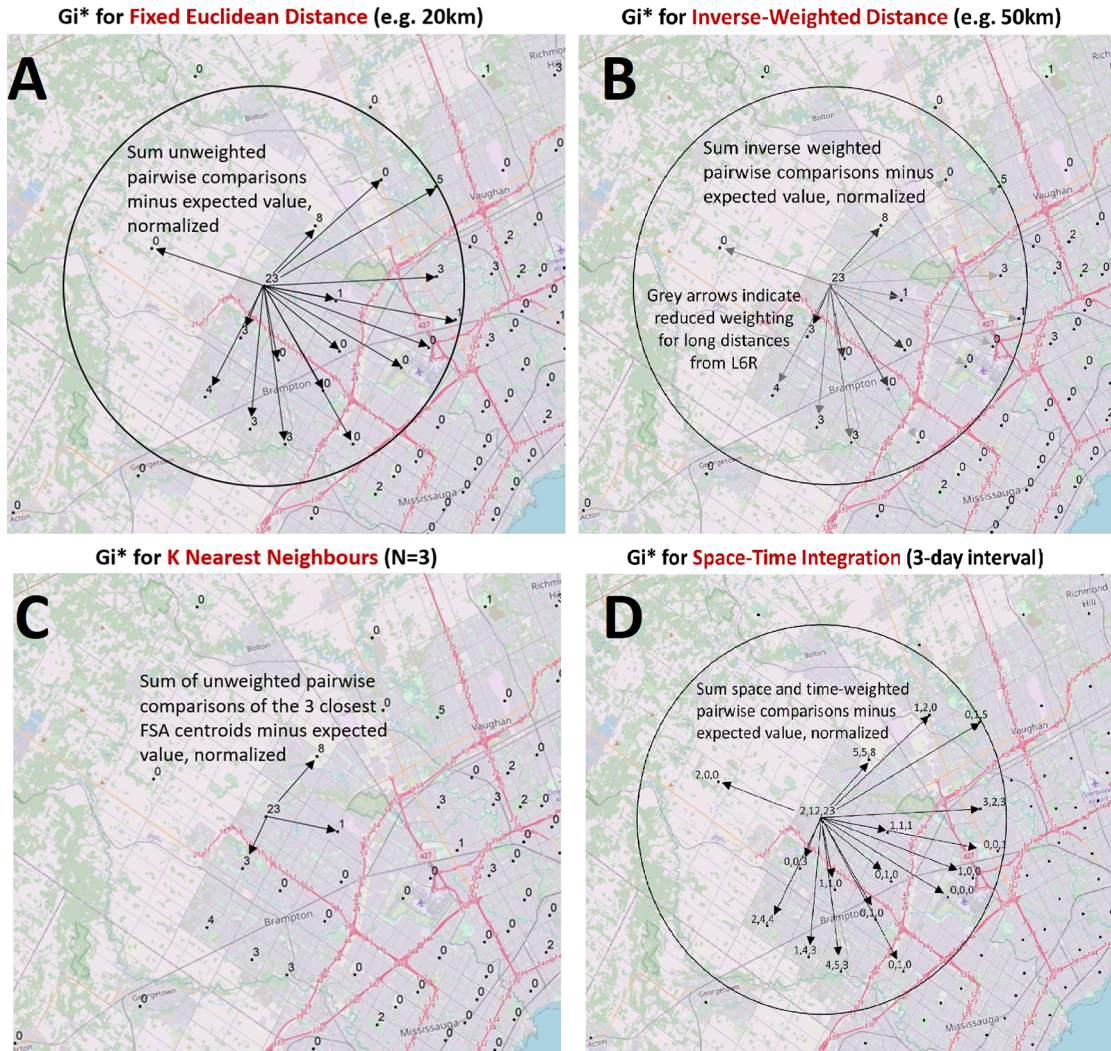


Figure 1. Modeling spatial relationships by area-to-area geostatistical analyses. Case counts aggregated at the centroids of FSAs or PCs over daily- or multi-day windows are compared to cases in neighboring FSAs or PCs and evaluated with geostatistical tests. These spatial relationships are defined by modeling spatial interactions between the FSA or PC being evaluated and the neighbors being tested. The panels describing these interactions indicate: (A) The fixed distance band weighs all neighbors within a specified distance equally (weights $[w_{ij}]$ are equal for features within circle; for all other features, $w_{ij} = 0$), (B) The inverse distance band is based on distance decay where neighbors further from the target feature is weighted less relative to closer neighbors ($w_{ij} = d_{ij}^{-\beta}$ where d_{ij} is the Euclidean distance between two points and β is the inverse distance power variable; $\beta = 1$ in this study), (C) K-Nearest Neighbors constructs a spatial relationship which assesses a feature with a specified number of its most proximate neighbors based on centroid distances (here, $K = 3$, where each K neighbor is weighted equally), (D) Space-Time integration defines feature relationships in terms of both a space (fixed distance) and a time window. Features that are proximate to one another in space and time are analyzed together (equal weights), as relationships are assessed relative to the location and time stamp of the target feature. Area-to-area geostatistical tests used include G_i^* analysis, Cluster and Outlier analysis, and Spatial Autocorrelation.

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{x} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{n \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij} \right)^2}{n-1}}} \quad (1)$$

where i and j are two separate features, w_{ij} is the spatial weight between these features (described in Figure 1), x_j is the value of the feature j (the feature being evaluated in this instance), n equals the total number features being considered, while \bar{x} and S equal:

Cluster and Outlier Analysis: Local Moran's I Spatial Autocorrelation: Global Moran's I



Figure 2. Geostatistical analysis of the distribution of COVID-19 cases. The Global and Local Moran's I statistics determine whether the numbers of cases in different PCs or FSAs in a region are distinct from one another or are related. Local Moran's I (also known as Cluster and Outlier analysis) computes a Moran's I value, or spatial autocorrelation, between counts within sets of FSAs or PCs. Global Moran's I evaluates the overall case clustering in a region. Panels shown indicate these relationships in Leamington FSA N8H: A) A positive local Moran's I index indicates that the regions neighboring N8H 3V5 has similar attributes (a high-case cluster; pink dot). B) A negative local Moran's I index demonstrates that cases the region surrounding the target are dissimilar (a high-case outlier in N8H 3V8; red dot). C) The Global Moran's I (Spatial Autocorrelation) index demonstrates whether cases are clustered, dispersed, or randomly distributed. In this example, cases were significantly clustered from Dec. 21 to 23, 2020 ($p = 0.03$; p -value signifies whether the null hypothesis [a random distribution] can be rejected). D) The case distribution was not significantly clustered on Dec. 2 to 4, 2020 ($p = 0.7$).

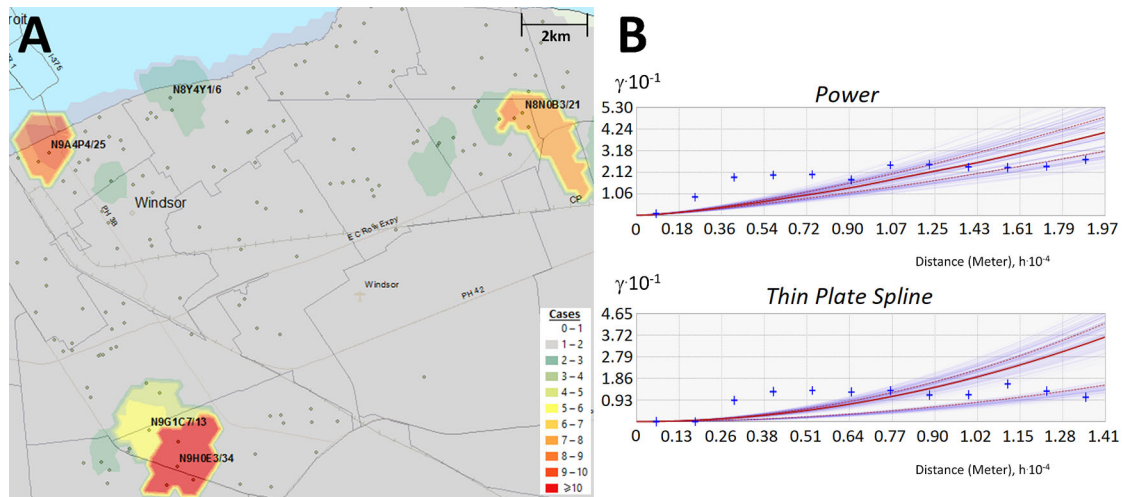


Figure 3. Interpolating the distribution of COVID-19 cases by Empirical Bayesian Kriging. Kriging estimates case counts by interpolating unsampled regions between counts at known locations. Panels indicate: (A) Topographic contours were generated by Empirical Bayesian kriging (EBK) when analyzing PC-level case data in Windsor (Dec. 18 to 20, 2020) and interpolate different case levels within the region; (B) EBK uses empirical semivariograms which describe the correlations between the distance separating a pair of locations with COVID-19-positive individuals, 'h', and their corresponding semivariance ' γ '. Larger values of γ indicate lower spatial autocorrelation between numbers of cases. A semivariogram model fits a curve through γ vs. distance that predicts γ at unsampled locations between known data points. The 'Power' semivariogram (B, top) is a parabolic mathematical model based on distance, while the Thin Plate Spline semivariogram (B, bottom) is a non-rigid model that passes exactly through bins of paired locations (grouped by distance and direction; blue crosses) while minimizing surface curvature. Red lines represent the median of the distribution of semivariogram models.

$$\bar{x} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{x})^2} \quad (3)$$

Space-time G_i^* analysis also uses this statistic to find hotspots that are both spatially and temporally clustered with other neighboring features (Figure 1D). Cluster and Outlier analysis identifies high-case regions that are geospatially clustered with other neighboring regions with a similar rate of cases (high-case clusters; Figure 2A), or those with neighboring regions with significantly lower COVID-19 case counts (high-case outliers; Figure 2B). The Global Moran's I index determines if the distribution of cases is clustered, dispersed, or random within a region.²⁶ The Moran's I index is positive for spatially clustered cases between postal codes, which is compared to the expected value (i.e., if the spatial relationship was random) to compute a z-score and significance level. Kriging is a geostatistical interpolation method which predicts values over a non-sampled region by analyzing a subset of locations with known values.²⁸ As a combination of geostatistical methods were applied to daily COVID-19 case data at the provincial [FSAs] and municipality [PCs] level for >450 days, efficient scripts were necessary to automate these analyses.

These geostatistical tools were used to identify and kriging COVID-19 hotspots, clusters, and outliers within Ontario at a high level of granularity. With the results of these analyses, we can ask questions regarding the distribution of hotspots relative to others in both time and space. For example, this study evaluates significant levels of COVID-19 over a series of consecutive days which we term 'streaks'. This allowed us to determine the true size and persistence of hotspots, compare and contrast geostatistical analyses of COVID-19 waves highlighting recurrent hotspots and clusters of PCs with high case counts. We also investigated the order in the progression of infection across a region in a statistically robust way. Empirical Bayesian Kriging (EBK) identified concentrations of COVID-19 that transcended the boundaries of FSAs and PCs. Ultimately, consistency between the results obtained by these different approaches reinforces the conclusions that might be drawn from them based on implementation as independent geostatistical tests.

Methods

Preparation of Ontario COVID-19 case data from ICES

Access to anonymized Ontario COVID-19 test results through the ICES Research and Analytic Environment (RAE) was obtained through Applied Health Research Questions (AHRQ) project #P0950.096 approved by the Ontario Ministry of Health and Long-Term Care (MOHLTC). ICES collects health care data for analysis and evaluation of the health system. Ontario-wide COVID-19 data tables maintained in ICES were pre-processed to ensure that only the chronologically first laboratory-confirmed viral RNA polymerase chain reaction (PCR) SARS-CoV-2 test of a specific individual was counted (while repeat infections occur, these are assumed to be significantly less frequent). PCs associated with each positive case were cross-referenced to their respective FSA (derived from the 2016 Canadian Census boundary file from Statistics Canada) to identify and filter out invalid entries. The PCs provided were sourced from the OLIS (Ontario Laboratories Information System) and the RPDB (Registered Persons) databases. The OLIS postal code was more likely to relate to a current address and was therefore prioritized. However, it is also the most error prone. Reasons for rejection include blank entries, invalid FSAs (e.g., out of province FSA), or intentional masking. The RPDB PC was used if the OLIS PC was rejected. Entries where both PCs were invalid were rare (<0.1% of all unique positive cases). Non-residents or those ineligible for the Ontario Health Insurance Plan (OHIP) are also not included. Cases were separately aggregated by FSAs and PCs, then converted into tables structured as rows (FSA or PC count) by columns (date) as input to ArcGIS Desktop. Cases were based on the date of sampling, rather than when test results were reported, to more accurately reflect the inception of infections.

As of June 2021, three separate and distinct waves of COVID-19 cases have been recognized in Ontario (and elsewhere). We delineated the inception and termination dates of each wave from the rates of change in infection rates (cases/day²) using the second derivative of daily case counts. Stochasticity in day-to-day variation due to testing and case reporting caused fluctuations in changes of daily counts, even when averaging cases over multiple consecutive days. Consecutive dates without fluctuations in counts (between -10 and +10 cases/day²) were rare. Therefore, dates were selected where the second derivative remained between -10 and +10 cases/day² immediately preceding rapid increases in cases that defined inception of a wave, or after infections ceased decreasing in counts, defining termination of a wave. The first wave (1) was found to span from March 26th to June 16th, 2020, second wave (2) began on September 25th, 2020 and continued until February 14th, 2021, and the third wave (3) covered March 17th to June 23rd, 2021.

Of the geostatistical methods tested, PC-level counts of infected individuals were the most geospatially resolved granular COVID-19 case data. Many PCs comprise of small geographic areas with low population densities. This made them more

prone to the effects of day-to-day stochasticity in patient ascertainment and laboratory reporting, especially when case counts were low. To address this, PC-level case data was also aggregated over multi-day sliding windows, effectively smoothing out this source of variation. Although 3 and 7-day windows were initially analyzed, longer windows were not as useful for determining the inception of hotspots, a goal of this analysis, and were eventually discounted. 3-day windows have been used in other geostatistical studies of COVID-19.^{13,14}

Access to, pre-processing and geostatistical analysis of COVID-19 daily cases and associated metadata required the RAE computer system at ICES. To ensure patient privacy, ICES stipulated that case counts could not be exported from the ICES system unless FSAs and/or PCs exhibiting between 1 and 5 daily COVID-19 cases were masked. We modified counts of FSAs meeting these criteria and assessed their impact on the geostatistical analysis. Gi* analysis of ground truth case data was compared with the analysis of two modified FSA-level case datasets (where all instances of between 1 and 5 case counts were either converted to a single count or randomized between 1 and 5). More than one third of the locations that were deemed significant hotspots by Gi* from actual case counts (from March 1st to September 24th, 2020) were not flagged when locations were masked and assigned a count of 1; greater than one half were missed when case counts were randomized (Extended data,²⁹ Section 1 - Table S1). Discordant hotspots identified from actual vs. masked case counts were reduced in FSAs with high-case counts (≥ 10) but were nevertheless unacceptably frequent (14% or 28% discordant, respectively, by masking low case counts as 1 or by randomization). We concluded that errors introduced by masking count values precluded analysis of masked data and required all work to be carried out within the RAE computing environment. All geostatistical analyses presented in this study were performed against unmasked, true case counts.

FSA and PC boundary files

Shapefiles which store geospatial data [points, lines or polygons] with related attribute information and depict the 2016 census geographic boundaries of FSAs across Canada were obtained from [Statistics Canada](#). PC boundary shapefiles created by DMTI (“DMTI_2019_CMPCS_LocalDeliveryUnitsRegion.shp”) were accessed through the [Scholar's Geoportal](#) at the University of Western Ontario. These files were used to compute latitude/longitude coordinates of the centroid for each FSA and PC (with the ArcGIS “Calculate Geometry” function), to validate PCs provided for each individual in the COVID-19 test dataset, and to convert spatial interpolation maps (such as those generated by kriging) into machine-readable text files (using ArcGIS “Intersection” function between the spatial map and with each boundary file). The ArcGIS “Split by Attribute” function was used to limit PC boundary files to the province of Ontario. This was necessary, both because data from other provinces were not available, and because of system limitations affecting processing and memory requirements for large geographic regions.

Minor discrepancies between the FSA and PC boundary files were identified which had little or no impact on our geostatistical analysis. Nearly 9% of all PCs in Ontario overlapped multiple FSAs. Most of these PCs were present in multiple locations and occurred in rural areas. Approximately 2.7% of PCs intersected with an incorrect FSA, which was likely the result of asynchronous creation of the corresponding shapefiles. These locations rarely contained many COVID-19 cases as only 11 PCs had >5 cases over any 3-day span in 2020, including: N9H 0E3 (Windsor, ON) which overlaps with the FSA N9G; N8N 0B3 (Tecumseh, ON) which overlaps the bordering FSA N8R, and M3H 5V9 (Toronto, ON) which intersected with neighboring FSA M3J. These discrepancies might impact the interpolation of cases in FSAs based on kriging analysis. Since the discrepant PCs were typically found on the border of FSAs, these conflicts had minimal effect on kriging accuracy.

Geostatistical analysis

COVID-19 cases from March 26th, 2020 to June 25th, 2021 at the FSA (Ontario-wide) and PC-level (for 6 populated areas: Hamilton, Kitchener/Waterloo, London Ottawa, Toronto, and Windsor, including Essex county) were evaluated using ArcGIS Desktop 10.7 [ESRI] with the *Geostatistical Analyst* extension. Daily cases compiled by FSA were evaluated with Cluster and Outlier analysis, Gi* ‘hotspot’ analysis, and with space-time integrated Gi* analysis. Cluster and Outlier analysis computes a Local Moran’s I index score for each feature. A positive score indicates that a feature is clustered with other similar high-case count regions, while a negative index indicates that a feature is an outlier. The z-score and p-value indicates the statistical significance of the index value. A false discovery rate (FDR) correction was applied to this and Gi* analysis, which changes the threshold of significance of p-value ≤ 0.05 to a reduced value within the 95% confidence interval [C.I.], thereby accounting for multiple testing and spatial dependency. Z-scores are also determined in Gi* analysis, with higher positive scores indicating greater clustering of cases in a region. Space-time Gi* analysis was based on a single day interval but incorporated temporal trends based on the days preceding and subsequent to each date. We set a minimum threshold of 10 confirmed cases for an FSA to be considered significant by these three approaches, as this threshold was found to provide reliable and conservative detection of known hotspots using the Gi* test. Over two thirds of FSAs with ≥ 10 cases between March 1st and Sept. 24th, 2020 were found to be significant by Gi* [99% C.I.] using both a fixed and inverse distance metric (Extended data,²⁹ Section 1 - Figure S1). Spatial relationships

were evaluated to determine which distance metric best reflected the relationship between the test location and its surrounding FSAs by fixed distance (all FSAs within a given threshold distance [10km] are equally weighted with the test FSA), inverse distance (FSAs closer to the test FSA are weighted more highly), and by K Nearest Neighbors (KNN; comparing counts in the test FSA and its closest K neighbors).

PC-level case data (binned over 3-day sliding windows) for each municipality were evaluated by Spatial Autocorrelation (Global Moran's I), Cluster and Outlier analysis, and EBK. Spatial Autocorrelation used FDR correction to account for multiple testing; it was not applied for PC-level Cluster and Outlier analysis due a limitation of the method in high-case count regions where all locations were deemed non-significant. The KNN distance metric was used to define the spatial relationship between features (with $K = 3$) for all PCs, including those with zero cases. PCs with ≥ 6 cases were reported if deemed significant by Cluster and Outlier analysis. This is consistent with case minimums used in other COVID-19 geostatistical studies.^{13,14} Kriging generates a topographic contour map of the number of COVID-19 cases interpolated over the evaluated region, as indicated below.

COVID-19 testing data compiled by ICES contained metadata on tested individuals, such as gender, age, and relevant predicate health conditions. To identify geospatially significant cases in these subgroups, case data was stratified by gender (Male: 241,393; Female: 247,152; those without gender information were ignored), by age (≥ 65 years old: 63,746; < 65 years: 424,854), and by presence or absence of at least one chronic health condition (187,242; includes chronic asthma, congestive heart failure, chronic obstructive pulmonary disease, dementia, hypertension and/or diabetes). 301,358 patients were without one of these conditions. Individuals with these conditions are grouped as an "at-risk population", consistent with provincial terminology. These case subsets were then evaluated with G_i^* (FSA-level), Spatial Autocorrelation, Cluster and Outlier analysis and kriging (PC-level).

Results from PC-level Cluster and Outlier analysis were further analyzed with a program script to identify significantly, highly clustered PCs over consecutive periods (i.e., termed high-case cluster "streaks"; *Local_Morans_Analysis.Recurrent_Clustered_PC_Identifier.pl*; see Data Availability section²⁹). We also identified which of these PC streaks occurred in close proximity (*Local_Morans_Analysis.Clustered_Streak_Pairing_Program.pl*) based on the neighbors of each postal code from PC centroid data (neighbors identified by *Ontario_City_Closest_Postal_Code_Identification.pl*; see *FSA and PC boundary files* subsection). Directional network graphs which visualize the interconnectivity of these paired high-case cluster streaks were created in MATLAB using built-in *digraph* and *plot* functions, where each edge (the connection between two PCs) represents a single pair of PCs with nearly concomitant or overlapping streaks.

Software automation was required to computationally evaluate daily and windowed case data over 16 months at the geographic resolution of FSAs and PCs across multiple cities with different geostatistical tests. Using the ArcPy package (ESRI), which enables Python programs to access ArcGIS's geoprocessing tools and extensions, program scripts were developed to perform geostatistical analyses across Ontario. These scripts set conditional variables which define the spatial relationship used (fixed, inverse or KNN), the threshold distance (or for KNN, the number of neighbors of the target feature), the distance method (how distances were calculated; set to 'Euclidean'), row standardization (set to 'none'), and whether FDR correction was used (applied to FSA-level G_i^* and Cluster and Outlier analysis, and PC-level Spatial Autocorrelation analysis). ArcPy scripts read in case data (in tabular format), iterated sequentially through each date skipping those without positive cases, applied the associated geostatistical tool, and converted the output into text files. Tabular output was subsequently filtered by removing PCs without cases to avoid producing outputs of prohibitive size.

These processes were then integrated and streamlined in a software package named the *Geostatistical Epidemiology Toolbox* (see Data Availability section²⁹), which is accessed through the ArcMap graphical user interface. This resource provides the same geostatistical operations and map imaging developed for this project in a simplified, user-friendly environment. This toolbox bundles manual pre-processing, analysis, and post-analytical steps into a set of linked, uninterrupted workflows. While the toolbox was designed to evaluate COVID-19 case data in Ontario, the software can be easily modified to accommodate other Canadian provinces and, with additional effort, other countries (see *Geostatistical Epidemiology Toolbox user manual*²⁹). To enable kriging functionality, the *Advanced Geostatistical Analyst* (ESRI) extension and the PC and FSA boundary shapefiles are required. These shapefiles are also required for PC-level Local Moran's I and Global Moran's I analyses. This package is not bundled with any of the program scripts developed to evaluate geostatistical output (besides data imaging), which are also provided in the accompanying Zenodo archive.²⁹

Empirical Bayesian Kriging analysis

Kriging estimates case counts in unsampled regions by interpolation between counts at known locations. These values are spatially correlated with degree of separation from centroids of PCs or FSA with known cases. A topographic contour map is generated which illustrates the interpolated case levels within the analyzed region. The spatial distribution of the

disease outbreak was determined for Hamilton, Kitchener/Waterloo, London, Ottawa, Toronto, and Windsor/Essex county by kriging analysis of PC-level COVID-19 case data over consecutive, overlapping 3-day windows. Kriging was also performed for other municipalities bordering Toronto (Ajax, Brampton, Markham, Mississauga, Pickering, Richmond Hill and Vaughan; see results in Extended data,²⁹ Section 3). Only PCs with ≥ 1 case were included in this analysis, as locations without confirmed cases greatly outnumbered PCs with cases ($\geq 95\%$ of PCs on any given date evaluated) leading to the severe depression of kriging signals. Several types of kriging (Ordinary, Universal, Simple and EBK) were evaluated for accurate detection of high case clusters in municipalities with high and moderate population densities (Toronto and London, respectively). Contour maps generated by EBK best matched regions of known cases used for analysis. Ordinary and Universal kriging yielded similar results for locations with high case counts (≥ 50) but lacked the sensitivity of EBK in regions with moderate cases. Simple kriging failed to define most contours correctly, regardless of case count. Therefore, EBK was chosen for all kriging analyses presented in this study.

Semivariograms describe how the distances separating COVID-19 positive individuals and the semivariance of counts at these locations are correlated. The best model that fits this relationship using empirically defined distance data was used to predict COVID-19 count levels at unsampled locations between the centroids of FSAs or PCs. Semivariogram models assessed included Power, Linear, Thin Plate Spline (TPS), Whittle, Exponential, and K-Bessel. Kriging parameters for these models were considered, such as sector type, search radius, neighborhood type and the number of neighbors considered for each location. Through empirical testing, the Power semivariogram model (based on a parabolic relationship to distance between locations) was chosen based on interpolated values that best represented actual case counts in low/moderately dense populations (Extended data,²⁹ Section 1 - Figure S2A). Kriging accuracy improved by inferring interpolated values from 3 neighboring locations with known counts. However, the Power semivariogram failed to detect some hotspots in Toronto, which exhibited the highest case counts and overall population density. The TPS semivariogram model, a non-rigid representation that passes exactly through points while minimizing surface curvature, performed best in this city (Extended data,²⁹ Section 1 - Figure S2B). We noticed that the separation between Power-based kriging contours was generally larger, in contrast with TPS-based contours which were tightly packed, more bifurcated and frequent. TPS, however, tended to generate elevated estimated counts that did not correlate with known hotspots. Using 5 neighboring locations for interpolation minimized these artifacts without affecting the contour map. However, with 10 neighbors, interpolation accuracy was affected according to a chi-square test (e.g., $p < 0.05$ for the 103 days in Wave 2). The Power semivariogram did not significantly differ from the actual counts ($p\text{-value} > 0.05$). However, interpolation with the TPS model in Toronto was consistently more similar to the actual case counts than the Power model, based on Chi-square test for all hotspots within Waves 1 and 2 (resulting in higher p-values for 85% and 78% over all dates within each wave). We performed kriging in Hamilton, Kitchener/Waterloo, London, Ottawa and Windsor/Essex with the Power model (3 neighbors), and in Toronto with the TPS model (5 neighbors). All other EBK parameters used were defaulted to those provided by the ArcGIS Advanced Geostatistics toolbox (Sector Type: 1 sector, Neighborhood Type: Standard, Search Radius: 1).

We have previously demonstrated that circumscribing the boundaries of analyzed regions by specifying these locations with zero counts improves kriging accuracy, especially at locations proximate to boundaries.³⁰ This approach was used in the present study of interpolated COVID-19 cases for each municipality in Ontario. Municipal borders were derived by combining all PCs associated with the region of interest. The optimal densities and distances of the zero-count buffer circumscribing municipalities by programmatically generating boundaries. Distances between the boundary to the city border were varied from 0.1, 0.5, 2, and 5 km, and densities between locations of 0.05, 0.25, 0.5 and 1 km. Kriging analysis was iterated for each of these defined zero-count borders, and the kriging layers generated were compared to the original kriging results. The location of each kriging contour properly coincided with regions of high case counts, regardless of the shape and density of the zero-count boundary locations. Raising the boundary density consistently increased the COVID-19 case value interpolated of a high-case region while simultaneously reducing the total area of the kriging contours around the region, regardless of semivariogram model used. This effect was less prominent in regions with higher density of cases (i.e., Toronto), and in regions further from the city boundaries (as the kriging contours exhibited consistent shape and magnitude > 6 km away from the city boundary). Considering these observations, we chose to use a moderate value for each parameter (0.5 km density and 1 km distance from the city boundary). The kriging tool in the *Geostatistical Epidemiology Toolbox*, however, automatically generates these boundaries 1 km from the selected border with 250 m of distance between each location.

Kriging analysis generates a series of shapefiles containing contour maps that define the interpolated level of COVID-19 cases of a region. An ArcPy script intersected map contours derived from kriging and either FSA or PC boundary files, which associated each kriging-derived case count with these locations. The results of these intersections were displayed as distinct contour distributions (or kriging 'symbolologies'): one for instances with a wider range of inferred case count levels (0-1, 1-5, 5-10, 10-15 ... 35-40, 40-50, > 50 cases), and another for low densities of inferred cases, which was

discriminated by single counts (0-1, 1-2 ... 9-10, >10 cases). In compliance with ICES’s reporting criteria, we defined any interpolated value >5 cases (either $\geq 5-6$ or $\geq 5-10$ contour, depending on the symbology type used) as a region of interest, and display PC or FSA designations for only these locations. The intersected area (m²) between contours and the corresponding FSA or PC was determined to compute the size of localized hotspots (e.g., the total area of the FSA or PC considered a region of interest by kriging).

Results

Changes in the distribution of COVID-19 infections over time were evaluated through geostatistical analysis of true (unmasked) case counts within both FSAs and PCs. PC-level analysis can map COVID-19 spread at a finer resolution, which more precisely identifies hotspots concentrated within individual postal codes (e.g., a localized hotspot). Area-to-area geostatistical analysis of cases in FSAs is better suited when COVID-19 cases are distributed across the entire FSA (e.g., an area with high disease burden by community-driven spread; Figure 4), rather than when cases are concentrated in a single location. This method does not, however, provide any information regarding the overall distribution of cases within the FSA. Spatial Autocorrelation analysis (Global Moran’s I) was therefore performed to identify FSAs where COVID-19 cases are clustered at the PC-level (potential evidence of disease transmission). These geostatistical methods combine analyses of COVID-19 cases at multiple levels of resolution, facilitating interpretation of the case distribution throughout the province.

Significant FSA hotspots were assessed by Gi* through comparison of cases within KNNs, a fixed distance, or by weighted inverse distance. During the first wave of COVID-19, when testing was not fully established across the province, FSAs with the minimum number of cases considered significant hotspots (≥ 10) were less common, and fewer

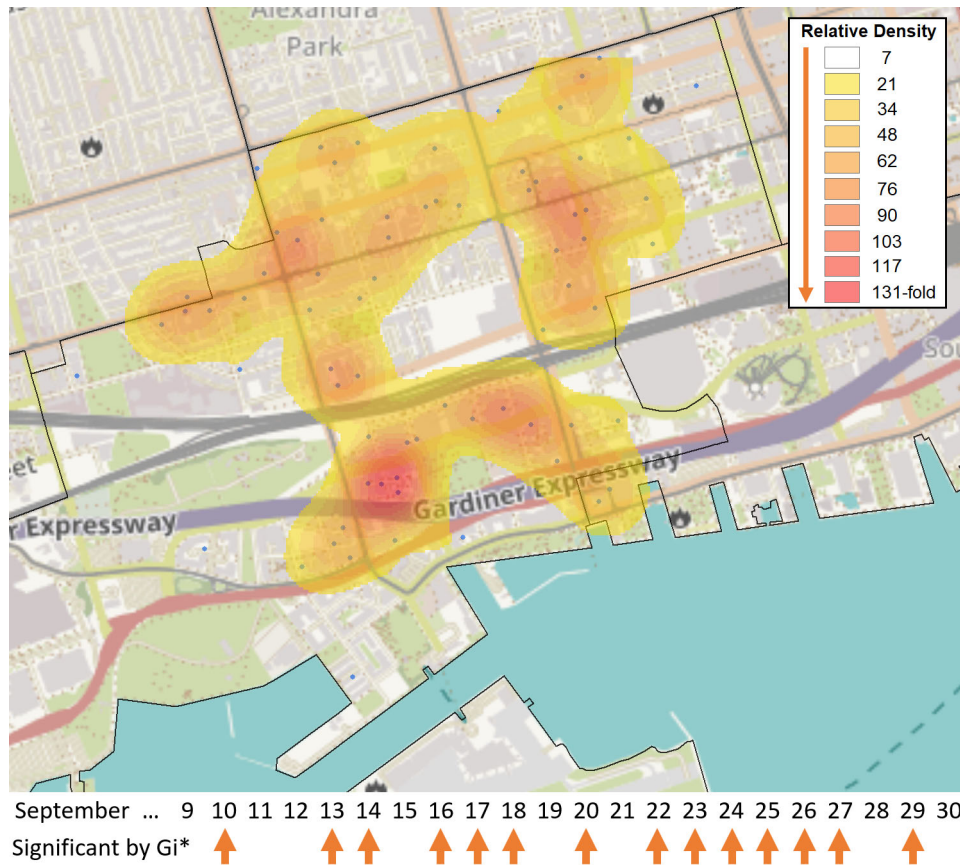


Figure 4. Broad distribution of COVID-19 cases across FSA M5V in Toronto. The FSA M5V in the entertainment district of Toronto was identified as a hotspot relative to the rest of Ontario by the Gi* test over 14 days in September 2020 (case counts ≥ 6 , with five days exhibiting at least 10 cases [Sept. 10th, 13th, 20th, 22nd, 23rd]). These cases were distributed across 71 separate PCs in M5V. A heat map generated with the *Kernel Density* function within the Spatial Analyst Toolbox of ArcGIS shows the relative density of postal codes with cases (blue dots), weighted by the number of days in which each individual postal code had ≥ 1 case (N = 118 cases total). Kriging analysis did not identify localized hotspots during this period, as none of the PCs exceeded 3 cases.

FSAs met these criteria (78 to 94%; Extended data,²⁹ Section 1 - Table S2). Later, fixed distance comparisons identified more FSAs with significantly higher disease burden than other metrics (53% of all FSAs in wave 3 versus 24% for KNN3 and versus 13% with inversely weighted FSAs; Extended data,²⁹ Section 1 - Table S2). We prioritized the standard Gi* analysis which utilized KNN (where K = 3) in this study, as it is moderately stringent (compared to results using the fixed and inverse distance metrics) and is unaffected by the shape of FSAs (which can be quite variable when comparing rural and urban areas).

Space-time Gi* analysis consistently called FSAs with ≥ 10 cases as significant in waves 2 and 3 (>91% and 95%, respectively; Extended data,²⁹ Section 1 - Table S2). This approach identified nearly twice as many FSAs as significant versus to standard Gi* comparing neighbors up to the same fixed distance (Extended data,²⁹ Section 1 - Table S2). FSAs deemed significant hotspots by space-time Gi* were tested against many more neighbors (<10km) relative to FSAs that were not hotspots (averaging 59 and 17 neighbors, respectively). This revealed a bias in significant FSA hotspots towards smaller, densely organized FSAs (which were typically urban), while those that did not achieve significance were often rural, covering larger areas and exhibiting lower COVID-19 counts. Additionally, 98% of the FSAs flagged were also significant hotspots on the days prior and after the central date in each window. This was uncommon in FSAs that were not determined to be significant hotspots (6%). As the vast majority of regions were flagged by this approach (>90%), we did not gain any insight from this space-time analysis. Instead, we evaluated temporal trends in the distribution of COVID-19 cases and identified potential transmission using Cluster and Outlier analysis of PCs that persisted across a consecutive series of dates (referred to as high-case cluster 'streaks'; described later in the Results section).

Comparison of kriging and Cluster and Outlier analysis of PC-level case data revealed that the same statistically significant PCs were often identified by both approaches on the same dates (March 26th to December 28th, 2020; ≥ 6 cases). A PC was considered significant by Cluster and Outlier analysis if the location was deemed a high-case cluster or outlier with an FDR-correction with 95% confidence, or a region of interest by kriging if ≥ 5 cases were interpolated within it. This comparison finds that both approaches are consistent in municipalities with lower overall case counts such as Hamilton (65.2% concordance; N = 66 PCs found significant by cluster/outlier analysis), Kitchener/Waterloo (85.1% concordance; N=47), London (69.0% concordance; N = 29), and Windsor/Essex (85.0% concordance; N = 160) (Extended data,²⁹ Section 1 - Table S3). The two methods showed <50% correlation to each other when evaluating cases in Ottawa (45.0% concordance; N = 211) and Toronto (49.7% concordance; N = 529). Nearly all discordant PCs were interpolated to have between 1-5 cases (i.e., marginally discordant). Kriging analysis commonly identified high-case outliers more often than high-case clusters. In Hamilton, 70% of high-case outliers but only 29% of high-case clusters were also found to be regions of interest by kriging (N = 59 and 7, respectively). Similar patterns were observed for all regions evaluated except Windsor/Essex, where overlap between kriging and high-case clusters and outliers were comparable (83% and 96% [N = 138 and 22], respectively; Extended data,²⁹ Section 1 - Table S3).

As incidence rates of COVID-19 have fluctuated since its initial appearance in Canada, the following geostatistical analyses have been stratified by periods with significant infectious burden, which provides a common frame of reference. We organize the analyses according to the three periods (or waves) with maximal occurrence of COVID-19 infections, where wave 1 extended from March 26th to June 16th, 2020, wave 2 from September 25th, 2020 to February 14th, 2021, and wave 3 from March 17th to June 23rd, 2021. The intervening periods between waves 1 and 2 and waves 2 and 3 are designated as inter-waves 1-2 and 2-3, respectively.

Wave 1

Kriging analysis of COVID-19 cases in this wave identified localized hotspots in Hamilton, Kitchener/Waterloo, London, Ottawa, Toronto, and Windsor/Essex (all localized hotspots are identified in Extended data,²⁹ Section 2). High disease burden FSAs (those flagged by Gi* analysis) were often found to include localized hotspots (>50% of flagged FSAs in cities evaluated; Extended data,²⁹ Section 1 - Table S4). Localized hotspots (regions within contours exceeding interpolated count >5 cases) were most frequently identified within high disease burden FSAs in Ottawa during the first COVID-19 wave (84%; Extended data,²⁹ Section 1 - Table S4). However, kriging also identified 34 additional localized hotspots within FSAs on dates within this period, suggesting that COVID-19 events of interest occur that are missed by Gi* analysis. Due to its high density, kriging analysis of cases in Toronto using the Power semivariogram (used in analysis of all other cities) was found to have poor overlap with Gi* analysis (only 31.6% of Gi*-significant FSAs contained a localized hotspot). The overlap between kriging and Gi* test results was improved using the TPS semivariogram (increased to 60.9%), and this was therefore used for all subsequent kriging analyses of Toronto. PC clustering was uncommon during this wave (<15% of FSAs flagged by Gi* and/or kriging were clustered). The exception was London where 27% of localized hotspots occurred in clustered FSAs [N = 11], and the majority of cases were often localized to a single PC in other cities that were analyzed (Extended data,²⁹ Section 1 - Table S4).

COVID-19 outbreaks frequently affected residents of long-term care residences during this wave. Geostatistical analyses confirmed the ability to identify these outbreaks. For example, an outbreak in a nursing home in Toronto in early April 2020 (M8V; Lakeshore Lodge) was identified by kriging, while the FSA in which it occurred was also significant by Gi* analysis (Figure 5A). A different outbreak in Toronto in April 2020 (M2K; Extencicare Bayview nursing home) was also detected by kriging analysis, but the FSA in which it was found was not a significant hotspot by Gi* (Figure 5B); it was significant only using a fixed distance metric of ≥ 20 km). Similarly, the FSA K1J (Ottawa, ON) was only significant by Gi* using an inverse distance weighted test, despite a mid-April 2020 outbreak in the Rothwell Heights assisted living community which was identified by kriging (Figure 5C). Localized hotspots within the FSA M9P were observed as early as March 31st, 2020 however the FSA was not flagged by Gi* analysis until April 6th, 2020 (Extended data,²⁹ Section 1 - Figure S3A). The localized hotspot covers the Village of Humber Heights Retirement Home in Toronto, which was reported to have a COVID-19 outbreak on April 12th, 13 days after the localized hotspot was first detected by kriging.³¹ In each of these instances, the interpolation of cases by kriging were similar to actual case counts in these FSAs. This was

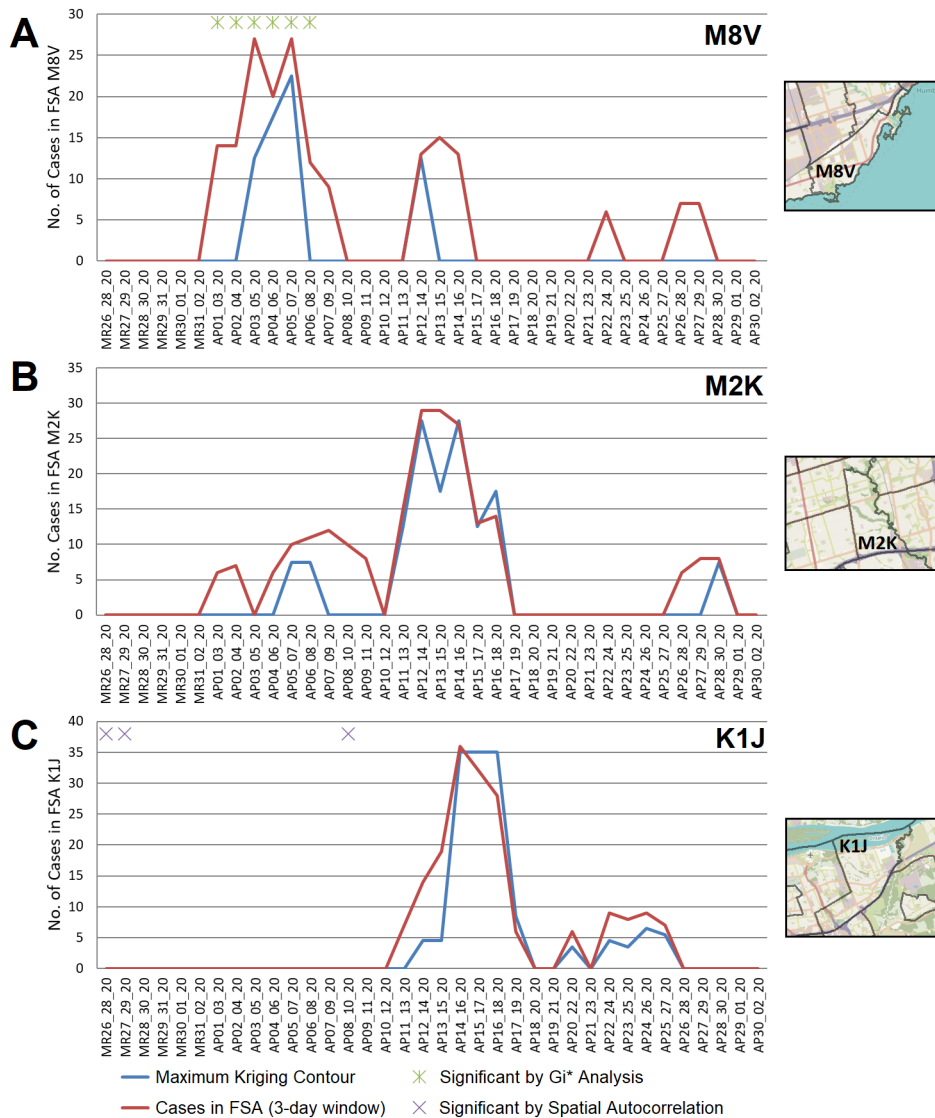


Figure 5. Geostatistical analysis of COVID-19 outbreaks during wave 1. FSAs exhibiting high-disease burden by Gi* analysis were identified with localized hotspots by kriging or as having significant case clusters by Spatial Autocorrelation. Panels indicate cases in all PCs within FSAs (A) M8V (B) M2K, and (C) K1J, which were amalgamated over 3-day sliding windows and masked (red line indicates actual case counts; privacy consideration require dates with fewer than 5 cases to be masked as zero). Kriging interpolated the number of cases over these FSAs (blue line), which closely approximates true case counts when counts are high. Green asterisks indicate dates where at least one day within a 3-day window was significant by Gi*; purple crosses indicate significantly clustered cases within PCs.

consistent with the concentration of cases within a single PC (a frequent occurrence in the first COVID-19 wave; Extended data,²⁹ Section 1 - Table S4). Hotspots where the case total exceeds the interpolated value is consistent with cases distributed among multiple PCs across an FSA. For example, localized hotspots in FSA M9V were identified sporadically, and the interpolated value does not always match total cases (Extended data,²⁹ Section 1 - Figure S3B). The localized hot spots cover the PC M9V 5B5 starting from April 4th. Outbreaks in the long-term care home within this PC (Humber Valley Terrance long-term care home in Toronto) were reported on April 16th, 2020.³² This FSA was not flagged by Gi* analysis until April 19th (Extended data,²⁹ Section 1 - Figure S3B).

Localized hotspots identified by kriging comprise a narrow range of PCs in each FSA. The area of interpolated contours that overlapped FSAs with significant disease burdens (by Gi* analysis) was compared to the total area of all corresponding FSAs. COVID-19 cases were highly localized, as the highest value contour (where kriging interpolates the greatest number of cases) constituted only a small portion of the total FSA area (Extended data,²⁹ Section 1 - Figure S4). While hotspots found by kriging often coincide temporally with Gi* results (Extended data,²⁹ Section 1 - Table S4), the interpolated hotspots contain fewer cases relative to the entire FSA, since the maxima of kriging contours cover only a portion of the overlapping FSA. The overall area indicated as a region of interest by kriging overlaps a greater fraction of the FSA area, but still covers a fraction of its area (<10% in the majority of cases). Compared to the other COVID-19 waves, however, kriging analysis of cases during this wave tended to span a larger area of each FSA (Extended data,²⁹ Section 1 - Figure S4A). This appears to be consistent with higher transmissibility (which, we speculate, may be related to lower levels of immunity during this period).

COVID-19 case counts were stratified into three categories: biological sex (males vs. females), age (exceeding vs. below 65 years old), and overall health (presence or absence of at-risk, pre-existing comorbidities). Stratified PC-level case counts were evaluated by Cluster and Outlier analysis across all waves (Table 1). In the first wave, the frequency of high-case outliers (a high-case PC with few cases amongst its immediate neighbors) is biased towards older, at-risk populations (Table 1A). Across all municipalities, PCs were flagged nearly 5-fold more often in the older population (305 and 62 PCs were identified as a high-case outlier on different dates in this wave, respectively). A similar pattern was observed in at-risk vs healthy populations (321 and 31 high-case outlier PCs, respectively), which is not surprising due to the considerable overlap between the elevated age and groups with at-risk comorbid diagnoses. High-case clusters were identified less often, but also showed this pattern, which is likely related to the high frequency of outbreaks in long-term care homes during this period (Table 1B). Windsor/Essex was an exception to this trend, since the frequencies of PCs designated as high-case outliers were comparable among both age and comorbidity groups (20 vs 17 for older vs younger populations, 19 and 17 for at-risk vs. healthy populations; Table 1A). The frequency of high-case clusters/outliers identified by sex-stratification varied by the region analyzed (e.g., more high-case outliers were identified in the female population in Toronto, in contrast within the male population in Windsor/Essex). This stratified analysis itself does not provide insight into the mode of COVID-19 transmission, that is, the development of a hot spot consisting of a particular subgroup does not indicate that viral transmission occurred exclusively between members of the particular subgroup.

Interwave 1-2

The first COVID-19 wave ended approximately two months after the imposition of public health prohibitions on mobility, limits on assembly, and social distancing (i.e., a lockdown) in Ontario. June 16th, 2020 defined the start of interwave 1-2, a period of much lower COVID-19 incidence than the preceding 2 ½ months. Fewer FSAs exhibited high disease burden by Gi* analysis (39.7% of those called in wave 1, despite the duration of interwave 1-2 being longer; Extended data,²⁹ Section 1 - Table S2). Despite the overall decrease in cases province-wide, the frequency of FSAs in the Windsor/Essex region that were designated hotspots by both Gi* and kriging analysis increased relative to the preceding wave (Extended data,²⁹ Section 1 - Table S4). This was due to multiple outbreaks in the Essex County FSAs in Leamington [N8H] and Kingsville [N9Y] (Extended data,²⁹ Section 1 - Figure S5). These outbreaks were publicly documented in guest farm workers, which is reflected by the Cluster and Outlier analysis of stratified case data (a strong bias of high-case clusters and outliers in healthy males <65 years in Windsor/Essex; Extended data,²⁹ Section 1 - Table S4). These same locations were correlated with interpolated kriging contour maxima and often consisted of isolated single PCs (Table 1). However, there were instances where the interpolated count values exceeded the actual case count (i.e., N8H, June 22-26th, 2020; Extended data,²⁹ Section 1 - Figure S5B) due to concurrent high case counts from neighboring FSAs (i.e., N0P 2J0; Extended data,²⁹ Section 1 - Figure S5C) bleeding into these contours from a single PC.

Wave 2

FSAs with high disease burden were more frequent within the second wave (Extended data,²⁹ Section 1 - Table S2). Even when accounting for its longer duration, there was a >10-fold increase in FSAs with high disease burden based on the Gi* test. The distribution of cases was more widespread across each FSA, as COVID-19 testing became more reliable and pervasive. Localized hotspots were detected less frequently within FSAs with high disease burden, i.e., significant FSAs by Gi* analysis overlapped less often with kriging contours of interest (>5 interpolated cases) in all municipalities except

Table 1. Number of postal codes with high case clusters and outliers by location and wave.


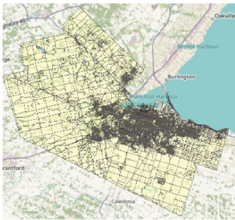

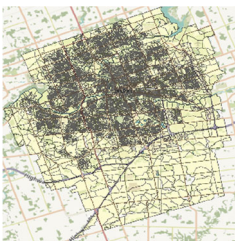
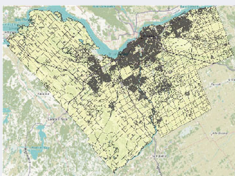
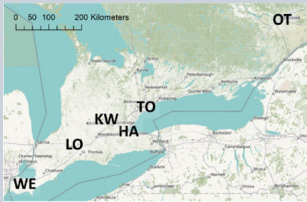


A. Distribution of high-case outlier postal codes in different municipalities									
City ¹	Wave ⁴	Estimated total cases ⁵	Number of high-case outlier PCs in stratified populations ⁶						
			All cases	Males	Females	≥65 years	<65 years	At-risk	Healthy
 <p>[No. of FSA/ PC/Area²/Pop.³]</p>									
(HA) Hamilton 	1	787	8	3	6	7	1	8	1
	1-2	410	0	0	0	0	0	0	0
	2	8788	100	20	43	31	45	41	24
	2-3	1464	29	6	4	0	25	2	8
	3	9755	150	36	43	0	131	9	84
[30/ 14,774/1,138/536,915]									
(KW) Kitchener/Waterloo 	1	511	25	14	11	17	4	24	0
	1-2	277	0	0	0	0	0	0	0
	2	4287	31	3	10	12	13	17	6
	2-3	560	11	5	0	0	11	0	8
	3	2733	30	3	0	0	26	0	15
[18/8,445/200.9/338,205]									
(LO) London 	1	363	11	3	3	6	5	6	3
	1-2	213	6	0	2	0	6	0	0
	2	3609	31	3	5	4	21	5	7
	2-3	347	2	0	0	0	0	2	0
	3	4363	28	6	1	0	26	0	10
[20/10,063/420.6/383,825]									
(OT) Ottawa 	1	1125	50	10	26	42	2	45	0
	1-2	1327	28	2	10	8	16	11	12
	2	6790	120	25	21	28	84	29	54
	2-3	1178	12	0	2	2	7	2	7
	3	8716	50	8	13	0	48	0	30
[47/ 21,963/2,790/934,031]									

Table 1. *Continued*

A. Distribution of high-case outlier postal codes in different municipalities									
City ¹	Wave ⁴	Estimated total cases ⁵	Number of high-case outlier PCs in stratified populations ⁶						
			All cases	Males	Females	≥65 years	<65 years	At-risk	Healthy
 [No. of FSA/ PC/Area ² /Pop. ³]									
(TO) Toronto  [101/47,450/630.2/2,732,000]	1	7144	261	81	162	213	33	219	10
	1-2	3148	25	3	10	6	13	10	4
	2	54,460	596	78	178	191	315	216	127
	2-3	7309	89	17	13	1	76	16	40
	3	52,037	428	45	72	3	357	25	181
(WE) Windsor/Essex  [24/ 10,758/1,785.1/415,293]	1	768	42	25	13	20	17	19	17
	1-2	1027	35	27	9	3	34	3	30
	2	7804	212	83	91	86	121	116	68
	2-3	530	6	3	0	0	3	3	3
	3	1800	43	11	8	0	41	2	31

¹ Map of the southern half of Ontario with two-letter abbreviations of each region of interest indicating their locations within the area. Additional maps below display the border and PC boundaries within each region.
² Area of region (km²).
³ Population based on 2016 census.
⁴ Interwaves 1-2 and 2-3 are displayed as '1-2' and '2-3', respectively.
⁵ Estimated COVID-19 cases in FSAs associated with region (days where cases within FSA were masked were assumed to be 1 case).
⁶ PCs identified as a high-case outlier on a particular 3-day interval with ≥ 6 cases when analyzing stratified case data.

B. Distribution of high-case cluster postal codes in different municipalities									
City	Wave ¹	Estimated total cases ²	Number of high-case cluster PCs in stratified populations ³						
			All cases	Males	Females	≥65 years	<65 years	At-risk	Healthy
(HA) Hamilton	1	787	2	0	0	1	0	1	0
	1-2	410	0	0	0	0	0	0	0
	2	8788	28	5	4	0	23	0	10
	2-3	1464	5	0	1	0	5	0	1
	3	9755	52	25	17	0	47	5	34
(KW) Kitchener/ Waterloo	1	511	3	0	3	3	0	3	0
	1-2	277	0	0	0	0	0	0	0
	2	4287	3	1	0	0	1	0	0
	2-3	560	2	0	0	0	2	0	0
	3	2733	1	0	0	0	0	0	0

Table 1. Continued

B. Distribution of high-case cluster postal codes in different municipalities									
City	Wave ¹	Estimated total cases ²	Number of high-case cluster PCs in stratified populations ³						
			All cases	Males	Females	≥65 years	<65 years	At-risk	Healthy
(LO) London	1	363	0	0	0	0	0	0	0
	1-2	213	0	0	0	0	0	0	0
	2	3609	9	0	1	1	4	4	0
	2-3	347	0	0	0	0	0	0	0
	3	4363	7	0	0	0	7	0	3
(OT) Ottawa	1	1125	8	1	0	6	0	7	0
	1-2	1327	5	0	1	0	5	0	3
	2	6790	21	2	4	0	17	0	7
	2-3	1178	1	0	0	0	1	0	0
	3	8716	28	6	7	0	22	0	17
(TO) Toronto	1	7144	22	0	7	2	8	9	1
	1-2	3148	4	0	4	4	0	4	0
	2	54,460	183	19	30	7	149	11	73
	2-3	7309	72	9	9	1	57	6	19
	3	52,037	172	11	12	0	136	3	51
(WE) Windsor/ Essex	1	768	2	2	0	0	2	0	2
	1-2	1027	17	7	1	0	15	0	12
	2	7804	36	13	4	0	23	0	11
	2-3	530	1	0	0	0	1	0	0
	3	1800	5	0	0	0	5	0	0

¹Interwaves 1-2 and 2-3 are displayed as '1-2' and '2-3', respectively.
²Estimated COVID-19 cases in FSAs associated with region (days where cases within FSA were masked were assumed to be 1 case).
³PCs identified as a high-case outlier on a particular 3-day interval with ≥ 6 cases when analyzing stratified case data.

Windsor/Essex (Extended data,²⁹ Section 1 - Table S4). Additionally, interpolated case counts by kriging tended to be less than actual FSA case counts, due to dispersion of cases across more PCs. Most municipalities contained more FSAs with localized hotspots that were not significant by Gi* analysis. Increased case counts across the province during this wave increased the significance threshold for Gi* analysis. This decreased the likelihood that FSAs in municipalities with moderate population densities were designated significant hotspots. For example, FSAs within Kitchener/Waterloo and London were not significant COVID-19 hotspots during wave 2 (Extended data,²⁹ Section 1 - Table S4). Hotspots in these regions were coincident with maximum contour levels interpolated by kriging. PCs had a greater proportion of high-case clusters and/or outliers in younger and healthy populations, in contrast with wave 1 (Table 1). The reduction of COVID-19 outbreaks (i.e., hotspots) in long-term care homes during this period was supported by reduced statistical significance of counts at these locations relative to increased background case counts across the province.

Localized hotspots detected in PCs by kriging were more frequent in Hamilton, London, Ottawa, Toronto and Windsor/Essex during this wave than in the preceding one (Extended data,²⁹ Section 1 - Table S4). For example, an interpolated hotspot in the N5Y FSA (London) with >5 interpolated cases over consecutive 3-day intervals (December 19-21 and December 20-22, 2020; Figure 6) represents a 'streak', since kriging analysis interpolated cases over this FSA throughout the previous week (however, it was below the kriging threshold for significance). This hotspot encompasses a 2 building apartment complex that was publicized on December 29 by the local health authorities as a COVID-19 hotspot, two weeks after it was first detected by kriging.³³ While the interpolated case count is lower than the total count in this FSA on this date (N=33, other cases are present at other locations), this hotspot encompasses only 4% of the FSA by area. Other localized hotspots during this wave are depicted (e.g. in Leamington ON; Extended data,²⁹ Section 1 - Figure S6) and catalogued in Extended data,²⁹ Section 2.

Significant clustering of cases in PCs in close proximity based on Spatial Autocorrelation analysis may be distinct from other tests that identify localized hotspots. Spatially clustered localized COVID-19 hotspots may not be confined within

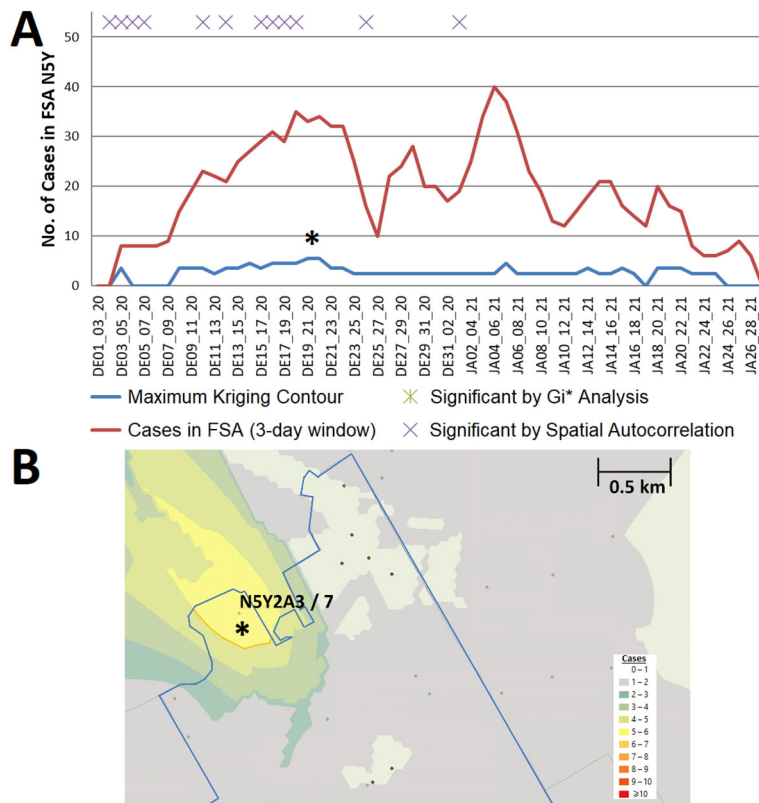


Figure 6. Example of a localized hotspot in London Ontario. (A) Geostatistical analysis of the FSA N5Y (London, ON) identifies a localized hotspot from Dec. 19th to the 22nd. Estimates from kriging were lower than actual case counts. (B) The kriging contour of a localized hotspot (yellow) interpolated with ≥ 5 cases encompasses 4% of the FSA area. Dots indicate PCs at least 1 case, however only those with ≥ 6 cases are labelled (with PC identifier/case count).

the same region of an FSA. For example, FSA M1P (Toronto) from Nov. 1-3 to Nov. 3-5, 2020 exhibited significant case clustering based on G_i^* analysis. However, a localized hotspot identified within this FSA during this time did not overlap this cluster (Extended data,²⁹ Section 1 - Figure S7A). A neighboring FSA M1C showed significant spatially correlated cases within the same timeframe (Nov. 6-8, 2020), however localized hotspots were not detected in this FSA (Extended data,²⁹ Section 1 - Figure S7B). No PC within this FSA had ≥ 6 cases during this period. This illustrates how combining Spatial Autocorrelation, G_i^* and kriging analyses can improve the interpretation of asymmetric COVID-19 case distributions.

Interwave 2-3

The interval between the end of wave 2 and inception of wave 3 was short (~31 days), however by comparison with interwave 1-2, the case counts were considerably higher (averaging 1134 and 154 per day, respectively). Thus, a far greater number of high disease burden FSAs were identified during this period (1.9 and 4.7-fold more often compared to wave 1 and interwave 1-2, respectively; Extended data,²⁹ Section 1 - Table S2). However, most of these occurred in Toronto and Windsor/Essex regions only (Extended data,²⁹ Section 1 - Table S4).

Wave 3

The frequencies of high disease burden FSAs in waves 2 and 3 were comparable upon adjusting for duration of each wave (Extended data,²⁹ Section 1 - Table S2). The regions detected by geostatistical analyses were different. High disease burden FSAs implicated in Hamilton and Kitchener/Waterloo by G_i^* analysis were significantly more frequent compared to the preceding wave, while FSAs within Toronto were less often significant (Extended data,²⁹ Section 1 - Table S4). By contrast, localized hotspots within FSAs were generally less frequent than in wave 2, most notably in the Windsor/Essex region which contained 389 localized hotspots in wave 2, but 80 in wave 3 (Extended data,²⁹ Section 1 - Table S4).

High-case clusters and outliers were seldomly identified in older (≥ 65) populations in wave 3 and interwave 2-3 (Table 1). Similarly, there is a considerable decrease in the number of case clusters/outliers in at-risk populations relative

to previous waves. For example, analysis of at-risk populations in Windsor/Essex identified 116 PCs as a high case outlier on a particular date during wave 2, but only 2 in wave 3 (Table 1A). These observations may be related to effectiveness of vaccination, based on the initial prioritization of older and at-risk cohorts for immunization.

The distribution of and relationships between clustered PCs with persistent COVID-19 cases

COVID-19 distributions were evaluated by Cluster and Outlier analysis to identify neighboring PCs with high concentrations of cases. PCs deemed part of a high-case cluster that were clustered over contiguous date ranges were then identified. These PCs were recognized as clustered over the duration of a time “streak”, which by definition, allows for a single day in which the PC was either not clustered or did not meet the minimum case count threshold of ≥ 6 cases within the 3-day sliding window. PCs that were part of a high-case cluster, but did not meet the case threshold, were not reported. This analysis identified 20 unique PCs with one or more high-case cluster streaks in Hamilton, 5 in Kitchener/Waterloo, 10 in London, 24 in Ottawa, 216 in Toronto, and 19 in Windsor/Essex (Extended data, 29 Section 1 - Table S5). Streaks were uncommon in Kitchener/Waterloo and London, as PCs in these cities rarely fulfilled the minimum case threshold. Most streaks were shorter than 4 sliding window periods, indicating that the most frequent streaks were brief due the large number of single day discontinuities from case reporting patterns and ascertainment bias. For these reasons, ≤ 3 -day bins were not considered evidence of a streak. However, 37 streaks longer than 3 days were interpreted to indicate persistent infections at the same locations. Streaks with more than 3 consecutive high-case clusters were uncommon in wave 1, with M6M 2J5 (West Park Long Term Care Center) as the only example of a streak exceeding 3 days. The lack of other locations with persistent infections may be related to inconsistent case reporting which was hampered by limited testing during this period. 4,34

We sought evidence of COVID-19 transmission by identifying pairs of PCs with high case counts that were adjacent to one another by both distance and time. In Figure 7A, we analyzed PCs and their 10 closest neighbors, where streaks were either concurrent, consecutive, or separated by short gaps within the same wave. Concurrent streaks of neighboring PC pairs spanned at least 3 and as many as 29 days in Hamilton, Ottawa, Toronto, and Windsor/Essex (Extended data, 29 Section 1 – Table S6). There were a greater number of discontinuous streaks across all municipalities, with gaps separating streaks in neighboring PCs varied from short to long intervals (Extended data, 29 Section 1 - Table S7). During the second wave, many pairs of PCs were comprised of neighboring apartment buildings, for example, within Toronto FSA M4H. Other long concurrent paired streaks (exceeding 1 week in duration) occurred in Toronto PCs M1R 1S9 and M1R 1T1, MIL 1K9 and MIL 1L1 (both PC pairs in North York, Toronto), and L0R 1C0 with its neighbors in Hamilton (Extended data, 29 Section 1 – Table S6). Pairs of concurrent streaks in close proximity did not occur during wave 1, since streaks were uncommon during this time. However, many neighboring PCs were both found to be clustered both spatially

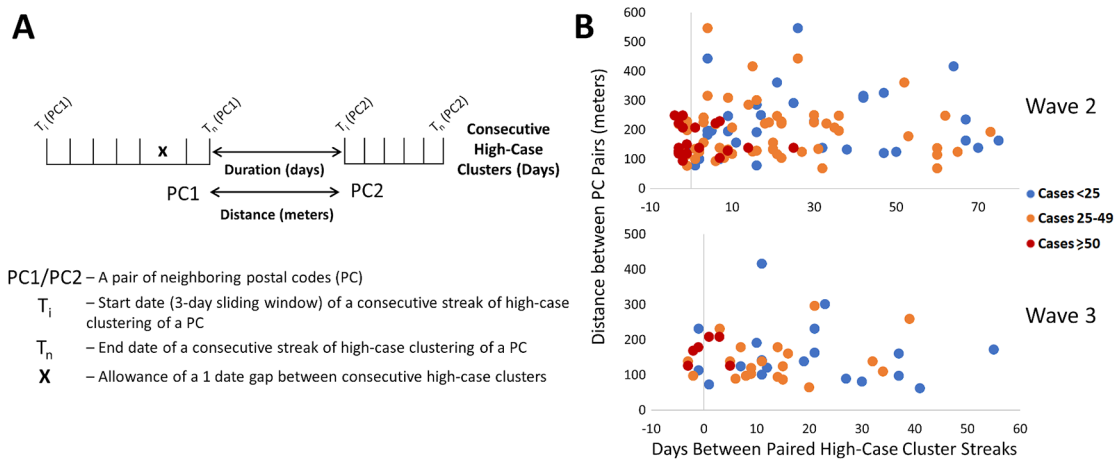


Figure 7. Contiguous or consecutive pairs of Toronto postal codes with COVID-19 case clusters lasting multiple days. Pairs of neighboring PCs with high-case count clusters were identified by Cluster and Outlier analysis. Panels show: (A) parameters measured in PC pairs that were elements of significant high-case clusters over multiple consecutive days (allowing for a single day gap), termed “streaks”; (B) characteristics of neighboring pairs of PCs with multi-day streaks in close proximity in Toronto (of ten closest neighbors based on distance between centroids) during waves 2 (top) and 3 (bottom). Wave 1 is not indicated because streaks were infrequent. Each point represents a pair of PCs, with color codes (see legend) indicating ranges of total number of cases in both, the X axis indicates the duration between PC pairs (with negative values representing number of days of overlap between end of first PC streak and beginning of the second), and the Y axis indicates spatial distance between PCs constituting the pair.

and temporally during interwave 1-2 and waves 2 and 3 (Extended data,²⁹ Section 1 - Table S7). While not concurrent, neighboring PC pairs in rural Essex county were clustered within 10 days of one another during interwave 1-2 (N0P 2L0, N0P 2P0 and N0R 1B0). This is confirmed by publicly reported outbreaks in agriculture facilities of this region during this time. During wave 2, 56 unique pairs of streaks were identified in neighboring PCs in Toronto (18 clustered concurrently, another 14 within a week of one other), 4 pairs in Hamilton (2 concurrent), 1 in London (concurrent), 4 in Ottawa (1 concurrent) and 14 in Windsor/Essex (10 concurrent; all involve PCs in Essex county). There were fewer incidences of streaks during the third wave, which may be related to the effectiveness of vaccination efforts which limit transmission (Figure 7B). During wave 3, we identified an additional 30 pairs of PCs with streaks in Toronto (7 are clustered concurrently, and an additional 4 pairs are offset by 1 week), and 6 in Hamilton (3 concurrent, and another pair offset by 2 days). Interestingly, 9 of the same PC pairs in Toronto clustered during the third wave were also flagged during the second wave, which could be an indication of an undetected reservoir of COVID-19 that persisted in this population during interwave 2-3. These pairs of PCs occurred in the Toronto neighbourhoods of East York (M4H 1J4 and M4H 1J5), Scarborough (M1H 2Y7 and M1H 2E9), North York (M1R 1T1 and M1R 1S9; M6M 5B3 and M6M 5B7), and Etobicoke (M9C 1G7 and M9C 1G6; M9V 3S6 and M9V 4A4; M9V 3S6 and M9V 4M1; M9V 4A4 and M9V 4M1; M9W 6L4 and M9W 6A7).

We examined evidence for likely community spread of COVID-19 between infected individuals in different pairs of neighboring PCs. This involved determining the distances separating them, the duration between streaks in the respective PCs, the total case counts among PC pairs during these streaks, and then assessing the significance of these characteristics. During waves 2 and 3 in Toronto, paired concurrent or consecutive streaks in such PCs were common (Figure 7B). Frequencies of these PC pairs were negatively correlated with the duration of the gap between the end of a streak in one member of the pair and the beginning of the other ($r = -0.57$ and -0.50 for waves 2 and 3, respectively). To assess the significance of these results, a multinomial logistic regression model was derived in R ("Streak_Analysis_using_Multinomial_Logistic_Regression_Models.Rmd" which uses the *multinom* function of the *nnet* package; available in Zenodo archive²⁹), evaluating (1) the total number of cases in each pair of PCs over the course of both streaks (classified as either <25, between 25 and 49, or ≥ 50 cases; the response variable of the regression), (2) the duration or gap (in days) separating streaks in a pair of PCs, and (3) the distances between the centroids of both PCs. The duration between streaks was significantly shorter for PC pairs with ≥ 50 cases relative to those with <25 cases ($p = 0.0005$ and 0.0141 for waves 2 and 3, respectively, using the Wald two-tailed z-test on the coefficients of the model). The same relationship was also significant when comparing streaks in pairs of PCs with ≥ 50 cases and those with 25-49 cases ($p = 0.0015$ and 0.0277 for waves 2 and 3, respectively). Total case counts and time between streaks in a pair of PCs were also inversely correlated ($r = -0.41$ in both waves 2 and 3; using the *CORREL* function in Excel), which is consistent with disease transmission between adjacent PCs being more common when cases were more abundant. These parameters were also inversely correlated ($r = -0.13$ and -0.17 , respectively) in Ottawa and Windsor/Essex, but under less stringent criteria that allowed ≥ 6 cases for only one of the PCs in a pair on a single day. The relationship between distance between PCs and total number of cases was not significant by the regression model in either of these waves. Increasing the number of PC neighbors considered over a larger range of distances did not alter this result (Extended data,²⁹ Section 1 - Figure S8). However, upon dividing groups of neighboring PCs thresholded according to average quartile values (distance [182 m] and average gap between streaks [17.5 days]) distance between neighbors was significant when combined with the degree to which they coincided temporally. Below average duration between streaks *and* distance separating PCs were 2.8-fold more likely to exhibit ≥ 50 cases than all other PC pairs ($p = 0.002$ by Mantel-Haenszel chi square test). Transmission of higher levels of COVID-19 between pairs of PCs is supported by abbreviated intervals between streaks and their proximity to one another.

Multiple paired PC streaks in Toronto were tightly grouped within a single neighborhood over a short period of time (Extended data,²⁹ Section 1 - Table S7). Directional network diagrams were created to illustrate the connectivity and temporal relationships between these and other paired high-case cluster streaks in close proximity (Extended data,²⁹ Section 1 - Figures S8 and S9 for waves 2 and 3, respectively). The preponderance of paired streaks were bimodal (9 paired streaks where only two neighbors interact [Extended data,²⁹ Section 1 - Figure S9D], of which 17 occurred in wave 3 [Extended data,²⁹ Section 1 - Figure S10A]), and in close succession (≤ 7 days between streaks). Multi-node networks involving ≥ 3 postal codes occurred in 3 distinct neighborhoods in wave 2, and another three neighborhoods within wave 3. These neighborhoods were often comprised of multiple adjacent apartment buildings, and rarely occurred in low-density residential areas. The PCs where streaks tended to coincide were often proximate to facilities where communities may have congregated (grocery stores, schools, parks, or religious sites). We investigated whether the participation of many PCs in a multi-node network was a result of higher case counts during a streak (which would suggest an increased probability of transmission between these locations). However, bimodal networks of two PCs with recurrent high case counts were not rare (e.g., >100 cases within a 9-day streak in M1R 1T1), so other factors are necessary to explain why some PCs form multi-node networks. Furthermore, the average numbers of cases within streaks in bimodal networks and multi-node networks were comparable (18.8 and 17.0 cases per streak, respectively).

In some instances, PCs began to accrue cases prior to their inclusion in significant clusters of PCs or reaching the minimum numbers of cases to be designated a COVID-19 hotspot (M4H 1K1, M4H 1K2, M4H 1K4, M4H 1L1, M4H 1L6, and M4H 1L7). These PCs were nevertheless elements of the M4H FSA streak. Similarly, M4H 1L3 and M4H 1L4 were elements of a cluster containing M4H 1J4 on Nov. 8-10, 2020, at least 2 weeks prior to the occurrence of the corresponding multi-day streaks. Consecutive streaks within a set of clustered PCs in Etobicoke (M9V 3S6, M9V 3Z8, M9V 4A4, M9V 4A9, M9V 4M1, M9V 4P1, and M9V 5G9; centroids were all within 0.63km of one another) occurred between mid-October and late December 2020 (Figure 8B). M9V 5G9 is separated from the other PCs by a wooded area,

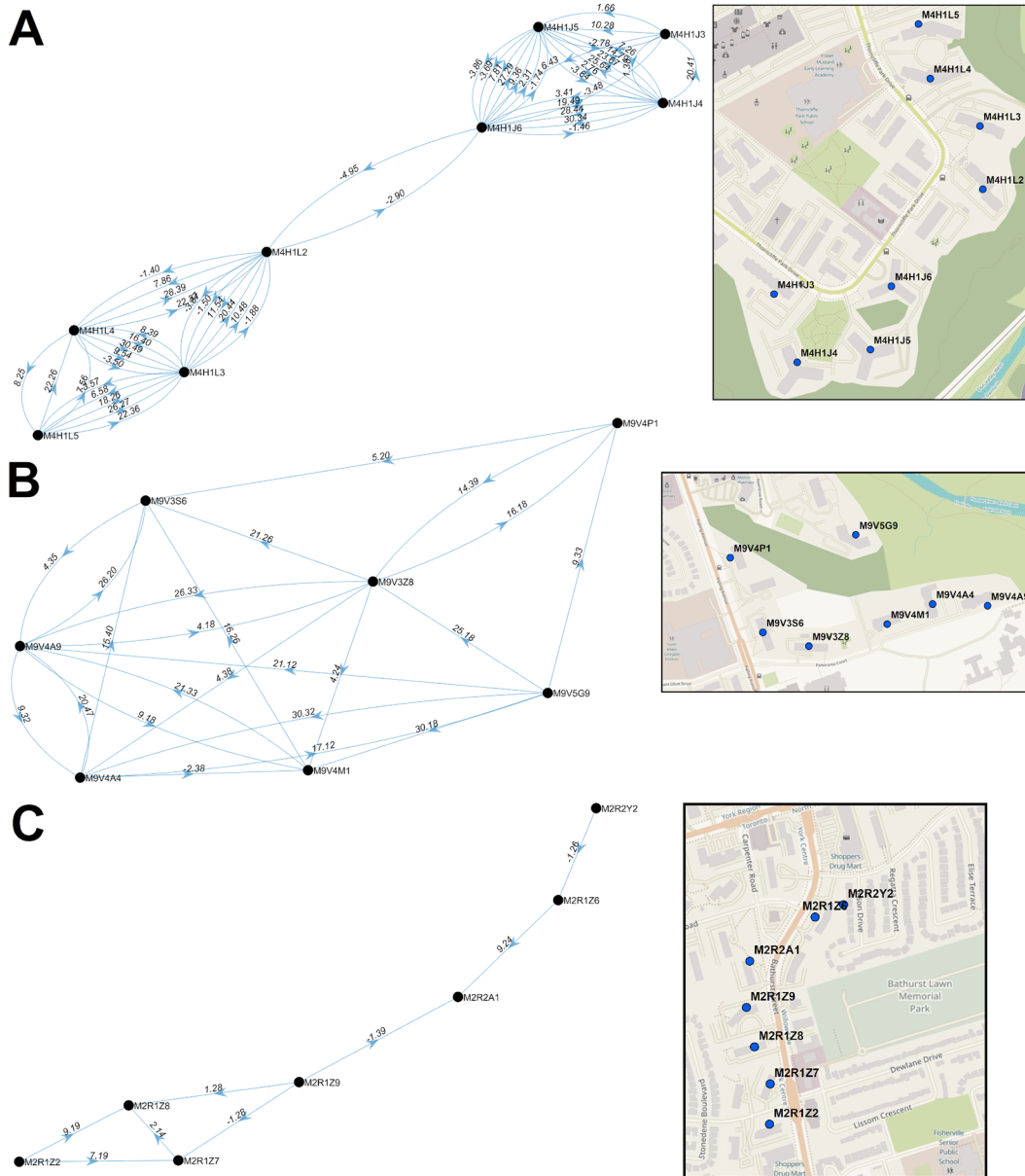


Figure 8. Multi-node directional networks of closely situated clustered postal codes. Directional acyclic graphs (left) organize adjacent high-case clustered PC streaks within the same FSAs ordered according to their occurrence. Panels indicate PCs within (A) M4H and (B) M9V in Toronto during wave 2. Each connection (or ‘edge’) linking two PCs represents a pair of clustered streaks occurring within 30 days of each other (all significant PC pairs are indicated in Extended data,²⁹ Section 1 - Table S7). Labels on each edge indicate the duration (in days) separating adjacent PC pairs and the number of cases which occurred in the combined streaks (“dates.cases”). Negative values, when present, indicate the number of overlapping days of concurrent streaks in pairs of PCs. Arrows indicate the temporal order of these paired streaks (from earlier to the later streak). Maps (right) show the physical locations of each PC within the corresponding networks.

but all are close to a major thoroughfare, Kipling Avenue. A striking directional acyclic network of COVID-19 cases involved 7 adjacent PCs whose streaks occurred concurrently and were occurred nearly consecutively along Bathurst Street, Toronto (M2R 1Z2, M2R 1Z7, M2R 1Z8, M2R 1Z9, M2R 2A1, M2R 1Z6, M2R 2Y2; within 0.57km of one another; [Figure 8C](#)). Curiously, the order of occurrence of COVID-19 infections among these PCs consistently followed a southerly direction along Bathurst Street.

The multi-node networks during wave 3 consisted of only three postal codes (M3C, M9V, M9W; Extended data,²⁹ Section 1 - Figure S10A). Increasing the number of neighbors (by relaxing distance constraints, from 10 to 30 PCs) for these PC networks results in multi-node networks that were previously identified as bimodal (M4H, M4C, and M4X; Extended data,²⁹ Section 1 - Figure S10B). This procedure also resulted in four additional bimodal pairs of PCs in wave 2 (Extended data,²⁹ Section 1 - Figure S9D); however, the total number of PCs comprising the multi-node networks in wave 2 was unchanged (Extended data,²⁹ Section 1 - Figures S8A-C).

Discussion

Geostatistical analysis revealed the locations and relationships between COVID-19 hotspots, utilizing daily postal code-level case data in Ontario Canada from March 26th, 2020 through June 25th, 2021 (including G_i^* , Spatial Autocorrelation, Clustering and Outlier analyses and interpolation by kriging). Previous efforts to integrate geostatistical and epidemiological data have not investigated COVID-19 at this level of geographic and temporal resolution. This approach more precisely localized COVID-19 hotspots, and enabled early detection of disease clusters, which could be exploited to prevent or limit further spread at or close to these hotspots (provided complete and up-to-date COVID-19 case data is made available in a timely manner). Concurrent or consecutive neighboring hotspots classified as persistent high-case clusters were flagged as possible evidence of disease transmission.

The COVID-19 epidemic in Ontario has been classified into three distinct waves of infections.³⁵ The recurrence of successive waves shortly after easing of public health measures strongly motivated development and implementation of novel strategies that anticipate the locations and timing of future hotspots. Network analysis demonstrated that these locations may form nuclei from which subsequent local outbreaks arise. COVID-19 clusters were considerably diminished in the older at-risk population during the third wave 3, most likely, because of province-wide vaccination compliance, which resulted in lower mortality in long-term care residences during this period.

The techniques implemented in this study could prove useful to assist with contact tracing of infectious diseases, which may be challenging to carry out in a timely fashion once community spread has already occurred. Contact tracing has been vital to identify potential sources of viral transmission from infected cases. Successful contact tracing is impacted by incomplete collection of information, specifically interviewees who provide imprecise information, omit critical interactions, or who are uncooperative.³⁶ The directional acyclic network graphs presented are consistent with transmission between individuals residing in multiple PCs in close proximity to each other. These tools could assist tracing investigations by focusing resources on the most likely locations where interactions between infected and uninfected persons occur. It may be possible to exploit observations of COVID-19 infections at consecutive, neighboring PCs to anticipate the occurrence and locations of subsequent infections. This was exemplified by the progression of cases along PCs on Bathurst Street, Toronto in wave 2 ([Figure 8C](#)). We suggest that these and similar patterns may be useful as features along with cellular phone mobility data from dates preceding the appearance of future hotspots for predictive machine learning models. Accuracy in predicting the timing of future hotspots, however, would be influenced by the lag between date of infection and presentation of symptoms and by compliance of symptomatic individuals in seeking testing. Application of geostatistical analyses and other technologies³⁷ may overcome the challenges posed by non-compliance or erroneous COVID-19 information from infected persons.

We also identified persistent, clustered postal codes that recurred between waves. Neighboring PCs comprising high-case clusters over a series of consecutive or overlapping dates were proposed as evidence of COVID-19 transmission. Nine PC pairs across 4 separate regions in Toronto were flagged in both waves 2 and 3, which were suggestive of recurrence. Without viral sequencing data, it is not possible to unequivocally establish whether these local outbreaks were independent events or due to re-emergence of the same infection from a low-level reservoir of asymptomatic or persistent mild infections during interwave 2-3. The same type of analysis did not yield evidence of inter-wave COVID-19 transmission in the other municipalities evaluated. Because of their comparatively lower population densities and smaller geographic size, it is less likely that the case thresholds required to define a COVID-19 hotspot within a 3-day window can be met in other Ontario municipalities. Increasing the aggregation of cases from 3- to 7-day sliding windows might improve detection of likely COVID-19 transmission in these areas.

Unlike traditional geostatistical epidemiology, we did not determine the likelihoods of hotspots based on simulations of random distributions of the same number of cases in the analyzed region,^{14,38} since background levels of COVID-19 were undetectable prior to the pandemic. Therefore, performing geostatistical tests on positive COVID-19 cases is valid statistically. FSA-level analysis on population-normalized case counts (cases per 100,000; where FSAs with <100 individuals were not analyzed) was also performed (March to November 2020).²⁹ Analyses could not be normalized by population densities (cases per 100,000) since these data are not publicly available for individual PCs. However, the smallest geographic units defined by Statistics Canada, termed Dissemination Areas, encompass an average of 15 postal codes (consisting of 400-700 persons).^{39,40}

A seroprevalence study of COVID-19 antibodies in Canada has suggested that infection rates were up to 3-fold higher than detected by RT-PCR testing during the first wave.³⁴ Similar studies have also indicated underreporting in the U.S.⁴¹ and Europe.^{42,43} This included asymptomatic and untested symptomatic infected individuals⁴² or who received negative results.⁴⁴ Our study assumes that the geographic and temporal distribution of unrecognized infections trends similarly with known cases. However, nonuniformity in the distribution of asymptomatic and undetected cases across Ontario could have altered the detection or localization of hotspots. It seems likely that significant hotspots detected by geostatistical methods that also fulfill minimum case count thresholds would seem to be less likely to present as false positives or be incorrectly localized.

Some geostatistical functions (e.g., Gi* and Cluster and Outlier analysis) could be computed rapidly (within seconds for a single day analysis), while others (e.g., intersection of kriging contours PC boundaries) required up to 10 minutes per date analyzed on the ICES research computing system. In order to scale these computations over the duration of this project, it was critical to automate analyses for Provincial and multiple Municipal jurisdictions for timely generation and interpretation of these results. In April 2021, the sponsor of our AHRQ project requested an analysis of recent COVID-19 count and location data (January-March 2021). We automated kriging of localized hotspots in Toronto, Windsor/Essex, Ottawa and London within 5 days from the receipt of the request. This entailed performing the software analysis, imaging the results as maps, and interpreting these results. The rapid turnaround of geostatistical analysis of ongoing infectious diseases could have practical value for time-sensitive, critical public health management decisions. These programs have been made as a publicly available package, reformatted as a user-friendly toolbox in ArcMap (the *Geostatistical Epidemiology Toolbox*).²⁹

FSA-level geospatial tests, such as Gi*, did not reveal the distributions of COVID-19 infections, a limitation of this approach. We attempted to mitigate this by integrating the results of multiple geostatistical methods at different levels of resolution (e.g., Spatial Autocorrelation determined if cases were clustered at the PC-level). The locations of individuals in the same FSAs and PCs were aggregated at the centroids of their respective boundaries. However, boundary shapes and areas covered by different FSAs or PCs can vary, even though the total populations within each are more similar. Rural FSAs can be larger than urban counterparts and are occasionally non-contiguous (e.g., the FSA N0P). While close proximity of COVID-19 positive individuals in high-density municipalities is more consistent with transmission, it was nevertheless feasible to detect hotspots in rural regions. However, high concentrations of cases are required to find statistically significant hotspots in these regions. Since we use PC centroids to assign cases, this also affects the contours defined by kriging analysis of rural regions. These are limitations of the imposed boundaries of FSAs and PCs and not of the geostatistical tests used. Due to privacy considerations imposed by data providers, we could not perform kriging analysis on exact locations of COVID-infected individuals. Kriging interpolation can overlap multiple FSAs (Extended data,²⁹ Section 1 - Figure S11), which can be advantageous over area-to-area geostatistical tests, as hotspots in close proximity that cross an FSA boundary could be missed with these tests. However, FSAs with few or no cases that are adjacent to a local hotspot could mistakenly be implicated. We account for this by ignoring FSAs without case counts, however this effect can still impact interpolation accuracy. For example, the interpolated case count for the FSA N8H in late June 2020 was much higher than the ground truth (Extended data,²⁹ Section 1 - Figure S5B). Over the same period, 109 cases were reported over a 3-day period within N0P 2J0, which is immediately north of N8H. The high-case kriging contour detecting this event intersects the boundary of N8H, thus causing this discrepancy (Extended data,²⁹ Section 1 - Figure S5C).

In late 2020, Ontario announced prioritization of COVID-19 vaccinations beginning with those with the highest risk of contracting severe illness. The subsequent group included individuals residing in hotspot communities.⁴⁵ These hotspots were identified in FSAs with the highest counts per 100,000 population. However, cases are not always uniformly distributed within FSAs with high disease burden (Figure 4). This study illustrates the importance of identifying hotspots at a higher resolution and describes how this level of precision provides targeted information about transmission that is often not feasible with FSA-level analyses. To minimize high rates of community transmission in the future, we suggest that targeted vaccination campaigns focus on highly constrained, localized hotspots with high case counts at the earliest possible juncture (rather than broadly across a region or FSA). For example, a hotspot in London was detected by kriging

over several weeks at levels below our pre-selected interpolated case thresholds (Figure 6). The results of this analysis surpassed these thresholds approximately one week before it was publicly revealed. Thus, targeting can achieve the necessary precision and at an earlier stage with these methods. Smaller targeted locations may also facilitate communication, promote compliance and implementation of individual public health measures.

Data availability

Zenodo: Data and Software Archive for ‘Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada, <http://doi.org/10.5281/zenodo.5585811>.²⁹

This project contains the following underlying and extended data, organized across 4 sections. Section 1 and 2 contain extended data, and Sections 3 and 4 respectively consist of data files containing and map image files displaying the underlying statistics from the various geostatistical analyses performed in this study.

Extended data

Zenodo: Data and Software Archive for ‘Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada’: <http://doi.org/10.5281/zenodo.5585811>.²⁹

This project contains the following extended data:

- Section 1 – An Excel document containing 7 additional tables described in this study which provide summary statistics of various geostatistical tests (“Section 1 – Tables S1-S4, S6”) and lists all identified single and paired high-case cluster streaks (“Section 1 – Tables S5, S7”). This section also contains 11 additional figures referred to in the manuscript (“Section 1 – Figures S1-S11”) with a Word document which describe them. Figure legends are as follows:
 - **Figure S1. Case Count of FSAs by Frequency.** Histograms indicate the frequency in which an FSA in (A) Ontario and (B) Toronto were reported to have $\geq X$ COVID-19 cases (where X is the value on the X-axis) on a specific day from March 1st to September 24th, 2020 (blue columns). G_i^* analysis identified which FSAs were significant hotspots by comparing them with neighboring FSAs at a fixed distance (indicated with red bars) and inversely weighted distances (indicated with green bars), at a threshold of $<50\text{km}$. A large fraction of FSAs in Ontario and Toronto contained fewer than 8 cases and were not statistically significant. Over two thirds of FSAs with at least 10 cases were significant by the G_i^* test at a 99% confidence level.
 - **Figure S2. Evaluation of Kriging Semivariogram Models.** Interpolation by different semivariogram models was evaluated using kriging to determine which model best approximated true case counts in regions of moderate-to-high population densities. The two best performing models chosen were ‘Power’ and ‘Thin Plate Spline’ (TPS). Each point in this figure represents PCs with ≥ 1 COVID-19 case and PCs with ≥ 6 cases have been labelled [PC/case count]. Panel (A): Kriging analysis of Hamilton (moderate-density of cases) indicates Power-modeled kriging (left) shows smooth fit to the actual case distribution compared to TPS, which is discontinuous (right). Panel (B): In Toronto (high-density of cases), interpolation with the ‘Power’ model more frequently under-estimated actual case counts, whereas the TPS model derived case counts that more precisely resembled the actual distribution of cases.
 - **Figure S3: Association of Kriging Interpolated and True Case Counts.** The distribution of COVID-19 cases within the FSAs (A) M9P and (B) M9V in Toronto, ON during wave 1 was analyzed with 3 methods: kriging (blue line), Spatial Autocorrelation (purple crosses) and G_i^* analysis (green asterisks). Kriging-interpolated case values (blue line) were similar to actual case counts (red line) for (A) (but not B), indicative that cases are concentrated within a few PCs. COVID-19 hotspots identified by G_i^* analysis (green asterisks) were often identified several days after one or more local hotspots were identified by kriging analysis.
 - **Figure S4. Histograms of the Fraction of FSAs Overlapped by Kriging Contours of Interest.** The percent overlap of the areas covered by kriging-derived contours and the corresponding FSAs were computed by area for all FSAs flagged by G_i^* analysis. Panels depict (A) wave 1, (B) interwave 1-2, (C) wave 2, (D) interwave 2-3, and (E) wave 3. Histograms illustrate the distribution of the fraction of each FSA that is geographically intersected with either the highest interpolated kriging contour level (left) or kriging contours interpolating at least 5 (right).

- **Figure S5. Hotspots Localized within High Disease Burden Regions.** The distribution of COVID-19 cases within the FSAs (A) N9Y (Kingsville, ON) and (B) N8H (Leamington, ON) during interwave 1-2 was evaluated by kriging, Spatial Autocorrelation and G_i^* analysis. Kriging-interpolated case values (blue line) were similar to actual case counts (red line). COVID-19 outbreaks within these FSAs were often identified by G_i^* analysis (green asterisks), however the distribution of cases within these regions were infrequently found to be spatially clustered (purple crosses). Panel (C) indicates Power-based kriging contours can be large in regions with low population densities due to the distribution of cases. Interpolation of cases within N8H from June 22nd to June 26th, 2020 exceeded than actual counts due to the contribution of a high-case postal code in a neighboring FSA (NOP), with the resultant kriging contour intersecting both FSAs.
- **Figure S6. A Localized Hotspot in Leamington, Ontario.** PC-level COVID-19 case data of the FSA: N8H (Leamington, ON) on consecutive dates in December 2020 evaluated by kriging interpolation (blue line), testing for significance by Spatial Autocorrelation (purple crosses) and G_i^* analysis (green asterisk). B) Map contours and distribution of cases. Small circles in this region represent PCs with one or more cases from December 21st to the 23rd, 2020. The maximal kriging contour was coincident with the hotspot in PC N8H 3V5 during this window.
- **Figure S7. Spatial Clustering and the Identification of Localized Hotspots.** Spatial autocorrelation analysis identifies clusters of postal codes with high case counts within FSAs. Panels (A) displays significantly clustered cases in Toronto FSA M1P from Nov. 3-5, 2020. This region was also significant by G_i^* and contained a localized hotspot consisting of 8 cases (M1P 4V6). Each dot represents a postal code with at least 1 but fewer than 6 cases. (B) Cases in M1C were also significantly clustered from Nov. 6-8, 2020, but no hotspots were identified. Topographic contours indicate kriging-interpolated case counts.
- **Figure S8. Contiguous or Consecutive Toronto Paired Streaks considering 50 Closest Neighbors.** High-case postal code clusters were identified by Cluster and Outlier analysis (Local Moran's I). High-case cluster streaks (PCs significant over multiple consecutive days, allowing for a single day gap) were identified and paired to other streaks within close proximity (considering 50 closest neighbors based on distance between PC centroids). Each point represents a paired PC streak with colors representing the total number of cases in both, the X axis indicates the duration between PC pairs (with negative values representing degree of overlap between end of first streak and beginning of the other), and the Y axis indicates spatial distance between the centroids of each PC forming a pair. All paired streaks are illustrated, regardless of wave.
- **Figure S9. Directional Acyclic Graph Network of Clustered Streaks in Toronto during Wave 2.** Directional networks identify streaks of clustered high-case counts (≥ 6 cases for each date evaluated) in close spatial and temporal proximity. Each connection (or edge) represents two neighboring PCs, where the value preceding the decimal represents the days separating pairs of streaks followed by the total number of cases in both ("dates.cases"). The arrows represent the order in which streaks occur in the pair of PCs. The left and right panels indicate the same network, based on either the 10 or 30 closest neighbors, respectively. Maps indicate the locations of PCs within the network (middle). The panels highlight different networks in neighborhoods according to the FSA: in which they occur: (A) M4H, (B) M9V and (C) M2R. Although increasing the number of neighbors considered for pairing added new edges, novel postal codes were not introduced. (D) Networks only connecting streaks between a single pair PCs are indicated.
- **Figure S10. Directional Networks of COVID-19 Clustering Streaks in Toronto during Wave 3.** Directional networks illustrate the relationship between PCs with consecutive COVID-19 clustering (i.e., streaks) in Toronto during Wave 3 of the COVID-19 epidemic, considering the (A) 10 closest and (B) 30 closest neighboring PCs for each streak. Simple networks involving two or three PCs were common and exhibited recurrent transmission between nodes. Increasing the number of allowable neighbors considered increased network complexity, including those not previously seen during Wave 2 (M3C, M4C and M4X). Networks for Wave 3 are less complex compared to those in Wave 2, possibly due to the increase of vaccinated individuals within the population.
- **Figure S11. Overlap of Localized Hotspots with Multiple FSAs.** Kriging contours can overlap multiple FSAs. While the kriging semivariogram model selected can affect their frequency, contours which intersect ≥ 2 FSAs can occur with either method. Panel A includes: a TPS-based kriging contour which intersected 9 separate FSAs in Toronto, and panel B shows a power-based kriging contour which interpolated ≥ 10 cases over 6 unique FSAs in London.

- Section 2 – All localized hotspots (identified through kriging analysis) catalogued for each municipality evaluated are provided in this section. Each data file lists the FSA containing the localized hotspot, the size of this hotspot (% by area), the 3-day window in which it was observed, the interpolated and true case values on this date, and whether the FSA was also significant by Gi* analysis (on any of the 3 days analyzed).

Underlying data

Zenodo: Data and Software Archive for ‘Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada’: <http://doi.org/10.5281/zenodo.5585811>.²⁹

This project contains the following underlying data:

- Section 3 – All underlying data from the geostatistical tests performed in this study are provided in this section. This includes the output from Ontario-wide FSA-level Gi* and Cluster and Outlier analyses, and PC-level Cluster and Outlier, Spatial Autocorrelation, and kriging analysis of 6 municipal regions. Kriging analysis output files are also provided for 7 municipalities immediately surrounding Toronto. This section also contains data files from our analyses of stratified case data (by age, gender, and at-risk condition). All coordinates presented in these data files are given in “PCS_Lambert_Conformal_Conic” format. Case values between 1-5 were masked (appear as “NA”).
- Section 4 – Sets of image files which map the results of our geostatistical analyses onto a map of Ontario or within the municipalities evaluated are provided. This includes: kriging analysis (PC-level), Local Moran's I Cluster and Outlier analysis (FSA and PC-level), normal and space-time Gi* analysis, and all images for all analyses performed on stratified data (by age, gender, and at-risk condition).

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

Software availability

The Zenodo repository (<http://doi.org/10.5281/zenodo.5585811>²⁹) also provides all program files pertaining to the (Geostatistical analysis software package to be used in ArcGIS), as well as all other scripts described in this manuscript. This code is available under the terms of the GNU General Public License v3.0.

Acknowledgements

The authors acknowledge advice from Dan Lizotte at Western University. ICES personnel: Refik Saskin, Minnie Ho, Diana An, and Ryan Godinho, Kamil Malikov and his staff at MOH, and Regine Lecocq at the Department of National Defence IDEaS program are recognized for their support and facilitation of this project.

References

1. Anderson RM, Heesterbeek H, Klinkenberg D, *et al.*: **How will country-based mitigation measures influence the course of the COVID-19 epidemic?** *Lancet*. 2020; **395**: 931–934.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Barrett K, *et al.*: **Estimation of COVID-19-induced depletion of hospital resources in Ontario, Canada.** *CMAJ Can. Med. Assoc. J.* 2020; **192**: E640–E646.
[Publisher Full Text](#)
3. Gardner BJ, Kilpatrick AM: **Contact tracing efficiency, transmission heterogeneity, and accelerating COVID-19 epidemics.** *PLoS Comput. Biol.* 2021; **17**: e1009122.
[PubMed Abstract](#) | [Publisher Full Text](#)
4. Detsky AS, Bogoch II: **COVID-19 in Canada: Experience and Response.** *JAMA*. 2020; **324**: 743–744.
[Publisher Full Text](#)
5. Government of Canada, O. of the A. G. of C: **Report 10 –Securing Personal Protective Equipment and Medical Devices.** 2021.
[Reference Source](#)
6. Liu J, *et al.*: **Community Transmission of Severe Acute Respiratory Syndrome Coronavirus 2, Shenzhen, China, 2020.** *Emerg. Infect. Dis.* 2020; **26**: 1320–1323.
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Carrat F, Valleron AJ: **Epidemiologic Mapping using the “Kriging” Method: Application to an Influenza-like Epidemic in France.** *Am. J. Epidemiol.* 1992; **135**: 1293–1300.
[PubMed Abstract](#) | [Publisher Full Text](#)
8. Desjardins MR, Whiteman A, Casas I, *et al.*: **Space-time clusters and co-occurrence of chikungunya and dengue fever in Colombia from 2015 to 2016.** *Acta Trop.* 2018; **185**: 77–85.
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Ali M, *et al.*: **Application of Poisson kriging to the mapping of cholera and dysentery incidence in an endemic area of Bangladesh.** *Int. J. Health Geogr.* 2006; **5**: 45.
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Greene SK, Peterson ER, Kapell D, *et al.*: **Daily Reportable Disease Spatiotemporal Cluster Detection, New York City, New York, USA, 2014–2015.** *Emerg. Infect. Dis.* 2016; **22**: 1808–1812.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Azevedo L, Pereira MJ, Ribeiro MC, *et al.*: **Geostatistical COVID-19 infection risk maps for Portugal.** *Int. J. Health Geogr.* 2020; **19**: 25.
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Briz-Redón Á, Serrano-Aroca Á: **A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain.** *Sci. Total Environ.* 2020; **728**: 138811.
[PubMed Abstract](#) | [Publisher Full Text](#)

13. Desjardins MR, Hohl A, Delmelle EM: **Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters.** *Appl. Geogr. Sevenoaks Engl.* 2020; **118**: 102202. [PubMed Abstract](#) | [Publisher Full Text](#)
14. Hohl A, Delmelle EM, Desjardins MR, et al.: **Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States.** *Spat. Spatio-Temporal Epidemiol.* 2020; **34**: 100354. [PubMed Abstract](#) | [Publisher Full Text](#)
15. Jia JS, et al.: **Population flow drives spatio-temporal distribution of COVID-19 in China.** *Nature.* 2020; **582**: 389–394. [PubMed Abstract](#) | [Publisher Full Text](#)
16. Kim S, Castro MC: **Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020).** *Int. J. Infect. Dis.* 2020; **98**: 328–333. [PubMed Abstract](#) | [Publisher Full Text](#)
17. Onovo AA, et al.: **Using Supervised Machine Learning and Empirical Bayesian Kriging to reveal Correlates and Patterns of COVID-19 Disease outbreak in sub-Saharan Africa: Exploratory Data Analysis.** 2020.2020.04.27.20082057. [PubMed Abstract](#)
18. Földváry L: **Geostatistical Investigations on the Spread of COVID-19.** *Int. J. Geoinformatics.* 2021: 75–84. [Publisher Full Text](#)
19. Fatima M, O'Keefe KJ, Wei W, et al.: **Geospatial Analysis of COVID-19: A Scoping Review.** *Int. J. Environ. Res. Public Health.* 2021; **18**: 2336. [PubMed Abstract](#) | [Publisher Full Text](#)
20. Kulldorff M, Information Management Services Inc: *SaTScanTM v8.0: Software for the spatial and space-time scan statistics.* 2009.
21. Anselin L, Syabri I, Kho Y: **GeoDa: An Introduction to Spatial Data Analysis.** *Geogr. Anal.* 2006; **38**: 5–22. [Publisher Full Text](#)
22. Rogan PK, Mucaki EJ: **Geostatistical Analysis of SARS-CoV-2 Positive Cases in the United States.** 2020. [Reference Source](#)
23. Getis A, Ord JK: **The Analysis of Spatial Association by Use of Distance Statistics.** *Geogr. Anal.* 1992; **24**: 189–206. [Publisher Full Text](#)
24. Ord JK, Getis A: **Local Spatial Autocorrelation Statistics: Distributional Issues and an Application.** *Geogr. Anal.* 1995; **27**: 286–306. [Publisher Full Text](#)
25. Anselin L: **Local Indicators of Spatial Association—LISA.** *Geogr. Anal.* 1995; **27**: 93–115. [Publisher Full Text](#)
26. Boots BN, Getis A: *Point Pattern Analysis.* Newbury Park, CA: Sage Publications; 1998.
27. Gribov A, Krivoruchko K: **Empirical Bayesian kriging implementation and usage.** *Sci. Total Environ.* 2020; **722**: 137290. [PubMed Abstract](#) | [Publisher Full Text](#)
28. Krivoruchko K: **Empirical Bayesian Kriging.** 2012. [Reference Source](#)
29. Mucaki EJ, Shirley BC, Rogan PK: **Zenodo Archive for 'Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada'.** 2021. [Publisher Full Text](#)
30. Rogan PK, et al.: **Meeting radiation dosimetry capacity requirements of population-scale exposures by geostatistical sampling.** *PLoS One.* 2020; **15**: e0232008. [PubMed Abstract](#) | [Publisher Full Text](#)
31. Rocca R, Westoll N: **Coronavirus: 7 dead, 37 cases at Humber Heights retirement home in Toronto [Internet].** *Global News;* 2020 [cited 2022 June]. [Reference Source](#)
32. Web Staff: **List of long-term care and retirement homes impacted by COVID-19 in GTA [Internet].** *CTV News Toronto;* 2020 [cited 2022 June]. [Reference Source](#)
33. Middlesex-London Health Unit: **Outbreak Declared as More Than 40 Cases Reported at Apartment Complex — Middlesex-London Health Unit.** 2020. [Reference Source](#)
34. Saeed S, et al.: **SARS-CoV-2 seroprevalence among blood donors after the first COVID-19 wave in Canada.** *Transfusion (Paris).* 2021; **61**: 862–872. [PubMed Abstract](#) | [Publisher Full Text](#)
35. Detsky AS, Bogoch II: **COVID-19 in Canada: Experience and Response to Waves 2 and 3.** *JAMA.* 2021; **326**: 1145–1146. [Publisher Full Text](#)
36. Garry M, Hope L, Zajac R, et al.: **Contact Tracing: A Memory Task With Consequences for Public Health.** *Perspect. Psychol. Sci. J. Assoc. Psychol. Sci.* 2021; **16**: 175–187. [PubMed Abstract](#) | [Publisher Full Text](#)
37. COVID-19 National Emergency Response Center, Epidemiology & Case Management Team, Korea Centers for Disease Control & Prevention: **Contact Transmission of COVID-19 in South Korea: Novel Investigation Techniques for Tracing Contacts.** *Osong. Public Health Res. Perspect.* 2020; **11**: 60–63. [PubMed Abstract](#) | [Publisher Full Text](#)
38. Kulldorff M: **Prospective time periodic geographical disease surveillance using a scan statistic.** *J. R. Stat. Soc. Ser. A Stat. Soc.* 2001; **164**: 61–72. [Publisher Full Text](#)
39. Statistics Canada: **Dissemination area: Detailed definition.** 2018. [Reference Source](#)
40. Statistics Canada: **Dissemination Area Reference Maps, by Census Tracts, for Census Metropolitan Areas and Census Agglomerations.** 2021. [Reference Source](#)
41. Jones JM, et al.: **Estimated US Infection- and Vaccine-Induced SARS-CoV-2 Seroprevalence Based on Blood Donations, July 2020–May 2021.** *JAMA.* 2021; **326**: 1400–1409. [PubMed Abstract](#) | [Publisher Full Text](#)
42. Byambasuren O, et al.: **Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: Systematic review.** *PLoS One.* 2021; **16**: e0248946. [PubMed Abstract](#) | [Publisher Full Text](#)
43. Grant R, et al.: **SARS-CoV-2 population-based seroprevalence studies in Europe: a scoping review.** *BMJ Open.* 2021; **11**: e045425. [PubMed Abstract](#) | [Publisher Full Text](#)
44. Charlton CL, 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**: e0029121.
45. Ontario Ministry of Health: **COVID-19: Guidance for Prioritization of Phase 2 Populations for COVID-19 Vaccination.** 2021. [Reference Source](#)

Open Peer Review

Current Peer Review Status:   

Version 2

Reviewer Report 08 August 2022

<https://doi.org/10.5256/f1000research.136332.r144874>

© 2022 Foldvary L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Lorant Foldvary

Department of Geodesy and Surveying, Associate Professor, Budapest University of Technology and Economics, Budapest, Hungary

Index as is.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: geodesy, gravimetry, geoinformatics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 24 May 2022

<https://doi.org/10.5256/f1000research.79823.r129171>

© 2022 Gayawan E. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Ezra Gayawan

Department of Statistics, Federal University of Technology, Akure, Nigeria

1. In the introduction, the authors, arguing against some geostatistical software, stated that "... however, these were not used in the present study because of their requirement to include disease background levels, which was not relevant for COVID-19." I wonder what

they really meant by “disease background levels” and how it restricts the use of the mentioned software including R for making geospatial analysis of COVID-19 data. As a fact, there are several of such analyses out there that used different methods and software particularly R. A Poisson geoaddivitive model, for instance, can be fitted to case count data on COVID-19 over time, which will generate some beautiful spatial maps without the need for any background disease as claimed.

2. The relationship between FSA and PC is not clear in the manuscript. This may be confusing for some readers, especially those not familiar with Canada. More detail is required in this aspect by clearly mentioning the benefits of why both FSA and PC were chosen rather than either one of them. An illustrative graph showing the FSA and PC regions would be helpful.
3. It is difficult to understand the illustration in Figure 1. A mathematical notation/formula should be used to support the illustration, clearly defining all the parameters involved in the formula.
4. The authors mention that "the software can be easily modified to accommodate other Canadian provinces and, with additional effort, other countries." However, there is no guide or link to a description of how the developed tool could be used and probably extended to other provinces or countries.
5. More details are required for the multinomial logistic model considered. Explicit details such as the response variable and the corresponding covariates should be clearly stated. If possible, the regression model should include stratification variables as covariates, such as categorized age, and "at-risk group" for example. The conclusion drawn from such analysis can be discussed in the Discussion session. It is not clear whether the developed toolbox was used for the estimation of the multinomial regression model or the Directional Acyclic Graph Network.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 21 Jul 2022

Peter Rogan, University of Western Ontario, London, Canada

Comment #1. In the introduction, the authors, arguing against some geostatistical software, stated that "... however, these were not used in the present study because of their requirement to include disease background levels, which was not relevant for COVID-19." I wonder what they really meant by "disease background levels" and how it restricts the use of the mentioned software including R for making geospatial analysis of COVID-19 data. As a fact, there are several of such analyses out there that used different methods and software particularly R. A Poisson geoaddivitive model, for instance, can be fitted to case count data on COVID-19 over time, which will generate some beautiful spatial maps without the need for any background disease as claimed.

Response:

We appreciate this comment, in particular about the Poisson geoaddivitive model. Depending on which R script is used, analysis with R may or may not require background levels to be input. We have removed the citation to R, as suggested by another reviewer.

Comment #2. The relationship between FSA and PC is not clear in the manuscript. This may be confusing for some readers, especially those not familiar with Canada. More detail is required in this aspect by clearly mentioning the benefits of why both FSA and PC were chosen rather than either one of them. An illustrative graph showing the FSA and PC regions would be helpful.

Response:

To address the lack of familiarity on Canadian Forward Sortation Areas and Postal Codes, the sentence which introduces FSAs and PCs in the Introduction has been expanded. We also relate FSAs conceptually to U.S. zip codes: "distributed at the level of individual forward sortation areas (FSA) and postal codes (PCs; a smaller geographic segment within an FSA). Area-to-area analyses were performed on COVID-19 case data at the provincial level by FSA, a geographical unit similar to zip codes in the United States which can encompass hundreds of PCs." We have also added the following to Table 1: map images illustrating the PC boundaries of the municipalities of interest, a general Ontario map indicating the locations of each of these regions, the total number of FSAs and PCs in each region, and their total area of each.

COVID-19 hotspot communities were often identified within FSAs. Prioritization for COVID-19 vaccinations (when supply was initially limited) was informed using hotspots identified at this level of granularity (Ontario Ministry of Health: COVID-19: Guidance for Prioritization of Phase 2 Populations for COVID-19 Vaccination. 2021). For this reason, we evaluated COVID-19 hotspots at this level of resolution. However, since the distributions of COVID-19

infections was often non-uniform in many FSAs (as we describe), this level of granularity was insufficient to fully elucidate its local distributions, and analysis of PCs was necessary.

This possibility is mentioned in the Introduction (second paragraph): "Available case data analyzed has been typically aggregated at the municipal or county level, however higher resolution spatial analyses has been limited as finer granularity for the distribution of COVID-19 cases has often been lacking."

Subsequently, we generally describe FSAs and PCs. However, it was not obvious that PC-level COVID-19 case data produced a high level of geographic resolution that had been previously unavailable. We have therefore added the following at the end of this paragraph: "Therefore, by evaluating COVID-19 case data at the PC level, we will be able to evaluate disease distributions at a high level of granularity."

Comment #3. It is difficult to understand the illustration in Figure 1. A mathematical notation/formula should be used to support the illustration, clearly defining all the parameters involved in the formula.

Response:

Formulation of the Getis-Ord local statistic is now presented in the third paragraph of the Introduction, with descriptions of each parameter within each equation, as requested.. We have also expanded the Figure 1 legend to describe how the weight function of these formulas differ depending on the spatial relationship being used. Weights of two features are equal in Fixed and K Nearest Neighbors, given that the two points are within a given distance metric (e.g. within 1km for Fixed, or is one of the K closest neighbors to the feature being evaluated for KNN; otherwise, $w_{i,j} = 0$). The inverse distance spatial relationship, however, computes weight as the inverse Euclidean distance between two features being evaluated. This is calculated as $w_{i,j} = d_{i,j}^{-\beta}$ where $d_{i,j}$ is the Euclidean distance between two points and β is the inverse distance power variable ($\beta = 1$ in this study). This is now provided in the legend of Figure 1.

Comment #4. The authors mention that "the software can be easily modified to accommodate other Canadian provinces and, with additional effort, other countries." However, there is no guide or link to a description of how the developed tool could be used and probably extended to other provinces or countries.

Response:

We have added a section to the Geostatistical Epidemiology Toolbox user manual (in the Zenodo archive) which details the steps required for evaluation other Canadian Provinces or other countries. We have modified the original statement in the main manuscript to direct readers to the Geostatistical Epidemiology Toolbox user manual (which now contains a section on modifying the toolbox for other countries and Canadian provinces): "While the toolbox was designed to evaluate COVID-19 case data in Ontario, the software can be easily modified to accommodate other Canadian provinces and, with additional effort, other countries (as described in the Geostatistical Epidemiology Toolbox user manual ²⁹)."

Using the toolbox to evaluate disease geostatistics in other Provinces is a relatively straightforward, requiring only alteration of a few lines of the main Python toolbox code ("GeostatisticalEpidemiologyToolbox_CytoGnomix.pyt"; see Toolbox guide sub-section "Evaluating Other Canadian Provinces" within the Zenodo archive file "Software.GeostatisticalEpidemiologyToolbox.zip").

Modifying the Toolbox to evaluate hotspots and interpolation in other countries would be more involved ("Evaluating Other Countries" section of the manual). The structure of the shape files that define the regions equivalent to an FSA or a postal code in other countries will likely follow a different naming convention (for example, county subdivisions and zip codes in the United States). We now indicate in the Toolbox user manual which column names used by the Toolbox are most likely to require modification ('PROV' and 'PRNAME'). Since it is not possible to generalize the structure of every shape file defining geographic elements of other countries, potential users should consider contacting us for assistance.

Comment #5. More details are required for the multinomial logistic model considered. Explicit details such as the response variable and the corresponding covariates should be clearly stated. If possible, the regression model should include stratification variables as covariates, such as categorized age, and "at-risk group" for example. The conclusion drawn from such analysis can be discussed in the Discussion session. It is not clear whether the developed toolbox was used for the estimation of the multinomial regression model or the Directional Acyclic Graph Network.

Response:

Postal codes were identified as a high-case cluster (by Cluster and Outlier analysis) over a period of >3 consecutive days (i.e., PC "streaks"), and established which of these streaks occurred in two neighboring PCs in close succession. We noticed that the number of cases reported in each paired streak, the interval of time separating each streak, and the physical distance between the PCs might be related. To establish this, we derived multinomial logistic regression models based on these factors, in which the total number of cases within the paired streaks was defined as the response variable. The relationship between the number of cases between streaks of pairs of PCs in close proximity and the time separating the occurrence of those two streaks was significant.

We now provide the R code used to derive these multinomial logistic regression models ("Streak_Analysis_using_Multinomial_Logistic_Regression_Models.Rmd") as a new section within the Zenodo archive ("Software.Additional_Programs_for_Cluster_Outlier_Streak_Identification_and_Pairing.zip"). The paragraph describing the multinomial logistic regression model (third paragraph of the Results has been updated to reflect this change: "The distribution of and relationships between clustered PCs with persistent COVID-19 cases"). This section has been revised to describe the software used to derive the model (including the name of the script and a citation of the Zenodo archive), and explicitly states which of the three variables was used as the response variable: The input data are also included in the archive.

The description now states: "To assess the significance of these results, a multinomial logistic regression model was derived in R

("Streak_Analysis_using_Multinomial_Logistic_Regression_Models.Rmd" which uses the *multinom* function of the *nnet* package; available in Zenodo archive ²⁹), evaluating (1) the total number of cases in each pair of PCs over the course of both streaks (classified as either <25, between 25 and 49, or ≥ 50 cases; the response variable of the regression), (2) the duration ..."

This analysis was not performed for stratified case data because it considerably decreases the number of high-case clusters detected per class (Table 1). This would result in far fewer and shorter PC "streaks", and the lower statistical power would make analysis would less meaningful.

The Geostatistical Epidemiology Toolbox was developed to perform and visualize geostatistical analyses through the ArcMap graphical user interface using ArcPy (which enables Python programs to access ArcGIS's geoprocessing tools and extensions). The directional acyclic graph networks, as well as the regression, were performed outside of this environment (using MatLab and R, respectively). Furthermore, these techniques examined a subset of our results. Thus, the toolbox does not and can not perform either task.

Competing Interests: PKR cofounded and BCS and EJM are employees of CytoGnomix Inc.

Reviewer Report 08 April 2022

<https://doi.org/10.5256/f1000research.79823.r129170>

© 2022 Tedijanto C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Christine Tedijanto 

Francis I. Proctor Foundation, University of California, San Francisco, San Francisco, CA, USA

The authors present a thorough geospatial analysis of COVID-19 cases in Ontario, Canada. In this work, they explore a number of different global and local clustering methods in order to demonstrate (1) the importance of higher resolution data to monitor and control COVID-19 and (2) potential transmission between hotspots. They find alignment between the different clustering approaches and evidence of neighboring areas experiencing high case burdens at similar times, implying transmission between communities. In general, I thought the analyses were interesting, but the work could be made more impactful by streamlining the text and focusing on a clearer through line.

1. For readers who may not be familiar with Ontario, FSAs, and PCs, it may be helpful to include additional background material such as (1) a map of Ontario showing the location of the six municipalities (and FSAs, if possible) of interest, (2) the number of FSAs/PCs and their approximate population sizes, and (3) overall case counts in your municipalities of interest in each of the waves and interwave periods.

2. The motivation behind the subgroup analyses (sex, age, health conditions) was not clear to me. It seems like clusters within each subgroup would tend to appear simply where the highest number of individuals in each subgroup reside, and I'm not sure we would expect the transmission to occur along subgroup lines (for example, we might expect younger family members to transmit to older ones rather than transmission occurring between older individuals). Please explain the reasoning and implications of this analysis further; it may be more suitable in the Extended Data as it seems less relevant to the overall conclusions of the manuscript.
3. I found it difficult to distinguish between the space-time G_i^* and the "streak" analyses. How do they relate to one another? Is there any additional information gained from the space-time G_i^* analysis?
4. Please describe in more detail the approach that was used to evaluate the kriging and semivariogram models. Why do you think the models performed differently in Toronto compared to other cities? In addition, please clarify – were PCs with 0 cases also excluded as neighbors (i.e., were neighbors in the kriging analysis restricted to neighbors with cases ≥ 1)?
5. Is there a way to incorporate uncertainty in the kriging analysis? I would also be cautious about using the term "significant" to describe areas with interpolated counts >5 , as this may be mistaken for statistical significance.
6. I thought the graphs in Figure 8 were a bit challenging to interpret. Visually, it may help to arrange the points similarly to how they are geographically located and/or to change the length of the edges in proportion to the duration separating adjacent pairs (the current label).
7. I would like further discussion about how the authors see these analyses being applied, especially in real-time. For example, the first paragraph of the Discussion says that this method "enabled early detection of disease clusters" – please provide additional evidence of this from the analyses. Additionally, based on the authors' experience of applying many different methods, are there specific ones that the authors would recommend for real-time use (or public health practice in general)?

Minor comments:

1. In the second paragraph of the introduction, the authors include R in a list of tools that were not used because they require background levels, but R is very general software with packages that may be used without disease background levels. I would consider removing it from this list.
2. In my opinion, some sections of the methods may be moved to the Extended Data (e.g. paragraphs beginning "Additional ArcPy scripts..." and "These processes were then integrated...") for conciseness. Tables 1 and 2 may also be better suited to the Extended Data.
3. Please clarify the paragraph about FSAs with case counts <5 – was it possible to analyze

these FSAs without masking, just within a separate computing environment? It was not clear to me whether they were masked or not in the final analysis.

4. In the fourth paragraph of the results, why were comparisons between kriging and the cluster/outlier methods restricted to March 26 to December 28, 2020?
5. In Figure 3B, what do the blue crosses represent? I thought they represented the observed data, in which case I would have expected the same blue crosses in the top and bottom panel. Please clarify.
6. Please provide additional information on the “relative density” metric in Figure 4. Is there a reason this particular area and time were highlighted in this figure? It may be interesting to show a hotspot that occurred during a wave rather than in a time window with low disease burden.
7. If possible, please combine all resources into one repository.
8. I did not see underlying data in any of the linked repositories. If the underlying data cannot be publicly shared, please provide a statement describing how the data may be accessed.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious disease epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 21 Jul 2022

Peter Rogan, University of Western Ontario, London, Canada

Major Comment #1:

...it may be helpful to include additional background material such as (1) a map of Ontario showing the location of the six municipalities (and FSAs, if possible) of interest, (2) the number of FSAs/PCs and their approximate population sizes, and (3) overall case counts in your municipalities of interest in each of the waves and interwave periods.

Response:

We agree that including regional statistics and maps of the areas being studied would be beneficial to the manuscript. We have therefore chosen to append this information to Table 1A of the manuscript, as the structure of the table was well suited for it. The first column of Table 1 includes a map of lower Ontario with the location of each region of interest, as well as maps of each region indicating their boundaries. We also include the population, the area (km²) and the number of FSAs and PCs in each region analyzed.

We have also added a new column which gives an estimate of the number of cases present in each region (the sum of cases from individual FSAs associated with the region of interest), divided by waves. These case numbers are estimated, since our contract with ICES has ended, and they have removed our access to actual case data. Thus, any dates in which case data was masked (i.e. the FSA had between 1-5 cases on a particular date) were considered to be a single case by this calculation (which is indicated in the footer of Table 1).

Major Comment #2:

The motivation behind the subgroup analyses (sex, age, health conditions) was not clear to me. It seems like clusters within each subgroup would tend to appear simply where the highest number of individuals in each subgroup reside, and I'm not sure we would expect the transmission to occur along subgroup lines (for example, we might expect younger family members to transmit to older ones rather than transmission occurring between older individuals). Please explain the reasoning and implications of this analysis further; it may be more suitable in the Extended Data as it seems less relevant to the overall conclusions of the manuscript.

Response:

The identification of COVID-19 clusters according to these subgroups was specifically requested by the Ontario Ministry of Health. There was interest in identifying differences in how COVID-19 cases were distributed across our regions of interest. This has merit since clustering of a particular subgroup was often correlated with known events occurring in those locations, e.g. high-case clusters of COVID-19 in predominantly young males in rural regions of Windsor/Essex reflecting outbreaks in migrant farm worker communities at this time). We have not moved this analysis (including Table 1) to the extended data.

The analysis we describe does not imply that transmission is occurring between or within subgroups. Nonetheless, the distribution of cases by subgroup was of particular interest to the Ontario Ministry of Health. It did not reflect the geographic bias in the residence of subgroup members. We have therefore added the following to the end of the fourth

paragraph of Results subsection "Wave 1" (page 17, paragraph 1 without markup): "This stratified analysis itself does not provide insight into the mode of COVID-19 transmission, that is, the development of a hot spot consisting of a particular subgroup does not indicate that viral transmission occurred exclusively between members of the particular subgroup."

Major Comment #3:

I found it difficult to distinguish between the space-time G_i^* and the "streak" analyses. How do they relate to one another? Is there any additional information gained from the space-time G_i^* analysis?

Response:

Space-time G_i^* analysis and Cluster and Outlier 'streaks' are indeed two separate analyses. Results using space-time G_i^* analysis (built-in to ArcGIS) is provided at the start of the Results. We found that space-time G_i^* analysis was generally uninformative, as the analysis called nearly all regions with high numbers of cases (>10 at the FSA level) as statistically significant (>90% in most waves). Therefore, we did not pursue it further. However, this was not explicitly stated, so we have added the following statement to the aforementioned Results paragraph: "As the vast majority of regions were flagged by this approach (>90%), we did not gain any insight from this space-time analysis."

We then investigated temporal trends using Cluster and Outlier analyses of "streaks" of consecutive days that particular locates were deemed hotspots. analysis This is a form of "space-time" analysis, but is completely independent of any functional tests built into ArcGIS. We aren't aware of any prior description of this method. We have modified the text to clarify that this was completely separate from space-time G_i^* analysis (page 14, paragraph 2): "Instead, we evaluated temporal trends in the distribution of COVID-19 cases and identified potential transmission using Cluster and Outlier analysis of PCs that persisted across a consecutive series of dates (referred to as high-case cluster 'streaks'; described later in the Results section)."

Major Comment #4:

1. Please describe in more detail the approach that was used to evaluate the kriging and semivariogram models.
2. Why do you think the models performed differently in Toronto compared to other cities?
3. In addition, please clarify – were PCs with 0 cases also excluded as neighbors (i.e., were neighbors in the kriging analysis restricted to neighbors with cases ≥ 1)?

Response:

(A) In the second paragraph of Methods subsection "Empirical Bayesian Kriging analysis", we describe our reasoning for prioritizing the Power and TPS semivariograms for low/medium and high case density regions, respectively. Some detail was excluded due to the overall length of the manuscript. To further elaborate, we chose to evaluate all available kriging semivariogram options in ArcGIS in moderate (London, ON) and high (Toronto, ON) density regions across several dates in Waves 1 and 2 of the COVID-19 pandemic. Each semivariogram was tested under various setting conditions (sector type [1 or 4 sectors],

search radius [the optimal radius determined by ArcGIS software, that value halved and that value doubled], neighborhood type [smooth or standard] and the number of neighbors considered for each location [5, 10, 15]). The kriging contours generated were converted to polygons in order to perform an 'intersect' in ArcGIS between these contours and PCs, and from this intersection, the values interpolated by kriging were compared to true case counts.

The Power semivariogram generally modeled counts better in low and medium-density municipal regions. Although TPS often performed well, it produced artifact contours in regions without cases, as described in the Methods ("TPS, in some instances, generated kriging artifacts that did not correlate with known hotspots"). However, the Power semivariogram model was sometimes inferior to TPS, with PCs having high case counts were being missed. Thus, we used TPS in Toronto, considering only 5 neighbors which minimized the kriging artifacts without sacrificing accuracy. This observation was supported by a chi-square test, whereby kriging interpolation with the TPS semivariogram utilizing ≥ 10 neighbors was frequently significantly different than ground truth (e.g. 72% of dates in Wave 2 were significantly different [$p < 0.05$] when considering 10 neighbors) This did not happen when considering 5 neighbors ($p > 0.05$ for all dates in Waves 1 and 2). This analysis justifies our selection for the number of neighbors parameter when performing kriging, and this analysis is now briefly described in the manuscript: "Using 5 neighboring locations for interpolation minimized these artifacts without affecting the contour map. However, with 10 neighbors, interpolation accuracy was affected according to a chi-square test (e.g., $p < 0.05$ for the 103 days in Wave 2)."

With similar tests, we found that the default kriging parameters (other than neighbor count) were optimal (or lead to equivalent results) regardless of semivariogram used. Multiple kriging parameters are used to define a search neighborhood, the window which delineates which measured locations will and will not be included to predict the value of a region. One such parameter divides this search window into multiple sectors. However, utilizing four sectors led to poorer interpolation of true case counts by chi-square test (significantly different [$p < 0.05$] for 37% of dates in Wave 1 and 82% of dates in Wave 2). Thus, a single sector was used. These parameters are now mentioned in the Methods section: "All other EBK parameters used were defaulted to those provided by the ArcGIS Advanced Geostatistics toolbox (Sector Type: 1 sector, Neighborhood Type: Standard, Search Radius: 1)."

The use of TPS in Toronto is now justified. Postal codes with > 5 cases are interpolated as such (≥ 5 -10 cases) over twice as often with TPS relative to Power (991 vs 414, respectively). Furthermore, the root mean square error of case count to interpolated case range of PCs with > 5 cases is higher using the Power semivariogram (6.4 with Power, 5.4 with TPS), indicating that TPS outperforms Power at interpolating case counts. A chi-square test comparing the interpolation of daily Toronto case counts using the TPS and Power semivariograms similarly showed that while use of the Power semivariogram was never significantly different from true case counts, the TPS semivariogram better matched ground truth (p -value TPS $>$ p -value Power for 85% and 78% of dates in Waves 1 and 2, respectively).

Similarity of Interpolation to True Cases by Chi-Square Test: Kriging utilizing the Power and TPS Semivariogram

Days with Greater Similarity [$>$ p-val] to actual cases Cases)			Days (≥ 1 PC with ≥ 6
$p_{TPS} > p_{POWER}$	$p_{POWER} > p_{TPS}$	$p_{TPS} = p_{POWER}$	
<i>Wave 1:</i>			
60	2	9	71
<i>Wave 2:</i>			
112	11	20	143

All six municipalities were evaluated with both the Power and TPS semivariograms across all dates evaluated, the output of which is available in the Zenodo archive.

(B) We believe that semivariogram performance in Toronto was due to its far higher population frequency and density relative to the other areas analyzed. Toronto has the highest density of postal codes (75.3 PCs per km²) compared to the other regions investigated (Hamilton: 13.0 PCs per km²; Kitchener/Waterloo: 42.0 PCs per km²; London: 23.9 PCs per km²; Ottawa: 7.9 PCs per km²; Windsor/Essex: 6.0 PCs per km²). While the PC density of Kitchener/Waterloo is somewhat comparable to Toronto, the population density of Kitchener/Waterloo is 2.5x smaller. Furthermore, the number of COVID-19 cases experienced between the two regions were far different, with Toronto having nearly 15-fold more cases compared to Kitchener/Waterloo (see Table 1).

Toronto exhibits far more PCs with cases than the other regions tested on most dates. The majority of these PCs have cases between 1-5 (<1% of PCs in Toronto had >5 cases across all dates analyzed). This pushes the interpolated case counts downwards (similar to how zero-case PCs depressed kriging values; see next section). Since the “Power” semivariogram is a parabolic mathematical representation of cases in the region, a large amount of low-case PCs may ‘smooth’ any contour peaks in the region. Since TPS semivariograms attempt to pass through binned data points exactly, these peaks are not lost. Thus, the true case count of high case areas are better represented within the kriging contour.

Kriging analysis was performed on COVID-19 case data that did not include PCs with zero cases. In the first paragraph of Methods subsection “Empirical Bayesian Kriging analysis”, we state: “Only PCs with ≥ 1 case were included in this analysis, as locations without confirmed cases were found to severely depress kriging signals.”

PCs with zero case counts are more frequent, even on dates where PCs with cases were at its maximum (e.g. from April 12 to the 14, 2021, 2377 PCs in Toronto had ≥ 1 case, which is only 5% of the total number of PCs in this region [2.4 to 3.8% for the other 5 regions of interest). These PCs were so pervasive that they depressed the interpolated values of actual cases in PCs with cases. The case count interpolated by kriging was never representative of (i.e. biased against) actual case counts when zero case PCs were included in our testing. We modified the relevant (above) to further explain this: “Only PCs with ≥ 1 case were included in this analysis, as locations without confirmed cases greatly outnumbered PCs with cases ($\geq 95\%$ of PCs on any given date evaluated) leading to the severe depression of kriging signals.”

Major Comment #5:
Is there a way to incorporate uncertainty in the kriging analysis? I would also be cautious

about using the term “significant” to describe areas with interpolated counts >5, as this may be mistaken for statistical significance.

Response:

The ArcGIS software is capable of determining uncertainty when performing Kriging (<https://desktop.arcgis.com/en/arcmap/10.5/extensions/geostatistical-analyst/kriging-in-geostatistical-analyst.htm>) with standard error maps (<https://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/using-ordinary-kriging-to-create-a-prediction-standard-error-map.htm>). We no longer have access to the ICES computer system to generate such plots. Furthermore, The ‘layer’ files are linked to the input used to generate the kriging layer which is no longer available to us. The main text has been edited to remove instances of the term “significant” for regions interpolated by kriging consisting of >5 case counts. This was changed to “region of interest”, “localized hotspot”, or avoided. The name of Supplemental Figure S3 was also changed.

Major Comment #6:

...graphs in Figure 8 were a bit challenging to interpret. Visually, it may help to arrange the points similarly to how they are geographically located and/or to change the length of the edges in proportion to the duration separating adjacent pairs (the current label).

Response:

The directional network graphs in Figure 8 were created with the MatLab *digraph* function. *digraph* does not provide any option to incorporate the recommended changes. The order of data points and edge length are automatically specified. Our only option was to label each path with distance, number of days separating pairs of neighboring streaks, PC identifiers, etc. The maps accompanying the graphs provide a geographic context for each network. Figure 8 and Extended Data Tables S9 and S10 remain unchanged.

Major Comment #7:

I would like further discussion about how the authors see these analyses being applied, especially in real-time. For example, the first paragraph of the Discussion says that this method “enabled early detection of disease clusters” – please provide additional evidence of this from the analyses. Additionally, based on the authors’ experience of applying many different methods, are there specific ones that the authors would recommend for real-time use (or public health practice in general)?

Response:

We provide an example within which the FSA N5Y in London Ontario was identified as a region of interest (>5 cases interpolated) by kriging, one week before it was publicized (Figure 6). To provide further evidence of early detection, additional examples (in a new

supplementary figure S3A and S3B) of identified localized hotspots prior to public announcement are described: “Localized hotspots within the FSA M9P were observed as early as March 31st, 2020 however the FSA was not flagged by Gi* analysis until April 6th, 2020 (Extended data, Section 1 - Figure S3A). The localized hotspot covers the Village of Humber Heights Retirement Home in Toronto, which was reported to have a COVID-19 outbreak on April 12th, 13 days after the localized hotspot was first detected by kriging.” “For example, localized hotspots in FSA M9V were identified sporadically, and the interpolated value does not always match total cases (Extended data, Section 1 - Figure S3B). The localized hot spots cover the PC M9V 5B5 starting from April 4th. Outbreaks in the long-term care home within this PC (Humber Valley Terrace long-term care home in Toronto) were reported on April 16th, 2020. This FSA was not flagged by Gi* analysis until April 19th (Extended data, Section 1 - Figure S3B).”

There were several instances where kriging identified a localized hotspot prior to public declaration within the same region. These was not found by either Gi* analysis or Cluster and Outlier analysis. While kriging analysis requires more time than these other statistics to generate results (minutes vs seconds per hotspot), software automation on high-powered systems at provincial and municipal scale improved performance of these processes. In April 2021, the Ontario Ministry of Health requested that we perform a full geostatistical analysis of COVID-19 in four municipal regions [London, Ottawa, Toronto and Windsor/Essex] from January and March 2021. We returned the results, including an evaluation of localized hotspots with kriging, within 5 days (see Discussion, paragraph 7). The “Streak” analysis method (based on results from the Cluster and Outlier analysis tool) was only used to trace disease spread; this was a retroactive analysis to identify neighboring, persistent, and overlapping COVID-19 clusters. We believe that additional work is required to prove the validity of use of this approach pro-actively.

Early hot spot detection is predicated on the availability of current and complete case data. Except for holiday periods, COVID-19 case data was preprocessed and updated on a weekly basis at ICES. As previously published, the number of cases on the final 3 days of the prior dataset was updated and usually increased in the subsequent release (additional testing data was added to those dates), especially in rural regions where testing was subject to adjustment. We have therefore added the following to the Discussion: “... and enabled early detection of disease clusters, which could be exploited to prevent or limit further spread at or close to these hotspots (provided complete and up-to-date COVID-19 case data is made available in a timely manner).”

Minor Comment #1:

In the second paragraph of the introduction, the authors include R in a list of tools that were not used because they require background levels, but R is very general software with packages that may be used without disease background levels. I would consider removing it from this list.

Response:

Geostatistical epidemiology has usually required pre-existing background levels of disease, which is not relevant at the beginning of the COVID-19 pandemic. R and the majority of available epidemiology packages include background level adjustment. While some R

packages may permit background to be ignored, we have removed R from the software listed.

Minor Comment #2:

In my opinion, some sections of the methods may be moved to the Extended Data (e.g. paragraphs beginning "Additional ArcPy scripts..." and "These processes were then integrated...") for conciseness. Tables 1 and 2 may also be better suited to the Extended Data.

Response:

Table 2 has been moved to the Extended Data, as requested. Table 1 remains, as previously discussed (Reviewer #2, Major Comment #2). The paragraph beginning "These processes were then integrated..." contains our description of the Geostatistical Epidemiology Toolbox, which is important for reproducibility and portability of the methods we describe. We highlight this in the paper, and F1000Research requires that the software be available as a condition of publication. We have reduced its length rather than removed it. The deleted, relevant information can now be found within the User Manual accompanying the toolbox ("UserManual_GeostatisticalEpidemiologyToolbox_CytoGnomix.pdf") The paragraph beginning with "Additional ArcPy scripts..." has been moved to the Zenodo archive ("Documentation_for_Each_Section_of_Zenodo_Archive.docx").

Minor Comment #3:

Please clarify the paragraph about FSAs with case counts <5 – was it possible to analyze these FSAs without masking, just within a separate computing environment? It was not clear to me whether they were masked or not in the final analysis.

Response:

All analyses described using unmasked true case count data within the ICES environment, the only exception being the G_i^* analysis that was performed to assess the impact of masking on the geostatistical results (described in the fourth paragraph of the Methods, where we found that masking significantly affected results: "We concluded that errors introduced by masking count values precluded analysis of masked data and required all work to be carried out within the RAE computing environment.")

We now explicitly state that no other geostatistical analysis described in this study used masked data, with the following statement in the Methods: "All geostatistical analyses presented in this study were performed against unmasked, true case counts." We also modified an earlier statement in the same paragraph regarding ICES stipulation that we mask FSAs/PCs with 1-5 cases to clarify that this practice applied only to data export: "To ensure patient privacy, ICES stipulated that case counts could not be exported from the ICES system unless FSAs and/or PCs exhibiting between 1 and 5 daily COVID-19 cases were masked." The following was also added to the Results (first para.): "Changes in the distribution of COVID-19 infections over time were evaluated through geostatistical analysis of true (unmasked) case counts within both FSAs and PCs."

Minor Comment #4:

In the fourth paragraph of the results, why were comparisons between kriging and the cluster/outlier methods restricted to March 26 to December 28, 2020?

Response:

The comparison between kriging and Cluster and Outlier analysis was part of a feasibility analysis (performed in January 2021) to compare the two methods. Expanding this analysis was unlikely to provide novel insights. We prioritized our limited access time to the data and computing environment to perform other analyses.

Minor Comment #5:

In Figure 3B, what do the blue crosses represent? I thought they represented the observed data, in which case I would have expected the same blue crosses in the top and bottom panel. Please clarify.

Response:

Blue crosses are bins of paired locations, grouped by distance and direction (as described in the Figure 3 description). How these bins are grouped by ArcGIS software is indeed different whether the Power or TPS semivariogram is being used. Additionally, how data points are binned can regionally differ, as the values and distances between points vary from location to location.

However, the two semivariograms from Figure 3B were not associated with each other. The original purpose of Figure 3B was to provide examples of "Power" and "TPS" semivariograms derived from this data. While not implicitly stated in the figure legend, we see now that how the inference of a relationship between the two semivariograms could be made. Figure 3B has therefore been updated with semivariograms that are representative of the kriging map presented in Figure 3A. The binned blue crosses for the "Power" and "TPS" semivariograms are similar, although not identical (for the reasons given earlier).

Minor Comment #6:

Please provide additional information on the "relative density" metric in Figure 4. Is there a reason this particular area and time were highlighted in this figure? It may be interesting to show a hotspot that occurred during a wave rather than in a time window with low disease burden.

Response:

Figure 4 represents a region where COVID-19 cases were highly dispersed over a period of time. There are indeed likely other additional regions that could have been used, however the example selected is indeed interesting and worthy of discussion. The time window (September 10th to the 29th, 2020) spans the very end of Interwave 1 (September 24th, 2020), leading into Wave 2, a time where the frequency of cases was rapidly increasing. The frequency of dates in which this FSA was called significant by Gi* analysis was similarly increasing. We therefore do not feel that replacing this figure for another example would be beneficial.

Minor Comment #7:

If possible, please combine all resources into one repository.

Response:

We have combined the data and software Zenodo archives into a single archive (titled 'Data and Software Archive for "Likely community transmission of COVID-19 infections between neighboring, persistent hotspots in Ontario, Canada"). All references to the Zenodo archive within the main text now point to this archive.

Minor Comment #8:

I did not see underlying data in any of the linked repositories. If the underlying data cannot be publicly shared, please provide a statement describing how the data may be accessed.

Response:

The underlying data can be found in the file "Section_3.All-Data_Files.Kriging_GiStar_Local_and_GlobalMorans.2020_2021.zip" in the Zenodo archive (<https://zenodo.org/record/5585812#.YtkdD3bMI2w>). It contains not only output from our geostatistical analyses, but other Ontario COVID-19 case data as well (where cases between 1-5 are masked, as required). This is mentioned in the "Data Availability" section of the manuscript:

Original statement: "Sections 3 and 4 contains the data files containing and the map image files displaying the underlying statistics from the various geostatistical analyses performed in this study."

This has now been clarified to indicate that Section 3 contains the data files, while Section 4 contains map images: "Sections 3 and 4 respectively consist of data files containing and map image files displaying the underlying statistics from the various geostatistical analyses performed in this study."

We also modified the "Extended Data" subsection that describes Section 3, changing the term "output files" to "underlying data" to make the location of the underlying data clearer.

Competing Interests: PKR cofounded and BCS and EJM are employees of CytoGnomix Inc.

Reviewer Report 05 April 2022

<https://doi.org/10.5256/f1000research.79823.r129168>

© 2022 Foldvary L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Lorant Foldvary

Department of Geodesy and Surveying, Associate Professor, Budapest University of Technology and Economics, Budapest, Hungary

Brief Report:

Q: Is the work clearly and accurately presented and does it cite the current literature?

A: The work is clearly and accurately presented, but the citation of the relevant literature is incomplete. For example, on page 3, in the Introduction: "Various geostatistical analyses have been used in the surveillance of COVID-19 spread across the globe"

Relevant examples are missing such as Földváry (2021)¹ but most essential referring to the solution at Johns Hopkins University.²

Q: Is the study design appropriate and does the work have academic merit?

Yes, the study design is fine, the academic merit, and the scientific value is notably high.

Q: Are sufficient details of methods and analysis provided to allow replication by others?

A: The used methods are fine, this is the strength of the study. Since the analysis is performed on commercially available software, the study might be replicated if the used data is accessible.

Q: If applicable, is the statistical analysis and its interpretation appropriate?

A: Yes, the statistical analysis are fine, the interpretation of the results is reasonable!

Q: Are all the source data underlying the results available to ensure full reproducibility?

A: The information on the source data in the manuscript is a bit confusing. Accordingly, improvement of this section is suggested (see the next paragraph). The data and software availability is provided, this content seems to be fine, although the reviewer has no capacity to check on the full reproducibility.

On the source data:

At distinct locations of the text some information on the data is provided, e.g. in the Geostatistical analysis subchapter of the Methods chapter (page 4) says: "COVID-19 cases from March 26th, 2020 to June 25th, 2021 at the FSA (Ontario-wide) and PC-level (for 6 populated areas: Hamilton, Kitchener/Waterloo, London Ottawa, Toronto, and Windsor, including Essex county) were evaluated using ArcGIS Desktop 10.7 [ESRI] with the Geostatistical Analyst extension". It is followed by the first information on the data (page 6, the same subchapter) by indicating that "COVID-19 testing data compiled by ICES contained metadata on tested individuals, such as gender, age, and relevant predicate health conditions", and only later, at the subchapter titled Preparation of Ontario COVID-19 case data from ICES (page 7), we get the introduction of the data source. Somehow the introduction of the data source should be provided more consistently, maybe in a separate chapter before "Methods" titled "Data". The importance of the data for any geostatistical study is essential since the quality and the reliability of the data validates the conclusions.

Q: Are the conclusions drawn adequately supported by the results?

A: Conclusions and interpretations of the results are adequately drawn.

Additional detailed comment:

Page 6, Methods: "COVID-19 testing data compiled by ICES contained metadata on tested

individuals, such as gender, age, and relevant predicate health conditions."
This is the first appearance of the abbreviation ICES, please provide here the full name of the organization!

References

1. Földvary L: Geostatistical Investigations on the Spread of COVID-19. *International Journal of Geoinformatics*. 2021. 75-84 [Publisher Full Text](#)
2. Dong E, Du H, Gardner L: An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*. 2020; **20** (5): 533-534 [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: geodesy, gravimetry, geoinformatics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 21 Jul 2022

Peter Rogan, University of Western Ontario, London, Canada

Comment #1:

The work is clearly and accurately presented, but the citation of the relevant literature is incomplete. For example, on page 3, in the Introduction: "Various geostatistical analyses have been used in the surveillance of COVID-19 spread across the globe" Relevant examples are missing such as Foldvary (2021) but most essential referring to the solution at Johns Hopkins University (Dong et al., 2020).

Response:

We now cite Földváry *et al.* (2021) in the Introduction of our manuscript, as suggested by the reviewer. We have omitted Dong *et al.* (2020), as the web resource described in this study simply displays the location of official COVID-19 counts across the world. This paper focuses on literature that performs geostatistical analyses of the distribution of COVID-19 cases beyond simple counts (e.g. disease spread directionality).

Comment #2:

The information on the source data in the manuscript is a bit confusing. Accordingly, improvement of this section is suggested (see the next paragraph). The data and software availability is provided, this content seems to be fine, although the reviewer has no capacity to check on the full reproducibility.

On the source data:

At distinct locations of the text some information on the data is provided, e.g. in the Geostatistical analysis subchapter of the Methods chapter (page 4) says: "COVID-19 cases from March 26th, 2020 to June 25th, 2021 at the FSA (Ontario-wide) and PC-level (for 6 populated areas: Hamilton, Kitchener/Waterloo, London Ottawa, Toronto, and Windsor, including Essex county) were evaluated using ArcGIS Desktop 10.7 [ESRI] with the Geostatistical Analyst extension". It is followed by the first information on the data (page 6, the same subchapter) by indicating that "COVID-19 testing data compiled by ICES contained metadata on tested individuals, such as gender, age, and relevant predicate health conditions", and only later, at the subchapter titled Preparation of Ontario COVID-19 case data from ICES (page 7), we get the introduction of the data source. Somehow the introduction of the data source should be provided more consistently, maybe in a separate chapter before "Methods" titled "Data". The importance of the data for any geostatistical study is essential since the quality and the reliability of the data validates the conclusions.

Response:

We agree that the description of the data source should come before our description of the geostatistical methods used to analyze it. We have reordered the sections "Preparation of Ontario COVID-19 case data from ICES" and "FSA and PC Boundary Files," which now appear before the "Geostatistical analysis" subsection of the Methods section.

Comment #3:

Page 6, Methods: "COVID-19 testing data compiled by ICES contained metadata on tested individuals, such as gender, age, and relevant predicate health conditions."

This is the first appearance of the abbreviation ICES, please provide here the full name of the organization!

Response:

ICES was previously an abbreviation for the 'Institute for Clinical Evaluative Sciences', but is no longer. ICES is now the proper legal name of the organization (see:

<https://www.ices.on.ca/About-ICES/New-name-and-logo>). Our agreement with ICES and representatives from ICES requested that the former abbreviation should not be included in the manuscript.

Competing Interests: PKR cofounded and BCS and EJM are employees of CytoGnomix Inc.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research