

G OPEN ACCESS

Citation: Iyaniwura SA, Ringa N, Adu PA, Mak S, Janjua NZ, Irvine MA, et al. (2023) Understanding the impact of mobility on COVID-19 spread: A hybrid gravity-metapopulation model of COVID-19. PLoS Comput Biol 19(5): e1011123. https://doi. org/10.1371/journal.pcbi.1011123

Editor: Virginia E. Pitzer, Yale School of Public Health, UNITED STATES

Received: December 18, 2022

Accepted: April 24, 2023

Published: May 12, 2023

Copyright: © 2023 Iyaniwura 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 author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supplementary material. The Bayesian inference codes are available on GitHub at https://github.com/iyaniwura/Hybrid_ Gravity_Metapopulaiton_model.

Funding: M.O. and N.J. received support for this project from the Michael Smith Foundation for Health Research and the Canadian Institutes of Health Research (CIHR Grant No.VR5-172683). The funders had no role in study design, data

RESEARCH ARTICLE

Understanding the impact of mobility on COVID-19 spread: A hybrid gravitymetapopulation model of COVID-19

Sarafa A. Iyaniwura^{1,2}*, Notice Ringa^{1,3}, Prince A. Adu^{1,3}, Sunny Mak¹, Naveed Z. Janjua^{1,3}, Michael A. Irvine^{1,4}, Michael Otterstatter^{1,3}

1 British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada, 2 Department of Mathematics and Institute of Applied Mathematics, University of British Columbia, Vancouver, British Columbia, Canada, 3 School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada, 4 Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

* iyaniwura@aims.ac.za

Abstract

The outbreak of the severe acute respiratory syndrome coronavirus 2 started in Wuhan, China, towards the end of 2019 and spread worldwide. The rapid spread of the disease can be attributed to many factors including its high infectiousness and the high rate of human mobility around the world. Although travel/movement restrictions and other non-pharmaceutical interventions aimed at controlling the disease spread were put in place during the early stages of the pandemic, these interventions did not stop COVID-19 spread. To better understand the impact of human mobility on the spread of COVID-19 between regions, we propose a hybrid gravity-metapopulation model of COVID-19. Our modeling framework has the flexibility of determining mobility between regions based on the distances between the regions or using data from mobile devices. In addition, our model explicitly incorporates time-dependent human mobility into the disease transmission rate, and has the potential to incorporate other factors that affect disease transmission such as facemasks, physical distancing, contact rates, etc. An important feature of this modeling framework is its ability to independently assess the contribution of each factor to disease transmission. Using a Bayesian hierarchical modeling framework, we calibrate our model to the weekly reported cases of COVID-19 in thirteen local health areas in Metro Vancouver. British Columbia (BC), Canada, from July 2020 to January 2021. We consider two main scenarios in our model calibration: using a fixed distance matrix and time-dependent weekly mobility matrices. We found that the distance matrix provides a better fit to the data, whilst the mobility matrices have the ability to explain the variance in transmission between regions. This result shows that the mobility data provides more information in terms of disease transmission than the distances between the regions.

collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author summary

Mathematical modelling was at the forefront of the fight against COVID-19. Many models were used to study the disease dynamics and the effectiveness of intervention strategies to control the disease spread. Due to the high infectiousness of COVID-19 and the high rate of human mobility, understanding the disease spread between regions became important. Many of the previously developed models for studying disease dynamics between regions described human mobility based on either the distances between the regions or using other forms of mobility data. In addition, some of these models can only be used to study the spread of COVID-19 from an epicentre to neighbouring regions/cities. These models are suitable for only the early stages of the disease outbreak. We developed a hybrid gravity-metapopulation modelling framework for studying the spread of diseases between regions and mobile device data as a proxy for human mobility between regions. Furthermore, our framework is suitable for studying disease dynamics at any epidemic stage. It accounts for disease spread from each region to the remaining regions, irrespective of the number of reported cases in each region and their population sizes.

This is a PLOS Computational Biology Methods paper.

Introduction

The pandemic of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which started in the city of Wuhan, Hubei province, China [1] has since spread all over the world with over 585 million reported cases and 6.4 million reported deaths, as of August 2022 [2]. In human populations, the virus can be transmitted through the inhalation of infectious droplets in aerosols, exposure to infectious respiratory fluids, coughing, sneezing, and having physical contact with an infected individual. It can also be transmitted indirectly when a susceptible individual comes in contact with a contaminated surface, such as door handles or other commonly shared surfaces or objects [3–5]. SARS-CoV-2 is the casual agent for the coronavirus disease 2019 (COVID-19), and is estimated to be more infectious compared to other coronaviruses such as the severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome coronavirus (MERS) [6, 7]. The COVID-19 disease was declared a public health emergency by the World Health Organization (WHO) on January 20, 2020 [8] and a pandemic on March 11, 2020 [9].

Due to the fast spread of COVID-19, during the early stages of the pandemic, governments around the implemented non-pharmaceutical interventions (NPIs) such as movement/travel restrictions, wearing of facemasks, closure of schools and businesses, physical distancing, etc. [10–14], to limit the spread of the disease. Although, the implementation of these NPIs helped in slowing down the spread of COVID-19, the disease still continues to spread under these restrictions. In addition, these NPIs have significant social and economic effects around the world [15–17], and could not be put in place for too long. The development of safe and effective COVID-19 vaccines brought some relief and were introduced to replace stringent NPIs [18, 19]. The first set of COVID-19 vaccines became available towards the end of 2020 [20]. These vaccines provide significant protection against the earlier strains of SARS-CoV-2 virus

[21–23]. However, the emergence of highly infectious mutant strains such as Omicron variant led to the continuous spread of the disease.

The first case of COVID-19 was reported in Wuhan, China, in December 2019 [24]. On the 12th of January 2020, the first case of the disease outside of China was confirmed in Thailand [25]. By January 30, 2020, COVID-19 has spread to 18 countries outside of China with a total of 7,818 confirmed cases worldwide [25]. The first confirmed case of COVID-19 in Africa was reported on February 14, 2020 [26, 27], in North America, January 21, 2020 [28], and in Europe, January 24, 2020 [29]. COVID-19 has spread more rapidly and widely around the world than previous outbreaks of coronaviruses. This spread can be attributed to globalization, settlement and population characteristics, and high human mobility [30]. Several studies have looked at the effect of human mobility on the spread of COVID-19 [31–33]. In Kraemer, Moritz UG, et al [31], real-time human mobility data was used to investigate the role of case importation in the spread of COVID-19 across cities in China. The impact of human mobility network on the onset of COVID-19 in 203 countries was studied in [32]. They used exponential random graph models to analyze country-to-country spread of the disease. Their study suggested that migration and tourism inflow contributed to COVID-19 case importation, and that a mixture of human mobility and geographical factors contribute to the global transmission of COVID-19 from one country to another. Human mobility data collected via mobile devices such as cell phones, smartwatches, e-readers, tablets etc., have also been used to study the spread of COVID-19 [34-37]. In [34], county-level cell phone mobility data collected over a period of 1 year in the US was used to study the spatio-temporal variation in the relationship between COVID-19 infection and mobility. They found that in the spring 2020, sharp drop in mobility often coincide with decrease in COVID-19 cases in many of the populous counties.

Mathematical models have been used to study the relationship between the spatio-temporal spread of COVID-19 and human mobility [37–45]. A city-based epidemic and mobility model together with multi-agent network technology and big data on population migration were used to simulate the spatio-temporal spread of COVID-19 in China [45]. In [46], a stochastic, data-driven metapopulation model was used to study the initial wave of COVID-19 in Belgium, and also to study different re-opening strategies. Their model incorporates the mixing and mobility of different age groups in Belgium. Another stochastic metapopulation model was used to study the spread of COVID-19 in Brazil [43]. This model assumes that epidemics start in highly populated central regions and propagate to the countrysides. For many states, they found strong correlations between the delay in epidemic outbreaks in the countrysides and their distance from central cities. In [47], an SEIR country-wide metapopulation model was used to study the spread of COVID-19 in England and Wales. The model was used to predict the COVID-19 epidemic peak in England and Wales, and also to study the effect of different non-pharmaceutical intervention strategies on the predicted epidemic peaks. Similarly, in [48] a stochastic SIR model was applied to describe the spatio-temporal spread of COVID-19 across 33 provincial regions in China and to also evaluate the effectiveness of various local and national intervention strategies. Their model incorporates an outflow mobility index for all the regions and the proportion of travelers between regions. More discussions on human mobility and COVID-19 transmission can be found in the systematic review article [49]. The relative contribution of mobility data to the observed variance in the COVID-19 transmission rates between regions still remains an unexplored problem.

We develop a hybrid gravity-metapopulation modeling framework for studying the spread of COVID-19 within and between different regions. An important feature of our framework is the ability to determine human mobility based on the distances between the regions or through empirical data such as those collected through mobile devices. In addition, our framework allows for the explicit incorporation of factors that affect disease transmission, such as facemasks, physical distancing, contact rates, etc., into a time-dependent disease transmission rate and the assessment of the contribution of each of these factors to actual disease transmission. As an illustration, we use a Bayesian hierarchical modeling framework to calibrate our model to the weekly reported cases of COVID-19 in the thirteen local health areas (LHAs) of Fraser health authority (Fraser Health), British Columbia (BC), Canada, from July 2020 to January 2021. The study area comprises 1.9 million population in the eastern sections of the Greater Vancouver area. We estimate region-specific scaling parameters for computing baseline disease transmission rates for each region, and a parameter for quantifying the contribution of mobility to disease transmission. In addition, we estimate a time-dependent piece-wise constant scaling parameter to account for the cumulative effect of the remaining factor that affect disease dynamics, which are not explicitly included in our model. We consider two main model structures in our example, which are determined by the mobility matrices used: one with a distance matrix (computed using the distances between the regions, based on the population weighted centroid) and another with time-dependent mobility matrices computed from mobile device data. The results from these two scenarios are used to test the hypothesis of whether the time-dependent mobility matrices, computed from mobile device counts, provide more information about human mobility, with respect to disease transmission between the regions than the distances between the region.

Materials and methods

Mathematical model

We develop a hybrid gravity-metapopulation model to study the dynamics of COVID-19, within and between regions. The model stratifies the population of each region into six compartments: susceptible (S), exposed (E), pre-symptomatic infectious (P), symptomatic infectious (I_1 and I_2), and recovered (R). Individuals in the pre-symptomatic infectious compartment are infectious (can transmit the disease) but do not show symptoms yet. Similar to [50, 51], we divided the infectious compartment into two classes so that the recovery time follows a Gamma distribution rather than an exponential distribution. This way, a symptomatic infectious individual spends the first half of their infectious period in I_1 and the other half in I_2 . We assume that there is no re-infection in our model due to relatively low infections across the study period as the size of the susceptible population is far greater than the size of the recovered population. In addition, we assume that all the individuals infected during our study period will not lose their COVID-19 immunity and be reinfected within this period [52]. Furthermore, asymptom-based.

A schematic diagram of the model illustrated for four (4) regions is shown in Fig 1, where the gray circles on the left represent the regions, while the black arrows show the interactions and movements of individuals between the regions. On the right, we have an illustration of the population dynamics in each of the regions, where the subscript *j* represents the *j*th region. The black arrows here show the transition of individuals through the different stages of COVID-19 at the rates indicated beside the arrows. The red dashed arrows indicate disease transmission. Observe that there is a red dashed arrow extending from each of the remaining three regions into region *j*, these arrows account for the contributions of infectious individuals in the three regions to disease transmission in the *j*th region. The ordinary differential equations (ODEs) for the model are given by (see Fig 1 for model schematic diagram and definition of state



Fig 1. Schematic of the model. An illustration of our hybrid gravity-metapopulation model for four regions. Interactions and movements between the regions (left) and stratification of the population of each region based on disease stages (right). Model compartments are defined as follows: Susceptible (S_j) ; exposed (E_j) ; pre-symptomatic infectious (P_j) ; symptomatic infectious $(I_{1j} \text{ and } I_{2j})$; and recovered (R_j) for region *j*. Our model assumes that there are no re-infections. The black arrows show the movement of individuals from one region to another (left) and the transition of individuals through the different stages of COVID-19 at the rates indicated beside the arrows (right). The red dashed arrows indicate disease transmission (see (1) for more details).

variables):

$$\begin{split} \frac{dS_j}{dt} &= -\beta_j \Psi_j S_j, \\ \frac{dE_j}{dt} &= \beta_j \Psi_j S_j - h_1 E_j, \\ \frac{dP_i}{dt} &= h_1 E_j - h_2 P_j, \\ \frac{dI_1}{dt} &= h_2 P_j - 2\gamma I_{1j}, \\ \frac{dI_{2j}}{dt} &= 2\gamma I_{1j} - 2\gamma I_{2j}, \\ \frac{dR_j}{dt} &= 2\gamma I_{2j}, \end{split}$$
(1)

where $\beta_j \equiv \beta_j(t)$ is the time-dependent disease transmission rate for region *j*. We aim to define the transmission rate (β_j) as a function of the different factors that affect disease transmission. This way, we would be able to evaluate the contribution of each of these factors to the overall disease transmission. Therefore, we define $\beta_j(t)$ as

$$\beta_j(t) = \exp(c_{0j} + c_1 m_j(t) + g(t)).$$
(2)

Here, c_{0j} is the scaling parameter for the baseline disease transmission rate for region *j*, c_1 is the scaling parameter used to remove biases from the time-series mobility data, and g(t) is a time-dependent piece-wise parameter used to account for other factors that affect disease transmission other than human mobility (e.g. facemask, social distancing, contact rates, etc.), which are not explicitly incorporated into the model. Movement within the j^{th} region is captured by a time-series mobility data represented by $m_i(t)$. This data is used as a proxy for the time-dependent contact rate in the region. We have defined our disease transmission rate, $\beta_i(t)$, as an exponential function to ensure that its value remains positive due to the way the time-series mobility data and the function g(t) are to be incorporated into $\beta_i(t)$. This definition will ensure that the estimated model parameters are identifiable (see Bayesian inference section for more details). Based on the definition of $\beta_i(t)$ in (2), $e^{c_{0j}}$ is the baseline disease transmission rate for region *j*, while $e^{c_1 m_j(t)}$ incorporates the effect of human mobility within the region into the transmission rate. Lastly, e^{g(t)} accounts for the effect of other factors that affect disease transmission, which are not explicitly incorporated into the model, on the disease spread. Although, the formulation in (2) explicitly incorporates only human mobility into the disease transmission rate, this formulation can be extended to include other factors that affect disease transmission such as facemaks, physical distancing, etc. See more details in the Discussion section.

The parameter $\Psi_j \equiv \Psi_j(t)$ in (1) is used to incorporate infectious interactions within the *j*th region, and their contribution to disease transmission in the region. In terms of a homogeneous single population model, this parameter would represent the probability of making an infectious contact in the population. Here, Ψ_j is defined as

$$\Psi_{j}(t) = \frac{1}{\hat{N}_{j}} \left[(1-\theta)(P_{j} + I_{1j} + I_{2j}) + \theta \sum_{i=1}^{M} \pi_{ji} \left(P_{i} + I_{1i} + I_{2i} \right) \right],$$
(3)

where $0 \le \theta \le 1$ is a parameter used to measure the effective contribution of human mobility to disease transmission in all the regions. Here, $\theta = 0$ implies that there are no infectious contacts due to mobility as defined by the intra-regional mobility matrix (π), and the regions are essentially uncoupled from each other. On the other hand, $\theta = 1$ means that all the infectious contacts in the system are due to human mobility. The parameter \hat{N}_j is the adjusted population size for region *j*, which incorporates the changes in the population size of the region due to movements in and out of the region. We define \hat{N}_j by

$$\hat{N}_j = (1-\theta)N_j + \theta \sum_{i=1}^M \pi_{ji} N_i,$$
(4)

where *M* is the total number of regions under consideration and N_j is the baseline population size of the *j*th region. The first term in (3) given by $(1 - \theta)(P_j + I_{1j} + I_{2j})/\hat{N}_j$ accounts for all the infectious contacts made by the residents of region *j* who are not moving within the region, while the second term, $(\theta/\hat{N}_j) \sum_{i=1}^{M} \pi_{ji} (P_i + I_{1i} + I_{2i})$, accounts for all the infectious contacts made in region *j* by the residents of the region who are moving within the region and the visitors from other regions. In (3) and (4), π_{ji} is the probability that an individual who migrated into region *j*, originated from region *i*, given that he/she is from one of the other regions under consideration. We compute this probability using two different approaches. The first approach uses the distances between the regions. In this case, π_{ji} is given by

$$\pi_{ji} = \frac{f(d_{ij})}{\sum_{i=1}^{M} f(d_{ij})}, \quad \text{with} \quad f(d_{ij}) = \frac{1}{\left(1 + d_{ij}\right)^{k}}, \tag{5}$$

Parameter	Description	Value	References
$\beta_j(t)$	Time-dependent disease transmission rate for region <i>j</i>	_	Computed using (2)
c _{0j}	$\exp(c_{0j})$ is the baseline disease transmission rate for region <i>j</i>	_	Estimated
<i>c</i> ₁	Scaling parameter for the time-series mobility data		Estimated
<i>g</i> (<i>t</i>)	Time-dependent piece-wise scaling parameters used to account for the effect of other factors that affect disease transmission	_	Estimated
h_1	Rate of transitioning from exposed to pre-symptomatic infectious	1/5 (days ⁻¹)	[53-55]
h_2	Rate of transitioning from pre-symptomatic infectious to symptomatic infectious	1 (days ⁻¹)	[54, 56-58]
γ	Infection recovery rate	1/5 (days ⁻¹)	[<u>54</u> , <u>57</u> , <u>58</u>]
θ	Measures the effective contribution of mobility to disease transmission	_	Estimated
d_{ij}	Distance from region <i>i</i> to region <i>j</i>	See Fig 3	Computed using [59]
π_{ji}	Probability that an individual moving in region <i>j</i> is from region <i>i</i>	See Figs $\frac{3}{2}$ & $\frac{4}{4}$	Computed using (5)
N_j	Baseline population size for region <i>j</i>	See Table C in <u>S1</u> <u>Text</u> .	[60]
\hat{N}_{j}	Adjusted population size for region <i>j</i>	_	Computed using (4)

Table 1. Model parameters, their descriptions, and values. The estimated parameters are presented in the Results section. The population sizes for the regions are presented in Table C in S1 Text.

https://doi.org/10.1371/journal.pcbi.1011123.t001

Where $d_{ij} \equiv d_{ji}$ is the distance from region *i* to region *j*, $k \in \mathbb{R}^+$ and *M* is the total number of regions considered. The second approach used to compute the probability π_{ji} involves using mobile device data.

The model parameters, their descriptions, and values are provided in <u>Table 1</u>. The estimated parameters are presented in the Results section.

As an illustration of concept, we consider the thirteen (13) local health areas (LHA) of Fraser health authority, British Columbia (BC), Canada. These regions include the communities of Abbotsford, Agassiz/Harrison, Burnaby, Chilliwack, Delta, Hope, Langley, Mission, Maple Ridge/Pitt Meadows, New Westminster, South Surrey/White Rock, Surrey and Tri-Cities. Fraser health authority (Fraser Health) is the largest of the five regional health areas in BC, with 12 acute care hospitals and providing health care to over 1.9 million people [65]. It has a width of 150 km. Fig 2 shows a map of Fraser Health with the 13 health regions shaded. The population sizes of the regions are detailed in Table C in <u>S1 Text</u>. We use a Bayesian hierarchical modeling framework to calibrate our model to the weekly reported cases of COVID-19 in these 13 LHAs, from July 2020 to January 2021. From the model calibration, we estimate the parameters c_{0j} , c_1 and g(t), which are used to construct and study the time-dependent disease transmission rate for each region, and to study the dynamics of the time-dependent piece-wise parameter g(t). We also estimate the parameter θ , used to quantify the effect of mobility, both within and between the regions, on disease transmission in the regions.

Data

Human population move between regions for many reasons including work, leisure, family visits, health reasons, e.t.c. The main goal of this work is to develop a mathematical modeling framework for studying and understanding the effect of human mobility on the spread of COVID-19 within and between regions. We consider the period from July 1, 2020 to January 27, 2021, inclusive. Although, movement restriction was imposed in Fraser Health during





some part of this period, we used the mobility data collected through mobile device counts as a proxy for quantifying movements between the regions and the contact rate within each region.

We quantify mobility between the regions using two approaches. The first approach uses the physical distances between the regions, based on population weighted centroid (left panel of Fig 3) and the formula in (5) to calculate the probability that an individual moving in region *j*, who came from one of the 13 regions, originated from region *i* (π_{ji}). The premise of using physical distance between regions is based on the concept of geographic distance decay, where spatial and social interactions decrease as the distance between regions increases [59]. The distances between our regions of interest are given in the left panel of Fig 3, while probabilities π_{ji} computed from these distances are presented in the right panel.

The diagonal entries of the probability matrix (π) represents the probability that an individual moving in a region is a resident of that region. It is important to note that the probability matrix is not symmetric, even though the distance matrix (left panel of Fig 3) is symmetric. In addition, each row of the probability matrix sums to 1. The second approach used to construct



Fig 3. Distance matrices. Physical distances (in km) between the local health areas (LHAs) based on population-weighted centroid (left) [59] and the probability matrix (π) computed using (5) (right). π_{ji} is the probability that an individual who came from one of the 13 regions to region *j*, originated from region *i*.

the probability matrix (π) is based on mobile device counts and uses Telus mobility (TELUS) data.

TELUS is a Canadian national telecommunications company that has network coverage in 99% of the populated areas of Canada. TELUS Insights provides anonymized geo-intelligence data, which reflects population location and mass movement patterns based on information about locations and population movement of TELUS mobile device users [66]. These data have helped answer a range of questions around location and public mobility patterns within Canada, including in infrastructure planning, health services, roads, and transit routes.

As TELUS subscribers use their mobile devices, they connect to various cellular towers for telecommunication services. These connections are used to determine the users' locations based on the nearest tower that relays signals to their devices. Every Telus user with their mobile network active would be included in the TELUS data, except the subscriber opt-out [67]. This network data provides insights into movement patterns and trends across Canada. To provide a layer of privacy, all the mobility data provided by TELUS are de-identified, aggregated into large data pools, rounded-up to the nearest 10 counts and all results are extrapolated to represent the entire population of a given region. This ensures that the data cannot be traced back to individual TELUS subscribers. The results of the TELUS application programming interface (API) implementation, which provide the numbers of mobile devices moving within and between geographical locations of interest and the neighbourhood that a mobile device resides in, depend on cellular tower locations at the time of the analysis.

We generate the mobility data for each region using a one-day bucket size and 120-minute minimum dwell time. We filtered for "non-residents", "moving residents" and "residents", which represent, respectively, the daily number of mobile devices residing in an LHA and spending at least two hours in another LHA (movement between regions), the daily number of mobile devices residing in an LHA and spending over two hours outside their census track within the LHA (movement within a region), and the total number of mobile devices residing in an LHA. To construct the weekly mobility matrices, we consider the "non-residents" and "moving residents" data. For each region and for a specified time interval (weekly), we





compute the number of mobile devices from the other 12 regions that visited the region and stayed there for at least 2 hours during the visit. This gives us the mobile device count for movement into the region (off-diagonal entries). For movement within the regions (diagonal entries), we used the "moving residents" data, from which we computed the number of mobile devices registered to a region and moving within the region. These information are used to construct a mobility matrix of device counts within the specified time interval for each region. We normalize each row of the matrix with the total number of devices in the row. This way, the i^{th} element of the j^{th} row represents the fraction of mobile devices that came into the j^{th} region (from the 13 regions) that originated from the *i*th region. These fractions can also be interpreted as the probability that an individual moving within the i^{th} region (whom originated from one of the 13 regions) is from the i^{th} region (π_{ii}). Using this approach we compute the probability/mobility matrices for each week from July 1, 2020 to January 27, 2021. The computed matrices for week 1 (July 1-7, 2020) and 30 (January 21-27, 2021) are shown in Fig 4, while the matrices for the remaining weeks are presented in Figs (B-F) in S1 Text. The distance matrix (right panel of Fig 3) and the constructed mobility matrices are used to describe the interaction between individuals from different regions. We considered two main scenarios in our Bayesian inference based on the distance and mobility matrices and investigated whether the mobility data is more informative, in terms of disease transmission than the distances between the regions.

We also used the Telus mobility data to compute the weekly mobility rate for the movements within each region. To compute these rates, we sum the daily device count in each region for "non-residents" and "moving resident", and divide it by the sum of the "residents" and "non-residents" device count for our entire study period. This gives us the proportion of mobile devices moving in each region with respect to the total number of devices in the region during our entire study period. For each week in our study period, we sum the computed proportion of mobile devices and divide by 7 to get the weekly average proportion of mobile devices moving in each of the regions, as shown in Fig 5. These mobility rates are used as



Fig 5. Weekly mobility rates. Weekly time-series mobility rates for each region from July 2020—January 2021, computed from the Telus mobility data.

proxy for the contact rates in the regions and are represented by $m_j(t)$ in the disease transmission rate ($\beta_j(t)$) defined in (2). We observe from Fig 5 that there is a sharp decline in mobility rate around the first week of September 2020 in most of the regions. Similarly, there is another decline in mobility rate around the first week on November. This decline is associated with the implementation of public health measures in BC.

We calibrate our model to the weekly reported cases of COVID-19 in the thirteen local health areas of Fraser Health, BC, obtained from the British Columbia Centre for Disease Control (BCCDC). We extracted these data from a line list generated by BCCDC Public Health Reporting Data Warehouse (PHRDW), based on symptom onset date or reported date where symptoms onset date is not available. The collected case data spans the period from July 2020 to January 2021, inclusive, and was incorporated into the model likelihood based on the computed disease incidence as shown in (6). The collected weekly reported cases of COVID-19 for the 13 regions are shown in Fig A in S1 Text. Similar to [50], our model incidence is computed as the number of individuals in the pre-symptomatic population (P), transitioning to the symptomatic infectious compartment (I_1).

Bayesian inference

Our hybrid gravity-metapopulation model (1) is fitted to the COVID-19 cases in all the thirteen regions using a Bayesian hierarchical modeling framework. Bayesian inference is a statistical technique for data analysis and parameter estimation, which is based on the Bayes' theorem. It has been applied to problem in many fields ranging from biology, physics, sport, epidemiology, ecology, and engineering, among others [68-70]. A Bayesian hierarchical modeling framework is one where the prior distribution of some of the model parameters depend on other parameters to be estimated. It allows the incorporation and estimation of model parameters at individual and population levels (see [71, 72] for more information on Bayesian hierarchical models). We implement our Bayesian inference model with the RStan package in R version 3.6.3 [73]. Stan is a free and open-source probabilistic programming language for statistical inference implemented in C++. It performs Bayesian inference on arbitrary user-defined models through Markov Chain Monte Carlo (MCMC), and can be invoked through other programming languages such as Python, Matlab, Julia and R. [74]. RStan is the R interface to Stan, which provides full Bayesian inference via the No-U-Turn sampler (NUTS), a variant of Hamiltonian Monte Carlo (HMC), approximate Bayesian inference via automatic differentiation variational inference (ADVI), and penalized maximum likelihood estimation via L-BFGS optimization [73].

In our Bayesian inference model, we construct the likelihood for the j^{th} region as

$$cases_i(t) \sim NegBin(incidence_i(t)), \psi),$$
 (6)

where NegBin(·) is the negative binomial distribution, $cases_j(t)$ and $incidence_j(t)$ are the weekly reported cases of COVID-19 and the incidence computed from the model (1), respectively, for region *j*. Here, ψ is the over-dispersion parameter. Using Bayesian inference framework implemented in Rstan gives us the flexibility to incorporate our prior knowledge into the model parameters and the ability to evaluate probabilistic statements of the data based on the model. In addition, this framework allows us to incorporate hierarchical structure into the model parameters, with the benefit of understanding variations in the parameters at individual and population levels. It enables us to construct the posterior distribution for the population mean and variance of the model parameters and those of the individual parameters for each region, which are conditioned on the population mean and variance. We have used a negative-binomial distribution to model the weekly reported cases of COVID-19 because of its

effectiveness and convenience in modeling nonnegative over-dispersed data. Uninformative priors were implemented in the Bayesian inference framework.

We incorporate the time-series mobility data (Fig 5) into our modeling framework using an exponential scaling approach for the disease transmission rate. The disease transmission rate is given by (2), where $c_{0j} \sim \mathcal{N}_+(0,1)$ is the scaling parameter for the baseline transmission rate for the j^{th} region ($e^{c_{0j}}$ is the baseline transmission rate) and $c_1 \sim \mathcal{N}_+(0,1)$ is the scaling parameter used to remove biases from the time-series mobility data ($m_j(t)$). Here, $e^{c_1 m_j(t)}$ models the time-varying effect of $m_j(t)$ on the disease transmission rate (β_j) for region j. The time-dependent piece-wise constant parameter g(t) is used to account for other factors that affect disease transmission, which are not explicitly accounted for in the model. This parameter is estimated every four weeks (except for the last interval which has 2 weeks). We also estimated the total prevalence of COVID-19 in all the 13 regions at the beginning of our study period. Similar to [50, 54], when building our Bayesian inference modeling framework, we simulated the incidence for our model (1) using known parameters values and then tested the ability of our framework to recover the values. We inspect the resulting posterior distribution for biases and their coverage of the true parameters.

Throughout this article, we used the Variational Bayes (VB) method with the meanfield algorithm implemented in RStan [75, 76] for our inferences, from which we estimate the total initial prevalence in all the 13 region and a parameter, θ , used to quantify the effective contribution of mobility to disease transmission in the regions (see the formulations in (3) and (4)). We estimate a fixed value of the parameter g(t) for every four weeks, starting from the beginning of our study period, and for the last two weeks. Thereby making it a time-dependent and piece-wise parameter. To ensure that the estimated parameters are identifiable and that the estimated values of g(t) from the second interval onward are relative to that of the first interval, we set g(t) = 0 for the first four weeks (first sub-interval). In addition, we rescaled the timeseries mobility data using the first week's mobility rate as a reference for the remaining rates. This was done by subtracting the mobility rate for the first week from those of the subsequent weeks. This way, the rescaled mobility rate for the first week is 0, while those for the remaining weeks are centered around 0. We used a Bayesian hierarchical modeling framework to estimate the parameters c_{0i} and g(t). We construct their population posterior distributions, which are used as priors for estimating the region specific c_{0i} for j = 1, ..., 13, and the interval specific g(t), respectively. The remaining parameters of the model are fixed and are as presented in Table 1.

We consider two main scenarios in our model calibration: one with a fixed distance matrix (computed from the distances between the regions, see Fig 3) and another with weekly mobility matrices (computed from Telus mobility data, see Fig 4 and Figs (B—F) in S1 Text). These two matrices are used to quantify mobility between the 13 regions. Performing inference based on these two scenarios enabled us to understand the effect of mobility on the posterior predictive distributions of the model and to determine which of the two mobility quantifiers best recreates the observed case data. It would also help us to identify, which of the two approaches provides more information on human mobility in terms of disease transmission. The two scenarios were ranked by comparing their leave-one-out predictions and standard errors, computed using the leave-one-out cross-validation (LOO) method [77–79], and using the widely applicable information criterion (WAIC) method [80, 81]. We compared the Variational Bayes (VB) method to the adaptive Hamiltonian Monte Carlo method No-U-Turn sampling. The results from both methods are found to produce comparable estimates of the posterior distribution with significant reduction in total computation time when VB is used [82]. For the case of a fixed distance matrix, the mean and/or median ELBO usually converges in 5, 000–6, 000 iterations of the stochastic gradient ascent algorithm, while it converge in 11, 000–12, 000 iteration for the weekly mobility matrices case (see [76, 83, 84] for more information about ELBO in the variational Bayes method).

Results

We considered two main scenarios when fitting our model to the weekly reported cases of COVID-19 (see <u>Methods</u>). Results for the two scenarios, for selected regions (Agassiz/Harrison, New Westminster, Maple Ridge/Pitt Meadows and Surrey), are presented in Fig.6. We selected these regions based on their population sizes and geographical locations, to show the diversity in reported cases and population sizes in the regions considered, and the model's ability in predicting cases irrespective of these factors. The results for the remaining regions are presented in Figs G and H in <u>S1 Text</u>.

For each model scenario, we present the posterior predictions of the weekly cases of COVID-19 in each region (columns 1 & 3 of Fig 6). We compute the time-dependent disease transmission rate, $\beta_j(t)$, using the estimated parameters and the formula in (2). These rates are presented in blue for the fixed distance matrix scenario (column 2 of Fig 6) and in gold for the weekly mobility matrices scenario (column 4 of Fig 6), together with the contribution of mobility to the transmission rate, $e^{c_1 m_j(t)}$ (green) with 50% credible interval (CrI) (darker bands) and 90% CrI (lighter bands).

We observe from these results that our model is able to capture the trends and reported cases of COVID-19 in each of the regions with a high degree of accuracy for both model scenarios. In addition, we see that there are significant changes in the computed disease transmission rate over time, which has a similar trend for all the regions. Even though there are no much changes in the time-series mobility data, its effect on the disease transmission rate for each region is still noticeable.

The mean estimate for the initial total prevalence in the 13 regions is 47.61 (90% CrI: 44.82 -50.31) for the distance matrix scenario and 50.19 (90% CrI: 47.37-53.04) for the weekly mobility matrix scenario. The mean estimate of the parameter used to quantify the effect of mobility on disease transmission in the regions (θ) for the distance matrix scenario is 0.53 (90% CrI: 0.44-0.60) and 0.90 (90% CrI: 0.72-0.98) for the scenario with weekly mobility matrices. This implies that movement between the regions contribute to a mean fraction of 0.53 and 0.90 of the total reported cases of COVID-19 in the regions for the distance and mobility matrix scenarios, respectively. The scaling parameter used to remove biases in the time-series mobility data (c_1) was estimated as 1.51 (90% CrI: 0.90-2.10) for the distance matrix and 2.11 (90% CrI: 1.52-2.69) for the mobility matrix scenario.

We estimated the scaling parameters for the baseline disease transmission rate, c_{0j} for j = 1, ..., 13, using Bayesian hierarchical modeling framework. These parameters are used to compute the baseline disease transmission rate for each region defined by $e^{c_{0j}}$ for j = 1, ..., 13. The mean estimate for the population mean and variance are 0.45 (90% CrI: 0.35—0.54) and 0.18 (90% CrI: 0.12—0.26), respectively, for the distance matrix and 0.21 (90% CrI: 0.05—0.36) and 0.35 (90% CrI: 0.24—0.47), respectively, for the mobility matrix scenario. The mean estimate for c_{0j} , for j = 1, ..., 13 with 90% credible interval (CrI) are presented in Tables A (distance matrix) and B (mobility matrix) in <u>S1 Text</u>. The estimated distribution for the baseline disease transmission rates for the regions are presented in Fig 7. We observe from the results in this figure that the predicted distributions for the larger and more urbanized regions with dense populations are similar for the distance and mobility matrix scenarios. These regions include Abbotsford, Burnaby, New Westminster, Surrey, and Tri-Cities. On the other hand, the predictions for the less densely populated smaller regions are relatively different for the two

scenarios. In addition, the variances in the distributions for the smaller regions are larger than those of the bigger regions with larger populations.

The time-dependent piece-wise parameter, g(t), was also estimated using a Bayesian hierarchical modeling framework with population mean and variance estimate with 90% credible interval given by -0.33 (-0.52, -0.14) and 0.30 (0.16, 0.47), respectively, for the distance matrix scenario, and -0.28 (-0.45, -0.10) and 0.32 (0.19, 0.50), respectively, for the weekly mobility matrix scenario. The mean estimates with 90% credible interval for the interval-specific parameters (g_2 - g_8) are given in Tables A (distance matrix scenario) and B (mobility matrix scenario) in S1 Text. It is important to emphasize that we have set $g_1 = 0$ (week 1–4) to ensure that



Fig 6. Observed and estimated COVID-19 cases. Weekly reported cases of COVID-19 and model prediction (columns 1 and 3). Disease transmission rate, $\beta_j(t)$ and the contribution of mobility to disease transmission, $e^{c_1 m_j(t)}$ (green curves in columns 2 and 4). Model types: fixed distance matrix (blue) and weekly mobility matrices (gold). Black dots are the weekly reported cases of COVID-19, the solid lines are the mean estimates of cases/parameters, the darker bands are the 50% CrI, while the lighter bands are the 90% CrI. Similar results for the remaining regions are presented in Figs G and H in S1 Text.

https://doi.org/10.1371/journal.pcbi.1011123.g006



Fig 7. Baseline disease transmission rate. The distributions for the baseline disease transmission rate, e^{c_0} , for each region computed using the estimated parameters c_{0j} for j = 1, ..., 13 (see Tables A and B in <u>S1 Text</u> for the estimates of c_{0j}). Scenarios: fixed distance matrix (blue) and weekly mobility matrices (gold).

the model parameters are identifiable, and to estimate g_2, \ldots, g_8 relative to g_1 . The distributions for the time-dependent effect of other factors that affect disease transmission, other than mobility, on the disease transmission rate, are given in Fig 8. We observe that the constructed distributions for the two scenarios agree well.

Lastly, we compare the estimated expected leave-one-out predictions and their standard errors, for the two model scenarios, computed using the leave-one-out cross-validation (LOO) method [77–79] and the widely applicable information criterion (WAIC) method [80, 81].

The comparison is summarized in Table 2, where the distance matrix scenario is ranked better than the mobility matrix scenario, in terms of their ability to capture the case data. Even



Fig 8. Contribution of other factors to disease transmission. The distributions for the contribution of other factors that affect disease transmission $(e^{g(t)})$ to the transmission rate $(\beta(t))$, computed every four weeks and for the last two weeks: $g_1 = 0$ (weeks 1–4), g_2 (weeks 5–8), g_3 (weeks 9–12), g_4 (weeks 13–16), g_5 (weeks 17–20), g_6 (weeks 21–24), g_7 (weeks 25–28), g_8 (weeks 29 & 30). Scenarios: fixed distance matrix (blue) and weekly mobility matrices (gold). The estimated means with 90% credible interval are presented in Tables A and B in S1 Text.

though the distance matrix scenario captures the case data better than the weekly mobility matrix scenario, the difference in the fits for the two approaches is not much.

Discussion

An important feature of our modeling framework includes the ability to explicitly incorporate factors that affect disease transmission into the transmission rate. This formulation allows us to effectively access the contributions of these factors to disease transmission in our model. In the example presented in this article, due to lack of adequate data, only time-series mobility data was incorporated explicitly into the disease transmission rate. The effect of other factors

Table 2. Model comparison using leave-one-out cross-validation (LOO) and the widely applicable or Watanabe-Akaike information criterion (WAIC). Model ranking (in descending order) is shown in the first column. The difference between the expected log pointwise predictive density (ELPD) for each scenario and that of the best scenario with standard errors are shown in the second column. In the third column, we have the Bayesian LOO estimate of the expected log pointwise predictive density (ELPD LOO) and its standard error. The LOO information criteria (LOOIC) and its standard error are given in the fourth column. Lastly, the computed Watanabe-Akaike information criterion (WAIC) for each model is shown in the fourth column.

Scenario	ELPD difference (SE)	ELPD LOO (SE)	LOOIC (SE)	WAIC (SE)
Distance matrix	0.0(0.0)	-1537.8(34.4)	3075.6(68.9)	3074.9(68.8)
Mobility matrix	-9.5(6.9)	-1547.3(36.2)	3094.6(72.4)	3093.4(72.3)

https://doi.org/10.1371/journal.pcbi.1011123.t002

that affect disease transmission was accounted for using a time-dependent piece-wise parameter. We attempted to incorporate the effect of facemasks into the model but could not get adequate data for facemask usage in each of the regions we considered. In this case, the disease transmission rate was formulated as follows

$$\beta_i(t) = \exp(c_{0i} + c_1 m_i(t) + c_2 f(t) + g(t)), \tag{7}$$

where c_{0j} for j = 1, ..., 13 are region-specific scaling parameters used to compute the baseline disease transmission rate for each region ($e^{c_{0j}}$ is the baseline transmission rate for region j). The parameters c_1 and c_2 are covariates for the mobility and facemask usage rates, respectively, and g(t) is a time-dependent piece-wise parameter that is used to incorporate the effect of other factors that affect disease transmission other than mobility and facemask. This formulation can always be extended to explicitly account for other factors that affect disease transmission based on data availability.

Our model captures the trends and reported COVID-19 cases in each region (see Fig 6, and Figs G and H in S1 Text). In addition, the results of the two model scenarios agree well, although, there is a slight different in the estimated time-dependent disease transmission rates and the contribution of mobility to disease transmission (columns 2 & 4) for some regions. There are significant changes in the computed disease transmission rate over time, which may be attributed to the intervention strategies implemented by the government during this period. Even though there are no much changes in the time-series mobility data, its effect on the disease transmission rate is still apparent for each region. The estimated total initial prevalence of COVID-19 in all the regions for the two scenarios agree well, as well as the estimates for the time-dependent piece-wise parameter (g(t)), used to incorporate the effect of other factors that affect disease transmission into the transmission rate (see Fig 8). However, the estimated effect of mobility on disease transmission is significantly different for the two scenarios. The mean estimate of this parameter was 0.53 (90% CrI: 0.44-0.60) for the distance matrix scenario and 0.90 (90% CrI: 0.72–0.98) for the weekly mobility matrix scenario. This can be interpreted as mobility contributing to 53% and 90% of the cases in the regions for the distance and mobility matrices scenarios, respectively. These results show that the weekly mobility data provides more information, in terms of disease transmission, than the distances between the regions. Note that the mobility referred to here is for both within and between the regions. To confirm that indeed the weekly mobility data provides more information, we considered a third scenario, where we used a fixed mobility matrix computed using the mobility data for the entire study from July 2020 to January 2021. For this scenario, we estimated the effect of mobility on disease transmission as 0.60 (90% CrI: 0.52–0.70) (see Fig (I-K) and Table D in S1 Text). As expected, the fixed mobility matrix does not provide more information about disease transmission than the weekly mobility matrices, even though it does better than the distance matrix.

The constructed distributions for the baseline disease transmission rate for the two model scenarios are similar for some of the regions and significantly different for other regions. These distributions are similar for the larger and more urbanized regions with dense population (Abbotsford, Burnaby, New Westminster, Surrey and Tri-Cities) and significantly different for the less densely populated smaller regions (see Fig 7). The difference in the predicted distributions for these two groups of regions may be attributed to their population size and mobility in the regions. Lastly, we compared the results obtained from the two scenarios using the leave-one-out cross-validation (LOO) and the widely applicable information criterion (WAIC) methods. This comparison ranks the distance matrix results better than those of the weekly mobility matrices, although, the computed LOO and WAIC for the two scenarios are very similar (see Table 2). We considered these two model scenarios in order to test the hypothesis of whether the time-dependent mobility matrices, computed from the mobile device data, provide more information about human mobility between the regions in terms of disease transmission than the distances between the regions. Based on our results, we conclude that even though the distance matrix provides a better fit to the data, the weekly mobility matrices have the ability to explain the variance in transmission between regions over time. The model for when the distance matrix is used is considered a *gravity model*, while the scenario where the weekly mobility matrices are used is referred to as a *metapopulation model*. Hence, our hybrid gravity-metapopulation model.

Unlike in other models used to study the effect of human mobility on disease spread, where mobility is described based on either the distances between regions or using mobile devices data or other forms of mobility data only [37, 45, 47], our hybrid gravity-metapopulation modeling framework provides the flexibility of switching between the two data types. In addition, our framework provides an approach for studying the spread of diseases between all the regions of interest, rather than from an epicenter or a large city to its neighboring smaller cities [43, 45], with the ability to quantify the effective contribution of mobility to disease spread between the regions. The models used to study disease spread from an epicenter to neighboring regions/cities are only suitable for studying disease spread at the early stages of the disease outbreak since there is a much higher probability of disease transmission from a person living in the epicenter to those living in the neighboring regions as shown in [45]. Also these models do not account for disease transmission between the smaller neighboring regions. Our modeling framework is suitable for studying disease dynamics at any stage of the epidemic and accounts for disease spread from each of the regions to all the remaining regions, irrespective of the number of reported cases in each region.

Overall, our modeling framework provides the ability to explicitly incorporate real data on factors the affect disease transmission into the disease transmission rate, and also allows independent assessment of the contribution of these factors to disease transmission in an epidemic. Furthermore, this framework allows us to quantify the effect of mobility on disease transmission in the regions. However, this work is not without limitations. We quantified the effect of mobility on disease transmission in the 13 LHAs of Fraser Health, BC, based on movements between these thirteen regions only. However, there are movement in and out of these regions to other parts of BC. Another limitation of this work is that some regions in Fraser Health are closer to regions in other regional health areas in BC, than they are to other regions in Fraser Health. For example, Burnaby is closer to Vancouver than it is to many of the LHAs in Fraser Health. As a result of this, the spread of COVID-19 in Burnaby may be influenced more by the number of cases in Vancouver than in other regions in Fraser Health, e.g. Hope, Chilliwack and Agassize/Harrision. In the example presented here, we explicitly incorporated only the time-series mobility data into the disease transmission rate and accounted for other factors that affect disease transmission through a piece-wise parameter. An interesting extensions

this work would be to incorporate the data for other factors that affect disease transmission explicitly into the model. This way, the effect of each factor on disease spread can easily be assessed. Another extension of this model is to include vaccination and the variants of concern of COVID-19. Since mobility rate varies by age, an exciting extension of this work would be to stratify the population of each region by age. This way, in addition to assessing the impact of mobility on disease spread, it would also be possible to assess the contribution of each age group to disease spread.

Supporting information

S1 Text. Supplementary methods and results. This document contains more details of the methods and results. (PDF)

Acknowledgments

We acknowledge the support of TELUS and Michelle Spencer in making the mobile device data available.

Author Contributions

Conceptualization: Sarafa A. Iyaniwura, Notice Ringa, Michael A. Irvine, Michael Otterstatter.

Data curation: Sarafa A. Iyaniwura, Notice Ringa, Prince A. Adu.

Formal analysis: Sarafa A. Iyaniwura.

Funding acquisition: Michael Otterstatter.

Investigation: Sarafa A. Iyaniwura, Michael A. Irvine, Michael Otterstatter.

Methodology: Sarafa A. Iyaniwura, Michael A. Irvine, Michael Otterstatter.

Resources: Michael A. Irvine, Michael Otterstatter.

Software: Sarafa A. Iyaniwura.

Supervision: Naveed Z. Janjua, Michael A. Irvine, Michael Otterstatter.

Validation: Sarafa A. Iyaniwura.

Visualization: Sarafa A. Iyaniwura, Sunny Mak.

Writing - original draft: Sarafa A. Iyaniwura, Notice Ringa.

Writing – review & editing: Sarafa A. Iyaniwura, Notice Ringa, Prince A. Adu, Sunny Mak, Naveed Z. Janjua, Michael A. Irvine, Michael Otterstatter.

References

- 1. World Health Organization (WHO). WHO's COVID-19 response timeline; (accessed April 17, 2022).
- World Health Organization (WHO). WHO Coronavirus (COVID-19) Dashboard; (accessed April 17, 2022).
- Karia R, Gupta I, Khandait H, Yadav A, Yadav A. COVID-19 and its modes of transmission. SN comprehensive clinical medicine. 2020; 2(10):1798–1801. https://doi.org/10.1007/s42399-020-00498-4 PMID: 32904860
- 4. Tan SWX, Chia YK, Lee PY, Ng TH, Wong OT, MARIMUTHU K. Air, surface environmental, and personal protective equipment contamination by severe acute respiratory syndrome coronavirus 2 (SARS-

CoV-2) from a symptomatic patient. Jama. 2020; 323(16):1610–1612. https://doi.org/10.1001/jama. 2020.3227 PMID: 32129805

- 5. World Health Organization (WHO). Scientific Brief: Transmission of SARS-CoV-2: implications for infection prevention precautions; (accessed August 3, 2021).
- Mallapaty S. Why does the coronavirus spread so easily between people? Nature. 2020; 579 (7798):183–184. https://doi.org/10.1038/d41586-020-00660-x PMID: 32157230
- Abdelrahman Z, Li M, Wang X. Comparative review of SARS-CoV-2, SARS-CoV, MERS-CoV, and influenza a respiratory viruses. Frontiers in immunology. 2020; p. 2309. <u>https://doi.org/10.3389/fimmu.</u> 2020.552909 PMID: 33013925
- World Health Organization (WHO). Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV); (accessed June 15, 2021).
- 9. World Health Organization (WHO). WHO Director-General's opening remarks at the media briefing on COVID-19. 2020; (accessed June 15, 2021).
- Perra N. Non-pharmaceutical interventions during the COVID-19 pandemic: A review. Physics Reports. 2021; 913:1–52. https://doi.org/10.1016/j.physrep.2021.02.001 PMID: 33612922
- Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. nature. 2020; 585(7825):410–413. https://doi.org/10.1038/ s41586-020-2293-x PMID: 32365354
- Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020; 584(7820):257–261. <u>https:// doi.org/10.1038/s41586-020-2405-7</u> PMID: 32512579
- 13. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. 2020;.
- Bo Y, Guo C, Lin C, Zeng Y, Li HB, Zhang Y, et al. Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. International Journal of Infectious Diseases. 2021; 102:247–253. https://doi.org/10.1016/j.ijid.2020.10.066 PMID: 33129965
- Cho SW. Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden. The Econometrics Journal. 2020; 23(3):323–344. https://doi.org/10.1093/ectj/ utaa025
- Demirgüç-Kunt A, Lokshin M, Torre I. The sooner, the better: The early economic impact of non-pharmaceutical interventions during the COVID-19 pandemic. World Bank Policy Research Working Paper. 2020; (9257).
- Hevia C, Neumeyer A. A conceptual framework for analyzing the economic impact of COVID-19 and its policy implications. UNDP Lac COVID-19 Policy Documents Series. 2020; 1:29.
- Junjiang L, Philippe G. Returning to a normal life via COVID-19 vaccines in the United States: a largescale Agent-Based simulation study. JMIR Medical Informatics. 2021; 9(4):e27419. https://doi.org/10. 2196/27419
- 19. Shayak B, Sharma MM. COVID-19 spreading dynamics with vaccination–allocation strategy, return to normalcy and vaccine hesitancy. medRxiv. 2020;.
- FDA U. FDA takes key action in fight against COVID-19 by issuing emergency use authorization for first COVID-19 vaccine; 2020.
- Thompson MG, Burgess JL, Naleway AL, Tyner HL, Yoon SK, Meece J, et al. Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers eight US locations, December 2020–March 2021. Morbidity and Mortality Weekly Report. 2021; 70 (13):495. https://doi.org/10.15585/mmwr.mm7013e3 PMID: 33793460
- Pilishvili T, Fleming-Dutra KE, Farrar JL, Gierke R, Mohr NM, Talan DA, et al. Interim estimates of vaccine effectiveness of Pfizer-BioNTech and Moderna COVID-19 vaccines among health care personnel —33 US sites, January–March 2021. Morbidity and Mortality Weekly Report. 2021; 70(20):753. https:// doi.org/10.15585/mmwr.mm7020e2 PMID: 34014909
- Self WH, Tenforde MW, Rhoads JP, Gaglani M, Ginde AA, Douin DJ, et al. Comparative effectiveness of Moderna, Pfizer-BioNTech, and Janssen (Johnson & Johnson) vaccines in preventing COVID-19 hospitalizations among adults without immunocompromising conditions—United States, March–August 2021. Morbidity and Mortality Weekly Report. 2021; 70(38):1337. <u>https://doi.org/10.15585/mmwr.</u> mm7038e1 PMID: 34555004
- 24. World Health Organization (WHO). Origin of the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), the virus causing COVID-19; (accessed August 13, 2022).
- 25. World Health Organization (WHO). WHO Timeline—COVID-19; (accessed August 13, 2022).

- World Health Organization (WHO). Report on the Strategic Response to COVID-19 in the WHO African Region; (accessed August 13, 2022).
- Haileamlak A. COVID-19 Pandemic Status in Africa. Ethiopian journal of health sciences. 2020; 30 (5):643. https://doi.org/10.4314/ejhs.v30i5.1 PMID: 33911823
- 28. (AJMC) AJMC. A Timeline of COVID-19 Developments in 2020; (accessed August 13, 2022).
- Spiteri G, Fielding J, Diercke M, Campese C, Enouf V, Gaymard A, et al. First cases of coronavirus disease 2019 (COVID-19) in the WHO European Region, 24 January to 21 February 2020. Eurosurveillance. 2020; 25(9):2000178. <u>https://doi.org/10.2807/1560-7917.ES.2020.25.9.2000178</u> PMID: 32156327
- Sigler T, Mahmuda S, Kimpton A, Loginova J, Wohland P, Charles-Edwards E, et al. The socio-spatial determinants of COVID-19 diffusion: the impact of globalisation, settlement characteristics and population. Globalization and health. 2021; 17(1):1–14. https://doi.org/10.1186/s12992-021-00707-2 PMID: 34016145
- Kraemer MU, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science. 2020; 368(6490):493–497. <u>https://doi.org/10.1126/science.abb4218 PMID</u>: 32213647
- Hâncean MG, Slavinec M, Perc M. The impact of human mobility networks on the global spread of COVID-19. Journal of Complex Networks. 2020; 8(6):cnaa041. <u>https://doi.org/10.1093/comnet/ cnaa041 PMID: 34191993</u>
- Galeazzi A, Cinelli M, Bonaccorsi G, Pierri F, Schmidt AL, Scala A, et al. Human mobility in response to COVID-19 in France, Italy and UK. Scientific reports. 2021; 11(1):1–10. <u>https://doi.org/10.1038/s41598-021-92399-2</u> PMID: 34162933
- Engebretsen S, Engø-Monsen K, Aleem MA, Gurley ES, Frigessi A, De Blasio BF. Time-aggregated mobile phone mobility data are sufficient for modelling influenza spread: the case of Bangladesh. Journal of the Royal Society Interface. 2020; 17(167):20190809. <u>https://doi.org/10.1098/rsif.2019.0809</u> PMID: 32546112
- **35.** Jewell S, Futoma J, Hannah L, Miller AC, Foti NJ, Fox EB. It's complicated: Characterizing the timevarying relationship between cell phone mobility and COVID-19 spread in the US. NPJ digital medicine. 2021; 4(1):1–11. https://doi.org/10.1038/s41746-021-00523-3 PMID: 34707199
- Oliver N, Lepri B, Sterly H, Lambiotte R, Deletaille S, De Nadai M, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle; 2020.
- Xiong C, Hu S, Yang M, Luo W, Zhang L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. Proceedings of the National Academy of Sciences. 2020; 117(44):27087–27089. https://doi.org/10.1073/pnas.2010836117
- Yechezkel M, Weiss A, Rejwan I, Shahmoon E, Ben-Gal S, Yamin D. Human mobility and poverty as key drivers of COVID-19 transmission and control. BMC public health. 2021; 21(1):1–13. <u>https://doi.org/ 10.1186/s12889-021-10561-x PMID: 33765977</u>
- Alessandretti L. What human mobility data tell us about COVID-19 spread. Nature Reviews Physics. 2022; 4(1):12–13. https://doi.org/10.1038/s42254-021-00407-1 PMID: 34877474
- Calvetti D, Hoover AP, Rose J, Somersalo E. Metapopulation network models for understanding, predicting, and managing the coronavirus disease COVID-19. Frontiers in Physics. 2020; 8:261. <u>https:// doi.org/10.3389/fphy.2020.00261</u>
- Zhang S, Yang Z, Wang M, Zhang B. "Distance-Driven" Versus "Density-Driven": Understanding the Role of "Source-Case" Distance and Gathering Places in the Localized Spatial Clustering of COVID-19 —A Case Study of the Xinfadi Market, Beijing (China). GeoHealth. 2021; 5(8):e2021GH000458. https:// doi.org/10.1029/2021GH000458 PMID: 34466764
- 42. Arenas A, Cota W, Gómez-Gardenes J, Gómez S, Granell C, Matamalas JT, et al. A mathematical model for the spatiotemporal epidemic spreading of COVID19. MedRxiv. 2020;.
- **43.** Costa GS, Cota W, Ferreira SC. Metapopulation modeling of covid-19 advancing into the countryside: an analysis of mitigation strategies for Brazil. medRxiv. 2020;.
- Hadjidemetriou GM, Sasidharan M, Kouyialis G, Parlikad AK. The impact of government measures and human mobility trend on COVID-19 related deaths in the UK. Transportation research interdisciplinary perspectives. 2020; 6:100167. https://doi.org/10.1016/j.trip.2020.100167 PMID: 34173458
- 45. Wei Y, Wang J, Song W, Xiu C, Ma L, Pei T. Spread of COVID-19 in China: analysis from a city-based epidemic and mobility model. Cities. 2021; 110:103010. <u>https://doi.org/10.1016/j.cities.2020.103010</u> PMID: 33162634
- 46. Coletti P, Libin P, Petrof O, Willem L, Abrams S, Herzog SA, et al. A data-driven metapopulation model for the Belgian COVID-19 epidemic: assessing the impact of lockdown and exit strategies. BMC infectious diseases. 2021; 21(1):1–12. https://doi.org/10.1186/s12879-021-06092-w PMID: 34053446

- 47. Danon L, Brooks-Pollock E, Bailey M, Keeling M. A spatial model of CoVID-19 transmission in England and Wales: early spread and peak timing. MedRxiv. 2020;.
- Oka T, Wei W, Zhu D. The effect of human mobility restrictions on the COVID-19 transmission network in China. PloS one. 2021; 16(7):e0254403. <u>https://doi.org/10.1371/journal.pone.0254403</u> PMID: 34280197
- Zhang M, Wang S, Hu T, Fu X, Wang X, Hu Y, et al. Human mobility and COVID-19 transmission: a systematic review and future directions. Annals of GIS. 2022; p. 1–14.
- Iyaniwura SA, Falcão RC, Ringa N, Adu PA, Spencer M, Taylor M, et al. Mathematical modeling of COVID-19 in British Columbia: an age-structured model with time-dependent contact rates. Epidemics. 2022; 39:100559. https://doi.org/10.1016/j.epidem.2022.100559 PMID: 35447505
- Conway JM, Tuite AR, Fisman DN, Hupert N, Meza R, Davoudi B, et al. Vaccination against 2009 pandemic H1N1 in a population dynamical model of Vancouver, Canada: timing is everything. BMC public health. 2011; 11(1):1–14. https://doi.org/10.1186/1471-2458-11-932
- Lavine JS, Bjornstad ON, Antia R. Immunological characteristics govern the transition of COVID-19 to endemicity. Science. 2021; 371(6530):741–745. https://doi.org/10.1126/science.abe6522 PMID: 33436525
- Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The lancet infectious diseases. 2020; 20 (5):553–558. https://doi.org/10.1016/S1473-3099(20)30144-4 PMID: 32171059
- Anderson SC, Edwards AM, Yerlanov M, Mulberry N, Stockdale J, Iyaniwura SA, et al. Estimating the impact of COVID-19 control measures using a Bayesian model of physical distancing. medRxiv. 2020;.
- 55. Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. New England Journal of Medicine. 2020; 382(12):1177–1179. https://doi.org/10.1056/NEJMc2001737 PMID: 32074444
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. New England journal of medicine. 2020;. <u>https://doi.org/10.1056/NEJMoa2001316 PMID: 31995857</u>
- 57. Tindale L, Coombe M, Stockdale JE, Garlock E, Lau WYV, Saraswat M, et al. Transmission interval estimates suggest pre-symptomatic spread of COVID-19. MedRxiv. 2020;.
- Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance. 2020; 25(17):2000257. https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257 PMID: 32372755
- Taylor PJ, Openshaw S. Distance decay in spatial interactions. In: Concepts and techniques in modern geography. Citeseer; 1975.
- 60. BCStats. Population Estimates & Projections for British Columbia; (accessed August 11, 2022).
- 61. Government BC. Health Authority Boundaries; (accessed March 8, 2023).
- 62. Government BC. Open Government Licence—British Columbia; (accessed March 8, 2023).
- 63. Earth N. 1:110m Cultural Vectors; (accessed March 8, 2023).
- 64. Earth N. Terms of Use; (accessed March 8, 2023).
- Health F. About Fraser Health; (accessed September 18, 2022).
- 66. TELUS. Telus Insight; (accessed September 10, 2022).
- 67. TELUS. Technology is evolving and so are we; (accessed March 2, 2023).
- **68.** Ellison AM. Bayesian inference in ecology. Ecology letters. 2004; 7(6):509–520. https://doi.org/10. 1111/j.1461-0248.2004.00603.x
- Von Toussaint U. Bayesian inference in physics. Reviews of Modern Physics. 2011; 83(3):943. https://doi.org/10.1103/RevModPhys.83.943
- 70. Hu B, Gonzales JL, Gubbins S. Bayesian inference of epidemiological parameters from transmission experiments. Scientific reports. 2017; 7(1):1–13. https://doi.org/10.1038/s41598-017-17174-8 PMID: 29196741
- 71. Congdon PD. Applied Bayesian hierarchical methods. Chapman and Hall/CRC; 2010.
- 72. McGlothlin AE, Viele K. Bayesian hierarchical models. Jama. 2018; 320(22):2365–2366. https://doi.org/ 10.1001/jama.2018.17977 PMID: 30535206
- 73. Stan Development Team. RStan: the R interface to Stan; 2020. Available from: http://mc-stan.org/.
- Gelman A, Lee D, Guo J. Stan: A probabilistic programming language for Bayesian inference and optimization. Journal of Educational and Behavioral Statistics. 2015; 40(5):530–543. https://doi.org/10. 3102/1076998615606113

- 75. Team TSD. Variational algorithm for approximate posterior sampling; (accessed May 18, 2021).
- Yao Y, Vehtari A, Simpson D, Gelman A. Yes, but did it work?: Evaluating variational inference. In: International Conference on Machine Learning. PMLR; 2018. p. 5581–5590.
- Magnusson M, Vehtari A, Jonasson J, Andersen M. Leave-one-out cross-validation for Bayesian model comparison in large data. In: International Conference on Artificial Intelligence and Statistics. PMLR; 2020. p. 341–351.
- Jiqiang G, Jonah G, Ben G, Sebastian W. Approximate leave-one-out cross-validation; (Accessed September 18, 2021).
- Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and computing. 2017; 27(5):1413–1432. https://doi.org/10.1007/s11222-016-9709-3
- Watanabe S, Opper M. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of machine learning research. 2010; 11(12).
- Aki V, Jonah G, Mans M, Yuling Y, Paul-Christian B, Topi P, et al. Widely applicable information criterion (WAIC); (Accessed September 18, 2021).
- Hoffman MD, Gelman A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 2014; 15(1):1593–1623.
- **83.** Hoffman MD, Johnson MJ. Elbo surgery: yet another way to carve up the variational evidence lower bound. In: Workshop in Advances in Approximate Bayesian Inference, NIPS. vol. 1; 2016. p. 2.
- 84. Gundersen G. The ELBO in Variational Inference; (accessed June 18, 2021).