COVID-19 Hospital Admissions and Wastewater Data in Canada: A Statistical Analysis

Warsame Yusuf
Public Health Agency of Canada
David Champredon
david.champredon@canada.ca
Public Health Agency of Canada

Research Article

Keywords: SARS-CoV-2, Wastewater-based surveillance, COVID-19 hospitalizations, Hierarchical Bayesian regression

Posted Date: June 13th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6628716/v1

Additional Declarations: No competing interests reported.

COVID-19 Hospital Admissions and Wastewater Data in Canada: A Statistical Analysis

Warsame Yusuf[1] and David Champredon[2*]

[1]Public Health Agency of Canada, 130 Colonnade Road South, Ottawa, K1A 0K9, Ontario, Canada.
[2]Public Health Agency of Canada, 370 Speedvale Avenue West, Guelph, N1H 7M7, Ontario, Canada.

*Corresponding author(s). E-mail(s): david.champredon@canada.ca;
Contributing authors: warsame.yusuf@phac-aspc.gc.ca;

Keywords: SARS-CoV-2, Wastewater-based surveillance, COVID-19 hospitalizations, Hierarchical Bayesian regression

1 Introduction

The SARS-CoV-2 pandemic witnessed the adoption of wastewater-based surveillance (WBS) by many jurisdictions worldwide, as early as the Spring of 2020 [1]. WBS relies on the fact that some pathogens are shed in stools and/or urine by infected individuals. Monitoring the pathogen concentration in wastewater (typically using quantitative polymerase chain reaction (qPCR) techniques) can provide a proxy of prevalence in an entire community [2, 3]. About four years after its large-scale debut, WBS is now deployed in many jurisdiction worldwide, monitoring various pathogens and substances of public-health concern [4]. This popularity stems in part by the ability of WBS to sample entire communities (including asymptomatic infections) at a relatively low cost and to quantify multiple pathogen targets (*e.g.,* SARS-CoV-2, Mpox, influenza, respiratory syncytial virus, measles) from a single sample [5, 6, 7, 8].

However, despite its coverage and flexibility, WBS does not provide a metric that is directly interpretable for public health professionals. Indeed, sewer systems are typically harsh and dynamic environments that can affect the pathogen concentration, yielding measurements displaying a large variance and, potentially, a bias. Hence, to support decision making, WBS needs to be translated into metrics that are directly

relevant to public health. During an epidemic, the burden level on hospitals – a critical component of most health systems – is an important metric. Multiple studies have focused on understanding the link between wastewater signal and hospital admissions. Many of them done during the COVID-19 pandemic found a correlation between the concentration of SARS-CoV-2 RNA in wastewater and COVID-19 hospital admissions with varying lead times [9, 10, 11, 3]. In Canadian studies, the optimal lead time when exploring correlations between SARS-CoV-2 concentration and hospital admissions varied between cities and between circulating variants. [12, 13, 14].

Additionally, a number of studies have used various mathematical models to estimate the number of COVID-19 hospital admissions from SARS-CoV-2 wastewater concentration, using various linear models, such as multivariate, Poisson, mixed-effects, and generalized linear models [11, 15, 16, 17]. Other studies used more advanced models that implement machine learning tools such as artifical neural networks and random forest algorithms [18, 19, 20]. The lag between infection and hospital admission being at least five days on average [21], using WBS to anticipate the burden level on hospitals can provide a key support for public health action.

Although these studies showed that WBS is often correlated to hospital admissions, there are a number of challenges that can affect the relationship. The emergence of new SARS-CoV-2 lineages may impact the amount of viral load shed by infected individuals into the sewer system and may affect the virulence of SARS-CoV-2, impacting the rate of hospitalization in infected individuals [18, 12, 22]. It is still not clear how fecal shedding and disease virulence are related. In studies that predicted hospitalizations from SARS-CoV-2 concentration in wastewater, variants of concerns were often used to split data into different time periods for analysis [12, 11, 13, 23]. Other studies took a more direct approach in incorporating variants in their models, using them as a fixed-effect [17], as a covariate [16], or as a time-varying intercept [14].

As new SARS-CoV-2 lineages continue to emerge, models that predict COVID-19 hospital admissions from wastewater will require to account for a virulence that may be variant-dependant. In this study, we developed a model that estimated COVID-19 hospital admissions from SARS-CoV-2 wastewater concentration in several cities across Canada guided by the presence of variants in each respective community. This model quantifies the severity of circulating SARS-CoV-2 variants by estimating the relationship between community prevalence (informed by wastewater) and acute infections (informed by COVID-19 hospital admissions). It also allows the severity of new lineages to be benchmarked to inform public health action.

2 Methods

2.1 Data

The wastewater data is from the Public Health Agency of Canada wastewater surveillance program. In collaboration with provincial and regional organizations, wastewater samples are collected from treatment plants across Canada. Once collected, samples are shipped to the National Microbiology Laboratory in Winnipeg, Manitoba, where qPCR is performed to measure the concentration of various pathogens such as influenza, respiratory syncytial virus, and SARS-CoV-2. In most cities, samples are

collected twice a week; while samples collected at treatment plants in Winnipeg are collected five times a week. We included in this study, SARS-CoV-2 wastewater concentration data from treatment plants in Vancouver, Edmonton, Regina, Winnipeg, Toronto, Halifax, and St. John's. For comparison with weekly hospital admission counts, we calculated the weekly mean SARS-CoV-2 concentrations at each treatment plant.

In cities with one treatment plant, we inferred that the weekly mean concentration represented SARS-CoV-2 levels for the entire city. To estimate a city-level SARS-CoV-2 concentration in wastewater, (cities usually have multiple wastewater treatment plants), we calculated a population-weighted mean concentration based on the catchment population of each treatment plant.

Anonymized and aggregated hospital admission data were retrieved from the Discharge Abstract Database (DAD), a database from the Canadian Institute for Health Information (CIHI) containing administrative hospital data received directly from inpatient facilities across Canada except Quebec [24]. Using this dataset, weekly counts of admitted patients with an ICD-10 code equal to U071 or U072 (indicating a diagnosis of COVID-19 infection) were retrieved up to March 31, 2024. Along with hospital admission data, the DAD captures the Forward Sortation Area codes (first three characters of postcodes in Canada) of admitted patients. CIHI provided us with weekly hospitalization counts aggregated at the wastewater treatment plant level, and we subsequently aggregated the data at the city level for this study.

To account for variations in city and treatment plant size, hospital admission data was normalized by taking weekly hospitalization counts and dividing by the sum of the corresponding total wastewater catchment population and then multiplied by 100,000 to obtain the number of admissions per 100,000 persons. Wastewater concentration data was not normalized to flow or any other biomarker. This was due to the fact that flow data was not available for all treatment plants included in the study.

The SARS-CoV-2 variant lineage data is taken from sequencing done on clinical samples collected from the COVID-19 National Genomics Database, a Canadian genomics database managed by the Public Health Genomics and Computational and Operational Genomics divisions of the Public Health Agency of Canada, with data provided by provincial and territorial laboratories within the Canadian Public Health Laboratory Network. The lineage data used in this study contains anonymized counts of each variant aggregated weekly at the provincial level. Using these counts, variant proportions were obtained, with a dominant variant determined by the variant with the highest proportion in a week.

To account for multiple sublineages circulating within a given week, many sublineages were merged into grouped variant lineage by their Pango lineage clade [25]. Using the `CCT` R package [26], we retrieved a list of lineages and their respective child sublineages from the Pango Designation GitHub repository [27]. From this list, we were able to merge sublineages found in the observed weekly variant data. For example, sublineages such as `JN.1.1, JN.1.1.5, JN.1.1.6` were merged into `JN.1` as they were reported be part of phylogenetic clade `24A`. On the other hand, sublineages such as `JN.1.11.1, LP.1, LP.1.2` were grouped into `JN.1.11.1` as they were reported to be part of clade `24B`.

2.2 Statistical models

We used a hierarchical model to estimate COVID-19 hospital admissions from the concentration of SARS-CoV-2 genes present in wastewater. In each province, the emergence of B.1.1.7 as the dominant variant ("Alpha" variant) marked the starting point in each of the data sources: wastewater concentration, hospital admissions, and variant data before the first date B.1.1.7 was declared dominant in provincial variant data were excluded from the analysis.

The hospital admission data we had access to was not linked to genetic sequencing. Hence, we could not associate unequivocally a COVID-19 hospitalization with a specific SARS-CoV-2 variant. We made the assumption that if a variant is observed more often than other variants during a time frame, hospital admissions observed during that time frame can be attributed to this "dominant" variant. Below, we define variant dominance based on a sustained high level of circulation.

Only variants identified as dominant were selected to be included in the statistical model. A variant was dominant if its abundance was larger than a threshold of 60%. Its period of dominance was defined by the start and end weeks for the interval of time where it sustained its abundance above that threshold. Moreover, dominant variants that did not exceed this threshold for 75% of the time within a minimum 8-week window were excluded from the model.

The hierarchical model performs a linear regression with three levels:

- the top "universal" level, which estimates hospital admissions from wastewater concentration across all variants and geographical locations
- the variant level which estimates hospital admissions from wastewater within each (dominant) variant period
- the geographical location level which estimates hospital admissions from wastewater for each variant in each city

The basic linear regression formula used in each level can simply be represented with the expression $H \sim mW + b$, where H represents the logarithm of hospital admissions, m represents the slope coefficient for the logarithm of wastewater concentration W, and b represents the intercept. At the universal level, the formula to estimate H at time t is:

$$H(t) \sim mW(t) + b \tag{1}$$

At the variant level, the formula to estimate the log number of hospital admissions H_v associated with variant v at time t is:

$$H_v(t) \sim m_v W_v(t) + b_v \tag{2}$$

with $v = 1, \ldots, V$ where V is the total number of variants identified in the data set. Finally, at the location level, the formula to estimate the log number of hospital admissions $H_{v,l}$ associated with variant v at the geographical location l at time t is:

$$H_{v,l}(t) \sim m_{v,l} W_{v,l}(t) + b_{v,l} \tag{3}$$

with $l = 1, \ldots, N_v$ where N_v is the number of geographical locations where variant v was observed.

The observed data is at the variant and location level, hence the data directly informs estimates for $m_{v,l}$ and $b_{v,l}$, which then inform – thanks to the hierarchical structure – estimates for variant-level parameters m_v, b_v, and ultimately "universal" parameters m, and b. A diagram illustrating the hierarchical relationship is shown in Figure 1.

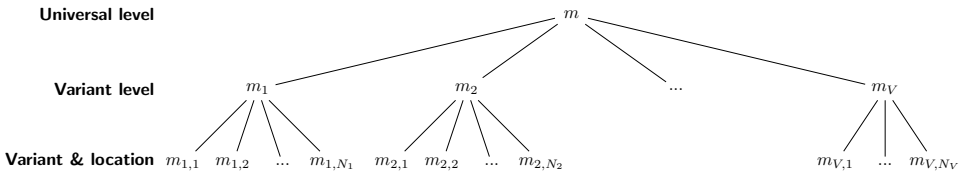

Fig. 1 Hierarchical diagram for the slope parameter m. The top level of the diagram represents the "universal" slope m of the linear regression linking any COVID-19 hospital admission with any SARS-CoV-2 variant. At the middle level, the parameter m_v represents the slope for variant v at any geographical location. At the bottom level, parameter $m_{v,\ell}$ represents the slope for variant v at the geographical location ℓ. Data may not be available for all geographical locations for a given variant, hence the number of locations having data for variant i can vary across variants and is noted N_i. Similar diagrams can be drawn for all other hierarchical parameters (*e.g.,* the intercept b).

The hierarchical model was implemented in R using the `nimble` package (version 1.2.1) [28, 29] and parameter estimation was done with the Hamiltonian Monte Carlo Markov Chain (MCMC) algorithm using the R package `nimbleHMC` (version 0.2.3). To test the validity of our implementation, synthetic data was simulated using predefined values for m and b at each hierarchical level, and the same number of variants and locations as found in the observation dataset. This created a dataframe containing simulated values of $H_{v,l}$ and $W_{v,l}$ for nine hypothetical variants and seven hypothetical geographical locations. Results of this preliminary check are shown in Appendix A and indicated no apparent concerns regarding the model implementation as the inference successfully estimated the parameter values used for the simulated data.

When running the inference on the real data, the MCMC algorithm was iterated 500,000 times with 250,000 iterations used as "burn-in", and executed on 4 chains.

In addition to the Hamiltonian MCMC algorithm, a naive linear regression was conducted to estimate the slope for each variant at each location. This served as an addiotional check of the MCMC inference (inferred slopes from the naive linear model shouldn't be disproportionaly different).

3 Results

3.1 Study population

Wastewater concentration and hospital admission data was collected from seven cities, representing a total community population of approximately 9.7 million, or 26% of

the Canadian population [30]. Community populations were derived from the catchment sizes of individual wastewater treatment sites, ranging from 42,981 to 1,447,246 persons.

Figure 2 shows the data used by the hierarchical model. In this figure, hospital admissions are plotted against observed wastewater concentration across various variant periods in the seven cities included in the study. As data was not available for all geographical locations for a given variant, some of the subplots of the figure are blank.

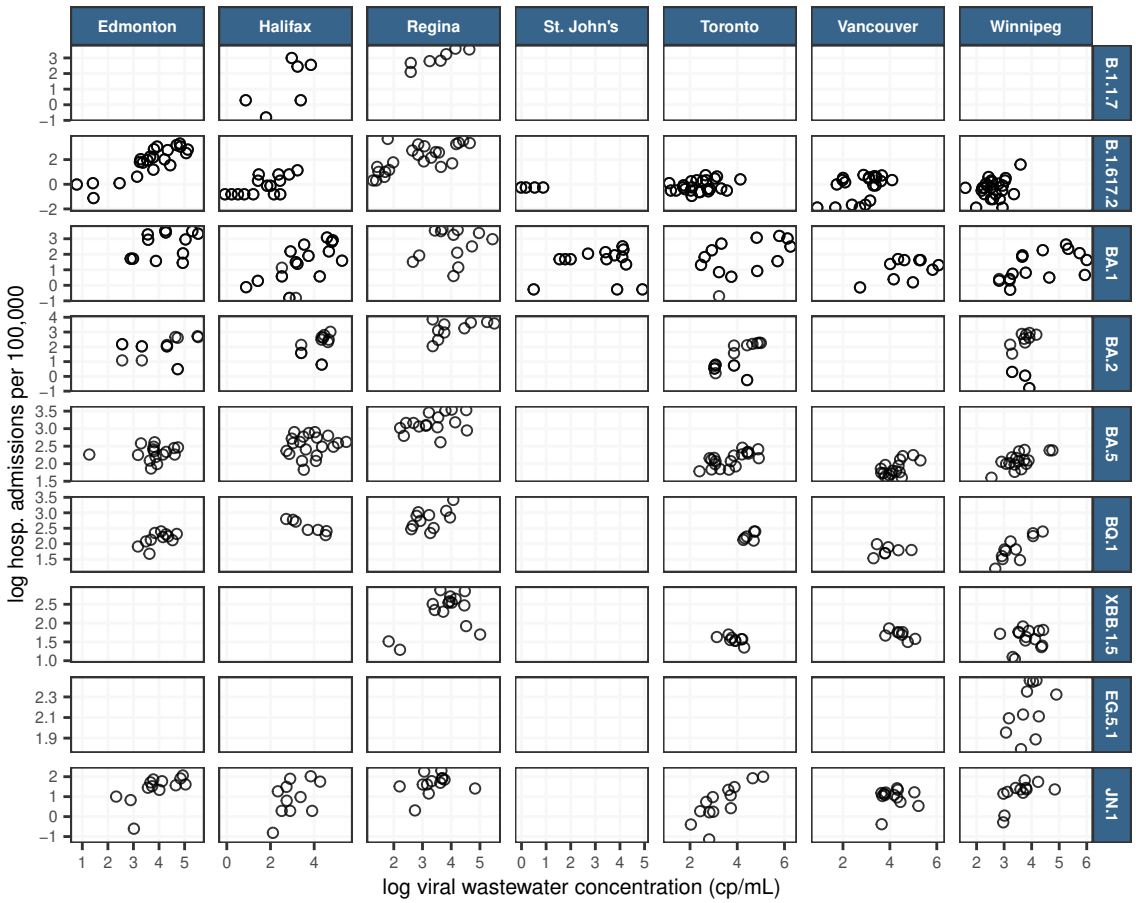

Fig. 2 Log-transformed COVID-19 hospital admissions and SARS-CoV-2 concentration in wastewater data used by the hierarchical model. Each subplot represents a variant at a given geographic location. Some locations have empty variant plots as no dominant variant was identified for inclusion. The X-axis represents weekly averaged wastewater concentration, measured in gene copies per mL (cp/mL). Y-axis represents weekly averaged hospital admissions measured in admissions per 100,000 persons.

The start and end dates determined for each variant at each geographic location are shown in Table B2. The date ranges represent the maximum period where the dominant period met the criteria specified in the methods, and there was sufficient wastewater and hospital admissions data to conduct a regression.

3.2 Hierarchical model

The Hamiltonian MCMC algorithm was run with 500,000 iterations on four chains, with the first 250,000 iterations used as a "burn-in", to estimate the values of model parameters m and b at each hierarchical level. To test if the MCMC algorithm had converged sufficiently, a Gelman-Rubin diagnostic was conducted to test the convergence of the universal and variant level parameters across the four chains. Figure E6 shows that most parameters had a potential scale reduction factor (PSRF) below 1.001.

The mean estimates for slope and intercept at the universal level (on the log scale) were inferred to be 0.353 and 0.471 respectively. Table 1 shows the estimates of m and b at the universal and variant levels.

Table 1 Inferred slope (m) and intercept (b) estimates from the Hamiltonian MCMC algorithm at the universal and variant levels. Algorithm was executed across 4 chains using 500,000 iterations and 250,000 iterations used as "burn-in". Values in brackets represent the 95% credible interval.

Variant	**slope** m	**intercept** b
B.1.1.7	0.406 [0.262; 0.669]	0.502 [-0.343; 1.353]
B.1.617.2	0.327 [0.219; 0.425]	-0.163 [-0.71; 0.396]
BA.1	0.351 [0.252; 0.457]	0.315 [-0.235; 0.853]
BA.2	0.38 [0.256; 0.551]	0.458 [-0.264; 1.159]
BA.5	0.31 [0.138; 0.432]	1.035 [0.299; 1.826]
BQ.1	0.344 [0.187; 0.492]	0.783 [0.096; 1.561]
XBB.1.5	0.327 [0.153; 0.465]	0.536 [-0.182; 1.335]
EG.5.1	0.355 [0.181; 0.546]	0.579 [-0.358; 1.635]
JN.1	0.371 [0.25; 0.533]	-0.07 [-0.81; 0.571]
Universal	0.353 [0.257; 0.459]	0.471 [0; 0.99]

To represent the results of the log-regressions in a more straightforward epidemiological way, we show the hospital admissions estimates given viral concentration in wastewater (not log transformed, but as they would be directly observed) in Figure 3. The hospital admissions values were calculated using the formula $h = e^b \times w^m$ where $h = \log(H)$ and $w = \log(W)$. The values for m and b were sampled from their posterior distributions and the colored area in Figure 3 illustrates different quantiles intervals (from light to dark, see figure caption for details).

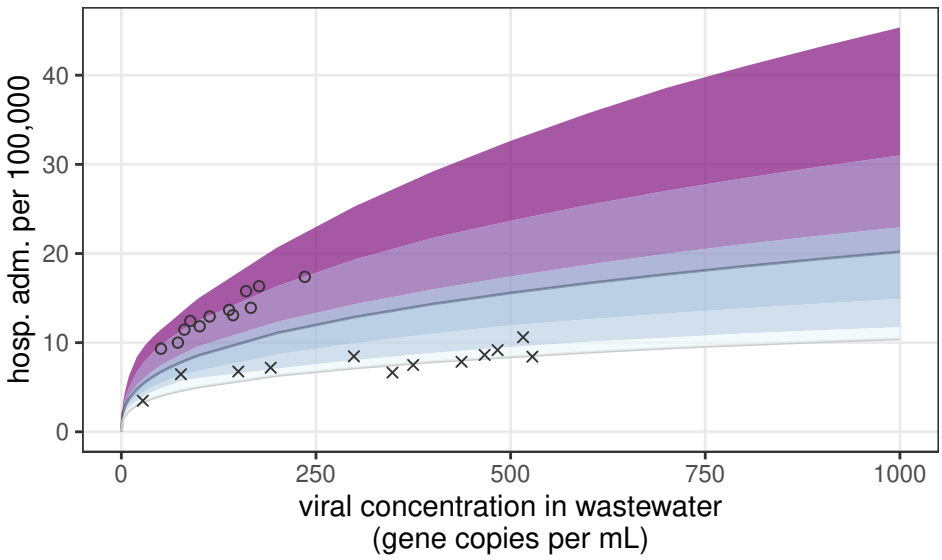

Fig. 3 Estimated SARS-CoV-2 hospital admissions per viral concentration of wastewater. Hospital admissions is measured per 100,000 persons. Different colours represent hospital admissions at different quantiles of the credible interval of the sampled posterior distribution of m and b, with darker colours representing the upper range of the credible interval. The lightest region represents the 2.5%-10% quantile, then the following regions indicate 10%-25%, 25%-50%, 50%-75%, 75%-90%, and 90%-97.5% for the darkest colour. Open circles represents an hypothetical new variant wave that could be categorized as a high-severity variant, whereas crosses represent another hypothetical low-severity new variant.

3.3 Sensitivity analyses and comparison to a simple linear regression

To test the robustness of our inference using the Hamiltonian MCMC algorithm, a number of sensitivity analyses were conducted to assess the impact of changing the dominance definition. The results of this sensitivity analysis are shown in Appendix D. Overall, the estimated slopes and intercepts at the universal and variant levels are comparable to what was found.

A naive linear regression was conducted to estimate, independently, the slope and intercept for each variant/location pair. Its results are shown in Appendix E and the differences with the Bayesian hierarchical model are as expected (estimates at the variant/location level from the hierarchical model are "pulled" to the universal level whereas estimates from the independent linear regressions are more scattered).

4 Discussion

Despite many advantages of WBS, translating viral concentrations measured in wastewater into practical clinical outcomes (*e.g.,* number of infections, hospitalizations) remains a challenge because of various factors, including environmental and in-sewer dynamics.

Our study, using data from seven large urban centres representing 26% of the Canadian population, inferred the relationship across locations and variants that circulated between late 2020 and March 2024, drawing information from multiple variants at different geographic locations. Looking forward in time, our study can also provide an expected range of COVID-19 hospital admissions for a given SARS-CoV-2 concentration level in wastewater and virulence of emerging variants. Indeed, estimates of the hierarchical parameters provide watermarks of severity levels for any future (re)emerging SARS-CoV-2 variants. The estimates of the "universal" regression parameters (slope m and intercept b) have a joint distribution from which we can extract confidence intervals to flag an unusually high virulence that would warrant informing local public authorities. For example, as a new variant emerges, data from wastewater and clinical surveillance can be overlayed on Figure 3 to categorize its population-level clinical severity by comparing the number of hospital admissions to wastewater concentrations, a proxy for the community-wide prevalence because it includes asymptomatic and mild infections. The open circles in Figure 3 indicate a hypothetical high-severity new variant (paired data would be obtained every week from the clinical and wastewater surveillance), whereas the crosses on the same figure show another hypothetical lower-severity variant. As future waves caused by new variants unfold, we can feed that data to the hierarchical model in order to determine updated levels for the severity categories for the next wave. This is a novel approach to inferring virulence of a SARS-CoV-2 variant as it gauges its virulence using wastewater surveillance that includes asymptomatic and mild infections, as opposed to relying solely on clinical information (*i.e.,* test positivity rates, hospital admissions).

Values of the regression intercept (parameter b) to infer hospital admissions represent the "baseline" hospitalization level for COVID-19 in a given location, that is, the number of hospital admissions when no SARS-CoV-2 transmission is apparently circulating. We found that the values of the intercept vary at both the variant and geographic level (Table 1). This is not surprising for two reasons. Firstly, delivery of healthcare in Canada is managed by provincial or territorial authorities, meaning that each province or territory has its own health care infrastructure and resource use. While the model estimates COVID-19 hospital admissions per capita to account for differences in population between locations, the varying number of available beds per capita and admission policies between provinces and territories would explain differing baseline levels. Secondly, as the hierarchical model does not draw data from variants equally between geographic locations (due to restrictions in variant dominance for the inclusion criteria), the weight of each location varies between variants depending on how much data is drawn (if any, at all) for each variant.

In this study we selected B.1.1.7 as our starting variant in the model. Commonly known as the "Alpha" variant, B.1.1.7 was one of the first publicized variants of concern that emerged in late 2020 across many countries [31]. Although other SARS-CoV-2 lineages were present in sample data prior its emergence, variants of concern became more clearly defined and classified after the presence of B.1.1.7 in sequencing data.

In this study, there were variants that did not meet the inclusion criteria in several geographic locations, while some variants met the inclusion criteria in virtually

all locations. These differences could suggest varying levels of variant invasion in the community. Variants such as B.1.617.2 (commonly known as Delta) and BA.1 (commonly known as the first Omicron variant) were well known variants of concern, and were included at all locations in our model. While increasing genetic diversity may have contributed to later variants not being included at all locations, there were still newer variants (such as BQ.1 and JN.1) that were included in most locations. In these two examples, they were included in all locations except St. John's. Given that St. John's variant data comes from the least populous province in our model in which the fewest samples were collected, it can still be inferred that these variants still had a high level of invasion in the community.

While other studies have examined the virulence of variants when predicting hospital admissions from SARS-CoV-2 wastewater concentration, this study provides new findings. Firstly, this study examines variants beyond the first and second waves of Omicron observed in late 2021 and early 2022 by classifying the virulence of BA.1, BA.2, and the subsequent lineages that emerged after up to JN.1 observed in early 2024. Previous studies have either captured up to the first Omicron wave [11, 16], or BA.2 [12, 13], or have classified data after the first Omicron wave as "post-Omicron" [23]. Secondly, the use of a hierarchical model generates a global range of SARS-CoV-2 virulence across different variants, allowing the virulence of emerging variants to be benchmarked. Although the World Health Organization has declared it to no longer be an international public health emergency since May 2023 [32], COVID-19 continues to be a public health issue with new mutations and waning of immunity giving rise to waves of COVID-19 hospitalizations. Using the hierarchical model in this study can allow researchers to classify the virulence of a new variant and help public health officials prepare hospital resources use accordingly.

In our study, we found that B.1.1.7 had a higher slope value (m) compared to B.1.617.2, suggesting higher virulence. This is contrary to previous studies, which have shown that mutations in its spike protein have led B.1.617.2 to be more transmissive and virulent than B.1.1.7 [33, 34]. Possible explanations for this discrepancy include the limited number of geographic locations for B.1.1.7 included in the model contributing a the wider uncertainty in the estimates for the slope m, and the introduction of vaccination campaigns in many Canadian cities during the B.1.617.2 period [12] that may have confounded the relationship between SARS-CoV-2 wastewater concentration and hospital admissions. Conversely, waning immunity from vaccination is a possible explanation as to why BA.1 and BA.2 had higher slope values than B.1.617.2 despite other studies showing higher virulence for B.1.617.2 [35, 36]. When examining newer Omicron variants to earlier lineages, we found that BA.5 had a lower virulence than BA.1 and BA.2. While studies in other countries found similar virulence between these variants [37, 38], a Canadian study found a higher risk of hospitalization in patients infected with BA.5 [39]. While BQ.1 had a slightly lower regression slope than BA.1 and BA.2, its slope was higher than BA.5. This is somewhat in line with existing literature that suggests that BQ.1 exhibits higher escape neutralisation than BA.5, but also higher escape neutralisation than BA.1 and BA.2 [40, 41]. JN.1 had a slightly higher regression slope than XBB.1.5, which was higher than EG.5.1. This is contrary to results found in previous studies [42, 43]. A possible explanation

could be that increased transmissibility and immune escape [44] allowed more data for JN.1 to be included than EG.5.1 and XBB.1.5, giving a more reliable estimate of its virulence. Decreased vaccine effectiveness against JN.1 may have also confounded JN.1's virulence estimate [45, 46].

Although this study develops a well-defined model to estimate COVID-19 hospital admissions from viral wastewater concentration, there are a number of limitations.

It is unclear if differences in fecal shedding dynamics had played a role in the differences in slopes (m) and intercepts (b) estimates between the variants. We interpret the slopes of each variant as an indication of virulence at the whole population level (*i.e.,* accounting for asymptomatic/mild infections). Variants with higher slopes suggest higher virulence, meaning that an infected individual is more likely to have a severe infection (and hence be hospitalized) than with variants associated with a lower regression slope. This interpretation assumes that the rate of fecal shedding, the amount of viral SARS-CoV-2 RNA that an individual sheds through their stool during their infection, remains similar across different variants. Research on fecal shedding dynamics between variants is limited, but could provide an alternate interpretation to varying slopes found in this model.

This study used a restrictive inclusion criteria when selecting data in the model. This was done to best ascertain hospital admissions to circulating variants. In the absence of genomic sequencing of admitted patients, ascertaining SARS-CoV-2 hospital admissions to variants can be difficult, particularly during periods of competing dominance. To define the time windows of each variant at each location, we examined clinical sample counts collected at the provincial level that corresponded to each city in the study. Sample counts varied between provinces. Larger provinces, such as Ontario and British Columbia, reported on average 765 and 1114 weekly samples respectively. Smaller provinces, such as Nova Scotia and Newfoundland and Labrador, reported on average 45 and 21 samples weekly. During the development of the model, the Shannon Index [47] was first attempted to establish variant dominance but as genetic drift increased over time (particularly after the emergence of BA.1), it became difficult to define variant periods. As such, a minimum threshold of 60% was selected to best ascertain hospital admissions to a particular variant in a community without being too restrictive in our variant data windows. But, lacking access to hospital admissions linked to genomic analysis, our identification of dominant variants, which relies on dominance for a sustained period of time, may be erroneous.

Although hospital admissions were normalized by population, wastewater concentration was not normalized by flow. This was in part due to the fact that not all cities included in the study had flow data available. Preliminary analyses had shown that flow-normalization of the cities with available flow data did not substantially impact the outcome of this study.

The model used in this study did not take take into consideration vaccination coverage. In many studies, vaccination has shown to be effective in reducing COVID-19 hospital admissions [48, 49, 50]. In Canada, SARS-CoV-2 vaccination coverage varied during the course of the COVID-19 pandemic. In 2021, following the rollout of SARS-CoV-2 vaccines, vaccination coverage was 80% [51]. By July 2024, however, only 3.9% of Canadians met COVID-19 vaccination recommendations at that time, which was to

receive a number of XBB.1.5 vaccine (the latest vaccine at the time) doses depending on their age group [52]. Varying vaccination coverage between different time periods may have confounded the impact of variants on SARS-CoV-2 hospital admissions and possibly on fecal shedding given that vaccination may reduce the viral load of an infectious person [53], and that vaccine effectiveness may vary between variants [54, 46].

This study did not take into consideration precipitation and degredation in the sewer system that may lead to potential losses of viral wastewater concentration [55, 56]. Future studies should take into account the viral fate of SARS-CoV-2 in the sewer system.

Despite its limitations, the hierarchical model used in this study provides a statistical framework for inferring COVID-19 hospitalizations from SARS-CoV-2 wastewater concentration. Unlike a simple independent linear regression, a hierarchical model can draw information from multiple sources with varying amounts of data to generate a weighted estimate of an outcome of interest. While a simple linear regression can provide estimates of variant virulence at a particular location for a specific variant, our hierarchical model provides a weighted estimate of a variant's virulence across all locations and past variants. That weighted estimate can then be used to provide, for an emerging variant, an estimate of SARS-CoV-2 virulence at the whole population level and categorize its severity level (Figure 3).

5 Conclusion

This study provides a modelling framework to assess the relationship between COVID-19 hospital admissions and SARS-CoV-2 wastewater concentration across different variants and large urban centres in Canada. The statistical inference can be used to analyze the virulence of past SARS-CoV-2 variants and also to categorize the virulence of emerging variants to support public health action.

6 Declarations

6.1 Ethics approval and consent to participate

Because the data used was publicly available through a mechanism set out by legislation or regulation and is protected by law, this study did not require a research ethics review in accordance to article 2.2(a) of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans [57] .

6.2 Consent for publication

Not applicable.

6.3 Availability of data and materials

The wastewater data is publicly available on the Public Health Agency of Canada website https://health-infobase.canada.ca/wastewater/. We are not authorized to share the hospital admission data, however, for the sake of allowing others to easily replicate our work, we provide perturbed synthetic hospital admission data that have the same characteristics as the real data set. The computer code for the statistical model is available at https://github.com/phac-nml-phrsd/hosp-ww-covid-canada.

6.4 Competing interests

None.

6.5 Funding

None.

6.6 Authors' contributions

All authors (W.Y. and D.C.) conceptualized the modelling, implemented the computer code, interpreted the results, wrote the main manuscript and approved the final manuscript.

6.7 Acknowledgments

None.

References

[1] Gertjan Medema, Leo Heijnen, Goffe Elsinga, Ronald Italiaander, and Anke Brouwer. Presence of sars-coronavirus-2 rna in sewage and correlation with reported covid-19 prevalence in the early stage of the epidemic in the netherlands. *Environmental Science & Technology Letters*, 7(7):511–516, 2020.

[2] Maria L. Daza-Torres, J. Cricelio Montesinos-López, Minji Kim, Rachel Olson, C. Winston Bess, Lezlie Rueda, Mirjana Susa, Linnea Tucker, Yury E. García, Alec J. Schmidt, Colleen C. Naughton, Brad H. Pollock, Karen Shapiro, Miriam Nuño, and Heather N. Bischel. Model training periods impact estimation of COVID-19 incidence from wastewater viral loads. *Science of The Total Environment*, 858:159680, February 2023.

[3] Ramanath Majumdar, Biruhalem Taye, Corey Bjornberg, Matthew Giljork, Danielle Lynch, Fadumasahra Farah, Intisar Abdullah, Kristin Osiecki, Iris Yousaf, Aaron Luckstein, Wendy Turri, Priya Sampathkumar, Ann M. Moyer, Benjamin R. Kipp, Roberto Cattaneo, Caroline R. Sussman, and Chanakha K. Navaratnarajah. From pandemic to endemic: Divergence of COVID-19 positive-tests and hospitalization numbers from SARS-CoV-2 RNA levels in wastewater of Rochester, Minnesota. *Heliyon*, 10(6):e27974, March 2024.

[4] Surabhi Singh, Amina Ismail Ahmed, Sumayya Almansoori, Shaikha Alameri, Ashraf Adlan, Giovanni Odivilas, Marie Anne Chattaway, Samara Bin Salem, Grzegorz Brudecki, and Wael Elamin. A narrative review of wastewater surveillance: pathogens of concern, applications, detection methods, and challenges. *Frontiers in Public Health*, 12:1445961, 2024.

[5] Eline F. De Jonge, Céline M. Peterse, Jaap M. Koelewijn, Anne-Merel R. Van Der Drift, Rudolf F.H.J. Van Der Beek, Erwin Nagelkerke, and Willemijn J. Lodder. The detection of monkeypox virus DNA in wastewater samples in the Netherlands. *Science of The Total Environment*, 852:158265, December 2022.

[6] Elisabeth Mercier, Patrick M. D'Aoust, Ocean Thakali, Nada Hegazy, Jian-Jun Jia, Zhihao Zhang, Walaa Eid, Julio Plaza-Diaz, Md Pervez Kabir, Wanting Fang, Aaron Cowan, Sean E. Stephenson, Lakshmi Pisharody, Alex E. MacKenzie, Tyson E. Graber, Shen Wan, and Robert Delatolla. Municipal and neighbourhood level wastewater surveillance and subtyping of an influenza virus outbreak. *Scientific Reports*, 12(1):15777, September 2022.

[7] Alexandria B Boehm, Bridgette Hughes, Dorothea Duong, Vikram Chan-Herur, Anna Buchman, Marlene K Wolfe, and Bradley J White. Wastewater concentrations of human influenza, metapneumovirus, parainfluenza, respiratory syncytial virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: A surveillance study. *The Lancet Microbe*, 4(5):e340–e348, May 2023.

[8] Emalie K Hayes, Madison T Gouthro, Jason J LeBlanc, and Graham A Gagnon. Simultaneous detection of sars-cov-2, influenza a, respiratory syncytial virus, and measles in wastewater by multiplex rt-qpcr. *Science of the Total Environment*, 889:164261, 2023.

[9] Qingyu Zhan, Kristina M. Babler, Mark E. Sharkey, Ayaaz Amirali, Cynthia C. Beaver, Melinda M. Boone, Samuel Comerford, Daniel Cooper, Elena M. Cortizas, Benjamin B. Currall, Jonathan Foox, George S. Grills, Erin Kobetz, Naresh

Kumar, Jennifer Laine, Walter E. Lamar, Alejandro M.A. Mantero, Christopher E. Mason, Brian D. Reding, Maria Robertson, Matthew A. Roca, Krista Ryon, Stephan C. Schürer, Bhavarth S. Shukla, Natasha Schaefer Solle, Mario Stevenson, John J. Tallon Jr, Collette Thomas, Tori Thomas, Dušica Vidović, Sion L. Williams, Xue Yin, and Helena M. Solo-Gabriele. Relationships between SARS-CoV-2 in Wastewater and COVID-19 Clinical Cases and Hospitalizations, with and without Normalization against Indicators of Human Waste. *ACS ES&T Water*, 2(11):1992–2003, November 2022.

[10] Tatjana Baldovin, Irene Amoruso, Marco Fonzo, Chiara Bertoncello, Vanessa Groppi, Gisella Pitter, Francesca Russo, and Vincenzo Baldo. Trends in SARS-CoV-2 clinically confirmed cases and viral load in wastewater: A critical alignment for Padua city (NE Italy). *Heliyon*, 9(10):e20571, October 2023.

[11] Hannes Schenk, Petra Heidinger, Heribert Insam, Norbert Kreuzinger, Rudolf Markt, Fabiana Nägele, Herbert Oberacher, Christoph Scheffknecht, Martin Steinlechner, Gunther Vogl, Andreas Otto Wagner, and Wolfgang Rauch. Prediction of hospitalisations based on wastewater-based SARS-CoV-2 epidemiology. *Science of The Total Environment*, 873:162149, May 2023.

[12] Nada Hegazy, Aaron Cowan, Patrick M. D'Aoust, Élisabeth Mercier, Syeda Tasneem Towhid, Jian-Jun Jia, Shen Wan, Zhihao Zhang, Md Pervez Kabir, Wanting Fang, Tyson E. Graber, Alex E. MacKenzie, Stéphanie Guilherme, and Robert Delatolla. Understanding the dynamic relation between wastewater SARS-CoV-2 signal and clinical metrics throughout the pandemic. *Science of The Total Environment*, 853:158458, December 2022.

[13] Lydia Cheng, Hadi A. Dhiyebi, Monali Varia, Kyle Atanas, Nivetha Srikanthan, Samina Hayat, Heather Ikert, Meghan Fuzzen, Carly Sing-Judge, Yash Badlani, Eli Zeeb, Leslie M. Bragg, Robert Delatolla, John P. Giesy, Elaine Gilliland, and Mark R. Servos. Omicron COVID-19 Case Estimates Based on Previous SARS-CoV-2 Wastewater Load, Regional Municipality of Peel, Ontario, Canada. *Emerging Infectious Diseases*, 29(8), August 2023.

[14] K. Ken Peng, Elizabeth M. Renouf, Charmaine B. Dean, X. Joan Hu, Robert Delatolla, and Douglas G. Manuel. An exploration of the relationship between wastewater viral signals and COVID-19 hospitalizations in Ottawa, Canada. *Infectious Disease Modelling*, 8(3):617–631, September 2023.

[15] Dustin T. Hill, Mohammed A. Alazawi, E. Joe Moran, Lydia J. Bennett, Ian Bradley, Mary B. Collins, Christopher J. Gobler, Hyatt Green, Tabassum Z. Insaf, Brittany Kmush, Dana Neigel, Shailla Raymond, Mian Wang, Yinyin Ye, and David A. Larsen. Wastewater surveillance provides 10-days forecasting of COVID-19 hospitalizations superior to cases and test positivity: A prediction study. *Infectious Disease Modelling*, 8(4):1138–1150, December 2023.

[16] Wouter A. Hetebrij, Ana Maria De Roda Husman, Erwin Nagelkerke, Rudolf F.H.J. Van Der Beek, Senna C.J.L. Van Iersel, Titus G.V. Breuning, Willemijn J. Lodder, and Michiel Van Boven. Inferring hospital admissions from SARS-CoV-2 virus loads in wastewater in The Netherlands, August 2020 – February 2022. *Science of The Total Environment*, 912:168703, February 2024.

[17] Maria L Daza-Torres, J Cricelio Montesinos-López, Heather N Bischel, Colleen C

Naughton, Angel N Desai, Marlene K Wolfe, Alexandria B Boehm, and Miriam Nuño. A mixed-effects model to predict covid-19 hospitalizations using wastewater surveillance. *Journal of Environmental Chemical Engineering*, 12(2):112485, 2024.

[18] Aikaterini Galani, Reza Aalizadeh, Marios Kostakis, Athina Markou, Nikiforos Alygizakis, Theodore Lytras, Panagiotis G. Adamopoulos, Jordan Peccia, David C. Thompson, Aikaterini Kontou, Apostolos Karagiannidis, Evi S. Lianidou, Margaritis Avgeris, Dimitrios Paraskevis, Sotirios Tsiodras, Andreas Scorilas, Vasilis Vasiliou, Meletios-Athanasios Dimopoulos, and Nikolaos S. Thomaidis. SARS-CoV-2 wastewater surveillance data can predict hospitalizations and ICU admissions. *Science of The Total Environment*, 804:150151, January 2022.

[19] Xuan Li, Huan Liu, Li Gao, Samendra P. Sherchan, Ting Zhou, Stuart J. Khan, Mark C. M. Van Loosdrecht, and Qilin Wang. Wastewater-based epidemiology predicts COVID-19-induced weekly new hospital admissions in over 150 USA counties. *Nature Communications*, 14(1):4548, July 2023.

[20] Swarna Kanchan, Ernie Ogden, Minu Kesheri, Alexis Skinner, Erin Miliken, Devyn Lyman, Jacob Armstrong, Lawrence Sciglitano, and Greg Hampikian. COVID-19 hospitalizations and deaths predicted by SARS-CoV-2 levels in Boise, Idaho wastewater. *Science of The Total Environment*, 907:167742, January 2024.

[21] Ahmed Alaa, Zhaozhi Qian, Jem Rashbass, Jonathan Benger, and Mihaela van der Schaar. Retrospective cohort study of admission timing and mortality following covid-19 infection in england. *BMJ Open*, 10(11), 2020.

[22] Archana Anand, Chenghua Long, and Kartik Chandran. NYC metropolitan wastewater reveals links between SARS-CoV-2 amino acid mutations and disease outcomes. *Science of The Total Environment*, 908:167971, January 2024.

[23] Qingyu Zhan, Helena Maria Solo-Gabriele, Mark E. Sharkey, Ayaaz Amirali, Cynthia C. Beaver, Melinda M. Boone, Samuel Comerford, Daniel Cooper, Elena M. Cortizas, Gabriella A. Cosculluela, Benjamin B. Currall, George S. Grills, Erin Kobetz, Naresh Kumar, Jennifer Laine, Walter E. Lamar, Jiangnan Lyu, Christopher E. Mason, Brian D. Reding, Matthew A. Roca, Stephan C. Schürer, Bhavarth S. Shukla, Natasha Schaefer Solle, Maritza M. Suarez, Mario Stevenson, John J. Tallon, Collette Thomas, Dušica Vidović, Sion L. Williams, Xue Yin, Yalda Zarnegarnia, and Kristina Marie Babler. Correlative Analysis of Wastewater Trends with Clinical Cases and Hospitalizations through Five Dominant Variant Waves of COVID-19. *ACS ES&T Water*, 3(9):2849–2862, September 2023.

[24] Discharge abstract database (dad) metadata. https://www.cihi.ca/en/discharge-abstract-database-dad-metadata.

[25] Áine O'Toole, Oliver G Pybus, Michael E Abram, Elizabeth J Kelly, and Andrew Rambaut. Pango lineage designation and assignment using sars-cov-2 spike gene nucleotide sequences. *BMC genomics*, 23(1):121, 2022.

[26] Cct r package. https://github.com/TheZetner/cct.

[27] Pango designation. https://github.com/cov-lineages/pango-designation.

[28] P. de Valpine, D. Turek, C.J. Paciorek, C. Anderson-Bergman, D. Temple Lang,

and R. Bodik. Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26:403–417, 2017.

[29] P. de Valpine, C. Paciorek, D. Turek, N. Michaud, C. Anderson-Bergman, F. Obermeyer, C. Wehrhahn Cortes, A. Rodríguez, D. Temple Lang, W. Zhang, S. Paganin, J. Hug, and P. van Dam-Bates. Nimble: Mcmc, particle filtering, and programmable hierarchical modeling. https://cran.r-project.org/package=nimble, 2024.

[30] Statistics Canada. Census profile. 2021 census of population. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E, 2023. Statistics Canada Catalogue number 98-316-X2021001. Ottawa. Released November 15, 2023. Accessed April 24, 2025.

[31] World Health Organization. Who announces simple, easy-to-say labels for sars-cov-2 variants of interest and concern. https://www.who.int/news/item/31-05-2021-who-announces-simple-easy-to-say-labels-for-sars-cov-2-variants-of-interest-and-concern, May 2021. Accessed: 2025-04-24.

[32] World Health Organization. Statement on the fifteenth meeting of the ihr (2005) emergency committee on the covid-19 pandemic. https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic.

[33] Paulo R.S. Sanches, Ives Charlie-Silva, Helyson L.B. Braz, Cíntia Bittar, Marilia Freitas Calmon, Paula Rahal, and Eduardo M. Cilli. Recent advances in SARS-CoV-2 Spike protein and RBD mutations comparison between new variants Alpha (B.1.1.7, United Kingdom), Beta (B.1.351, South Africa), Gamma (P.1, Brazil) and Delta (B.1.617.2, India). *Journal of Virus Eradication*, 7(3):100054, September 2021.

[34] Katherine A Twohig, Tommy Nyberg, Asad Zaidi, Simon Thelwall, Mary A Sinnathamby, Shirin Aliabadi, Shaun R Seaman, Ross J Harris, Russell Hope, Jamie Lopez-Bernal, Eileen Gallagher, Andre Charlett, Daniela De Angelis, Anne M Presanis, Gavin Dabrera, Cherian Koshy, Amy Ash, Emma Wise, Nathan Moore, Matilde Mori, Nick Cortes, Jessica Lynch, Stephen Kidd, Derek Fairley, Tanya Curran, James McKenna, Helen Adams, Christophe Fraser, Tanya Golubchik, David Bonsall, Mohammed Hassan-Ibrahim, Cassandra Malone, Benjamin Cogger, Michelle Wantoch, Nicola Reynolds, Ben Warne, Joshua Maksimovic, Karla Spellman, Kathryn McCluggage, Michaela John, Robert Beer, Safiah Afifi, Sian Morgan, Angela Marchbank, Anna Price, Christine Kitchen, Huw Gulliver, Ian Merrick, Joel Southgate, Martyn Guest, Robert Munn, Trudy Workman, Thomas Connor, William Fuller, Catherine Bresner, Luke Snell, Amita Patel, Themoula Charalampous, Gaia Nebbia, Rahul Batra, Jonathan Edgeworth, Samuel Robson, Angela Beckett, David Aanensen, Anthony Underwood, Corin Yeats, Khalil Abudahab, Ben Taylor, Mirko Menegazzo, Gemma Clark, Wendy Smith, Manjinder Khakh, Vicki Fleming, Michelle Lister, Hannah Howson-Wells, Louise Berry, Tim Boswell, Amelia Joseph, Iona Willingham, Carl Jones, Christopher Holmes, Paul Bird, Thomas Helmer, Karlie Fallon, Julian Tang, Veena Raviprakash, Sharon

Campbell, Nicola Sheriff, Victoria Blakey, Lesley-Anne Williams, Matthew Loose, Nadine Holmes, Christopher Moore, Matthew Carlile, Victoria Wright, Fei Sang, Johnny Debebe, Francesc Coll, Adrian Signell, Gilberto Betancor, Harry Wilson, Sahar Eldirdiri, Anita Kenyon, Thomas Davis, Oliver Pybus, Louis Du Plessis, Alex Zarebski, Jayna Raghwani, Moritz Kraemer, Sarah Francois, Stephen Attwood, Tetyana Vasylyeva, Marina Escalera Zamudio, Bernardo Gutierrez, M. Estee Torok, William Hamilton, Ian Goodfellow, Grant Hall, Aminu Jahun, Yasmin Chaudhry, Myra Hosmillo, Malte Pinckert, Iliana Georgana, Samuel Moses, Hannah Lowe, Luke Bedford, Jonathan Moore, Susanne Stonehouse, Chloe Fisher, Ali Awan, John BoYes, Judith Breuer, Kathryn Harris, Julianne Brown, Divya Shah, Laura Atkinson, Jack Lee, Nathaniel Storey, Flavia Flaviani, Adela Alcolea-Medina, Rebecca Williams, Gabrielle Vernet, Michael Chapman, Lisa Levett, Judith Heaney, Wendy Chatterton, Monika Pusok, Li Xu-McCrae, Darren Smith, Matthew Bashton, Gregory Young, Alison Holmes, Paul Randell, Alison Cox, Pinglawathee Madona, Frances Bolt, James Price, Siddharth Mookerjee, Manon Ragonnet-Cronin, Fabricia F. Nascimento, David Jorgensen, Igor Siveroni, Rob Johnson, Olivia Boyd, Lily Geidelberg, Erik Volz, Aileen Rowan, Graham Taylor, Katherine Smollett, Nicholas Loman, Joshua Quick, Claire McMurray, Joanne Stockton, Sam Nicholls, Will Rowe, Radoslaw Poplawski, Alan McNally, Rocio Martinez Nunez, Jenifer Mason, Trevor Robinson, Elaine O'Toole, Joanne Watts, Cassie Breen, Angela Cowell, Graciela Sluga, Nicholas Machin, Shazaad Ahmad, Ryan George, Fenella Halstead, Venkat Sivaprakasam, Wendy Hogsden, Chris Illingworth, Chris Jackson, Emma Thomson, James Shepherd, Patawee Asamaphan, Marc Niebel, Kathy Li, Rajiv Shah, Natasha Jesudason, Lily Tong, Alice Broos, Daniel Mair, Jenna Nichols, Stephen Carmichael, Kyriaki Nomikou, Elihu Aranday-Cortes, Natasha Johnson, Igor Starinskij, Ana Da Silva Filipe, David Robertson, Richard Orton, Joseph Hughes, Sreenu Vattipally, Joshua Singer, Seema Nickbakhsh, Antony Hale, Louissa Macfarlane-Smith, Katherine Harper, Holli Carden, Yusri Taha, Brendan Payne, Shirelle Burton-Fanning, Sheila Waugh, Jennifer Collins, Gary Eltringham, Steven Rushton, Sarah O'Brien, Amanda Bradley, Alasdair Maclean, Guy Mollett, Rachel Blacow, Kate Templeton, Martin McHugh, Rebecca Dewar, Elizabeth Wastenge, Samir Dervisevic, Rachael Stanley, Emma Meader, Lindsay Coupland, Louise Smith, Clive Graham, Edward Barton, Debra Padgett, Garren Scott, Emma Swindells, Jane Greenaway, Andrew Nelson, Clare McCann, Wen Yew, Monique Andersson, Timothy Peto, Anita Justice, David Eyre, Derrick Crook, Tim Sloan, Nichola Duckworth, Sarah Walsh, Anoop Chauhan, Sharon Glaysher, Kelly Bicknell, Sarah Wyllie, Scott Elliott, Allyson Lloyd, Robert Impey, Nick Levene, Lynn Monaghan, Declan Bradley, Tim Wyatt, Elias Allara, Clare Pearson, Husam Osman, Andrew Bosworth, Esther Robinson, Peter Muir, Ian Vipond, Richard Hopes, Hannah Pymont, Stephanie Hutchings, Martin Curran, Surendra Parmar, Angie Lackenby, Tamyo Mbisa, Steven Platt, Shahjahan Miah, David Bibby, Carmen Manso, Jonathan Hubb, Meera Chand, Gavin Dabrera, Mary Ramsay, Daniel Bradshaw, Alicia Thornton, Richard Myers, Ulf Schaefer, Natalie Groves,

Eileen Gallagher, David Lee, David Williams, Nicholas Ellaby, Ian Harrison, Hassan Hartman, Nikos Manesis, Vineet Patel, Chloe Bishop, Vicki Chalker, Juan Ledesma, Katherine Twohig, Matthew Holden, Sharif Shaaban, Alec Birchley, Alexander Adams, Alisha Davies, Amy Gaskin, Amy Plimmer, Bree Gatica-Wilcox, Caoimhe McKerr, Catherine Moore, Chris Williams, David Heyburn, Elen De Lacy, Ember Hilvers, Fatima Downing, Giri Shankar, Hannah Jones, Hibo Asad, Jason Coombes, Joanne Watkins, Johnathan Evans, Laia Fina, Laura Gifford, Lauren Gilbert, Lee Graham, Malorie Perry, Mari Morgan, Matthew Bull, Michelle Cronin, Nicole Pacchiarini, Noel Craine, Rachel Jones, Robin Howe, Sally Corden, Sara Rey, Sara Kumziene-SummerhaYes, Sarah Taylor, Simon Cottrell, Sophie Jones, Sue Edwards, Justin O'Grady, Andrew Page, Alison Mather, David Baker, Steven Rudder, Alp Aydin, Gemma Kay, Alexander Trotter, Nabil-Fareed Alikhan, Leonardo De Oliveira Martins, Thanh Le-Viet, Lizzie Meadows, Anna Casey, Liz Ratcliffe, David Simpson, Zoltan Molnar, Thomas Thompson, Erwan Acheson, Jane Masoli, Bridget Knight, Sian Ellard, Cressida Auckland, Christopher Jones, Tabitha Mahungu, Dianne Irish-Tavares, Tanzina Haque, Jennifer Hart, Eric Witele, Melisa Fenton, Ashok Dadrah, Amanda Symmonds, Tranprit Saluja, Yann Bourgeois, Garry Scarlett, Katie Loveson, Salman Goudarzi, Christopher Fearn, Kate Cook, Hannah Dent, Hannah Paul, David Partridge, Mohammad Raza, Cariad Evans, Kate Johnson, Steven Liggett, Paul Baker, Stephen Bonner, Sarah Essex, Ronan Lyons, Kordo Saeed, Adhyana Mahanama, Buddhini Samaraweera, Siona Silveira, Emanuela Pelosi, Eleri Wilson-Davies, Rachel Williams, Mark Kristiansen, Sunando Roy, Charlotte Williams, Marius Cotic, Nadua Bayzid, Adam Westhorpe, John Hartley, Riaz Jannoo, Helen Lowe, Angeliki Karamani, Leah Ensell, Jacqui Prieto, Sarah Jeremiah, Dimitris Grammatopoulos, Sarojini Pandey, Lisa Berry, Katie Jones, Alex Richter, Andrew Beggs, Angus Best, Benita Percival, Jeremy Mirza, Oliver Megram, Megan Mayhew, Liam Crawford, Fiona Ashcroft, Emma Moles-Garcia, Nicola Cumley, Colin Smith, Giselda Bucca, Andrew Hesketh, Beth Blane, Sophia Girgis, Danielle Leek, Sushmita Sridhar, Sally Forrest, Claire Cormie, Harmeet Gill, Joana Dias, Ellen Higginson, Mailis Maes, Jamie Young, Leanne Kermack, Ravi Gupta, Catherine Ludden, Sharon Peacock, Sophie Palmer, Carol Churcher, Nazreen Hadjirin, Alessandro Carabelli, Ellena Brooks, Kim Smith, Katerina Galai, Georgina McManus, Chris Ruis, Rose Davidson, Andrew Rambaut, Thomas Williams, Carlos Balcazar, Michael Gallagher, Áine O'Toole, Stefan Rooke, Verity Hill, Kathleen Williamson, Thomas Stanton, Stephen Michell, Claire Bewshea, Ben Temperton, Michelle Michelsen, Joanna Warwick-Dugdale, Robin Manley, Audrey Farbos, James Harrison, Christine Sambles, David Studholme, Aaron Jeffries, Leigh Jackson, Alistair Darby, Julian Hiscox, Steve Paterson, Miren Iturriza-Gomara, Kathryn Jackson, Anita Lucaci, Edith Vamos, Margaret Hughes, Lucille Rainbow, Richard Eccles, Charlotte Nelson, Mark Whitehead, Lance Turtle, Sam Haldenby, Richard Gregory, Matthew Gemmell, Claudia Wierzbicki, Hermione Webster, Thushan De Silva, Nikki Smith, Adrienn Angyal, Benjamin Lindsey, Danielle Groves, Luke Green, Dennis Wang, Timothy Freeman, Matthew Parker, Alexander Keeley, Paul Parsons,

Rachel Tucker, Rebecca Brown, Matthew Wyles, Max Whiteley, Peijun Zhang, Marta Gallis, Stavroula Louka, Chrystala Constantinidou, Meera Unnikrishnan, Sascha Ott, Jeffrey Cheng, Hannah Bridgewater, Lucy Frost, Grace Taylor-Joyce, Richard Stark, Laura Baxter, Mohammad Alam, Paul Brown, Dinesh Aggarwal, Alberto Cerda, Tammy Merrill, Rebekah Wilson, Patrick McClure, Joseph Chappell, Theocharis Tsoleridis, Jonathan Ball, David Buck, John Todd, Angie Green, Amy Trebes, George MacIntyre-Cockett, Mariateresa De Cesare, Alex Alderton, Roberto Amato, Cristina Ariani, Mathew Beale, Charlotte Beaver, Katherine Bellis, Emma Betteridge, James Bonfield, John Danesh, Matthew Dorman, Eleanor Drury, Ben Farr, Luke Foulser, Sonia Goncalves, Scott Goodwin, Marina Gourtovaia, Ewan Harrison, David Jackson, Dorota Jamrozy, Ian Johnston, Leanne Kane, Sally Kay, Jon-Paul Keatley, Dominic Kwiatkowski, Cordelia Langford, Mara Lawniczak, Laura Letchford, Rich Livett, Stephanie Lo, Inigo Martincorena, Samantha McGuigan, Rachel Nelson, Steve Palmer, Naomi Park, Minal Patel, Liam Prestwood, Christoph Puethe, Michael Quail, Shavanthi Rajatileka, Carol Scott, Lesley Shirley, John Sillitoe, Michael Spencer Chapman, Scott Thurston, Gerry Tonkin-Hill, Danni Weldon, Diana Rajan, Iraad Bronner, Louise Aigrain, Nicholas Redshaw, Stefanie Lensing, Robert Davies, Andrew Whitwham, Jennifier Liddle, Kevin Lewis, Jaime Tovar-Corona, Steven Leonard, Jillian Durham, Andrew Bassett, Shane McCarthy, Robin Moll, Keith James, Karen Oliver, Alex Makunin, Jeff Barrett, and Rory Gunson. Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (B.1.617.2) compared with alpha (B.1.1.7) variants of concern: A cohort study. *The Lancet Infectious Diseases*, 22(1):35–42, January 2022.

[35] Camino Trobajo-Sanmartín, Ana Miqueleiz, Marcela Guevara, Miguel Fernández-Huerta, Cristina Burgui, Itziar Casado, Fernando Baigorria, Ana Navascués, Carmen Ezpeleta, and Jesús Castilla. Comparison of the Risk of Hospitalization and Severe Disease Among Co-circulating Severe Acute Respiratory Syndrome Coronavirus 2 Variants. *The Journal of Infectious Diseases*, 227(3):332–338, February 2023.

[36] Sharon K. Greene, Alison Levin-Rector, Nang T. T. Kyaw, Elizabeth Luoma, Helly Amin, Emily McGibbon, Robert W. Mathes, and Shama D. Ahuja. Comparative hospitalization risk for SARS-CoV-2 Omicron and Delta variant infections, by variant predominance periods and patient-level sequencing results, New York City, August 2021–January 2022. *Influenza and Other Respiratory Viruses*, 17(1):e13062, January 2023.

[37] Mary-Ann Davies, Reshma Kassanjee, Petro Rousseau, Erna Morden, Leigh Johnson, Wesley Solomon, Nei-Yuan Hsiao, Hannah Hussey, Graeme Meintjes, Masudah Paleker, Theuns Jacobs, Peter Raubenheimer, Alexa Heekes, Pierre Dane, Jamy-Lee Bam, Mariette Smith, Wolfgang Preiser, David Pienaar, Marc Mendelson, Jonathan Naude, Neshaad Schrueder, Ayanda Mnguni, Sue Le Roux, Kathleen Murie, Hans Prozesky, Hassan Mahomed, Liezel Rossouw, Sean Wasserman, Deborah Maughan, Linda Boloko, Barry Smith, Jantjie Taljaard, Greg Symons, Ntobeko A. B. Ntusi, Arifa Parker, Nicole Wolter, Waasila Jassat, Cheryl Cohen, Richard Lessells, Robert J. Wilkinson, Juanita Arendse, Saadiq Kariem,

Melvin Moodley, Milani Wolmarans, Keith Cloete, Andrew Boulle, and Western Cape and South African National Departments of Health in collaboration with the National Institute for Communicable Diseases in South Africa Affiliations. Outcomes of laboratory-confirmed SARS-CoV -2 infection in the Omicron-driven fourth wave compared with previous waves in the Western Cape Province, South Africa. *Tropical Medicine & International Health*, 27(6):564–573, June 2022.

[38] Nicole Wolter, Waasila Jassat, Sibongile Walaza, Richard Welch, Harry Moultrie, Michelle Groome, Daniel Gyamfi Amoako, Josie Everatt, Jinal N Bhiman, Cathrine Scheepers, Naume Tebeila, Nicola Chiwandire, Mignon Du Plessis, Nevashan Govender, Arshad Ismail, Allison Glass, Koleka Mlisana, Wendy Stevens, Florette K Treurnicht, Zinhle Makatini, Nei-yuan Hsiao, Raveen Parboosing, Jeannette Wadula, Hannah Hussey, Mary-Ann Davies, Andrew Boulle, Anne Von Gottberg, and Cheryl Cohen. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: A data linkage study. *The Lancet*, 399(10323):437–446, January 2022.

[39] Shannon L. Russell, Braeden R. A. Klaver, Sean P. Harrigan, Kimia Kamelian, John Tyson, Linda Hoang, Marsha Taylor, Beate Sander, Sharmistha Mishra, Natalie Prystajecky, Naveed Z. Janjua, James E. A. Zlosnik, and Hind Sbihi. Clinical severity of Omicron subvariants BA.1, BA.2, and BA.5 in a population-based cohort study in British Columbia, Canada. *Journal of Medical Virology*, 95(1):e28423, January 2023.

[40] Panke Qu, John P. Evans, Julia N. Faraone, Yi-Min Zheng, Claire Carlin, Mirela Anghelina, Patrick Stevens, Soledad Fernandez, Daniel Jones, Gerard Lozanski, Ashish Panchal, Linda J. Saif, Eugene M. Oltz, Kai Xu, Richard J. Gumina, and Shan-Lu Liu. Enhanced neutralization resistance of SARS-CoV-2 Omicron subvariants BQ.1, BQ.1.1, BA.4.6, BF.7, and BA.2.75.2. *Cell Host & Microbe*, 31(1):9–17.e3, January 2023.

[41] Xiao-Lin Jiang, Ka-Li Zhu, Xue-Jun Wang, Guo-Lin Wang, Yi-Ke Li, Xue-Juan He, Wen-Kui Sun, Peng-Xiang Huang, Jin-Zhong Zhang, Hui-Xia Gao, Er-Hei Dai, and Mai-Juan Ma. Omicron BQ.1 and BQ.1.1 escape neutralisation by omicron subvariant breakthrough infection. *The Lancet Infectious Diseases*, 23(1):28–30, January 2023.

[42] Catherine Quinot, Freja Kirsebom, Nick Andrews, Julia Stowe, Mary Ramsay, Gavin Dabrera, Meaghan Kall, Jamie Lopez Bernal, and Alex Allen. Severity of COVID-19 sub-lineages XBB/XBB 1.5/XBB1.16, EG.5.1. and JN.1. in England. *The Lancet Regional Health - Europe*, 43:100975, August 2024.

[43] Matthew E Levy, Vanessa Chilunda, Richard E Davis, Phillip R Heaton, Pamala A Pawloski, Jason D Goldman, Cynthia A Schandl, Lisa M McEwen, Elizabeth T Cirulli, Dana Wyman, Andrew Dei Rossi, Hang Dai, Magnus Isaksson, Nicole L Washington, Tracy Basler, Kevin Tsan, Jason Nguyen, Jimmy Ramirez, Efren Sandoval, William Lee, James Lu, and Shishi Luo. Reduced Likelihood of Hospitalization With the JN.1 or HV.1 Severe Acute Respiratory Syndrome Coronavirus 2 Variants Compared With the EG.5 Variant. *The Journal of Infectious Diseases*, 230(5):1197–1201, November 2024.

[44] Delphine Planas, Isabelle Staropoli, Vincent Michel, Frederic Lemoine, Flora

Donati, Matthieu Prot, Francoise Porrot, Florence Guivel-Benhassine, Banujaa Jeyarajah, Angela Brisebarre, Océane Dehan, Léa Avon, William Henry Bolland, Mathieu Hubert, Julian Buchrieser, Thibault Vanhoucke, Pierre Rosenbaum, David Veyer, Hélène Péré, Bruno Lina, Sophie Trouillet-Assant, Laurent Hocqueloux, Thierry Prazuck, Etienne Simon-Loriere, and Olivier Schwartz. Distinct evolution of SARS-CoV-2 Omicron XBB and BA.2.86/JN.1 lineages combining increased fitness and antibody evasion. *Nature Communications*, 15(1):2254, March 2024.

[45] Dan-Yu Lin, Yi Du, Yangjianchen Xu, Sai Paritala, Matthew Donahue, and Patrick Maloney. Durability of xbb.1.5 vaccines against omicron subvariants. *New England Journal of Medicine*, 390(22):2124–2127, 2024.

[46] Jennifer L. Nguyen, Marianna Mitratza, Hannah R. Volkman, Leonie De Munter, Thao Mai Phuong Tran, Catia Marques, Mustapha Mustapha, Srinivas Valluri, Jingyan Yang, Andrés Antón, Irma Casas, Eduardo Conde-Sousa, Laura Drikite, Beate Grüner, Giancarlo Icardi, Gerrit Luit Ten Kate, Charlotte Martin, Ainara Mira-Iglesias, Alejandro Orrico-Sánchez, Susana Otero-Romero, Gernot Rohde, Luis Jodar, John M. McLaughlin, and Kaatje Bollaerts. Effectiveness of the BNT162b2 XBB.1.5-adapted vaccine against COVID-19 hospitalization related to the JN.1 variant in Europe: A test-negative case-control study using the id.DRIVE platform. *eClinicalMedicine*, 79:102995, January 2025.

[47] Selene Ortiz-Burgos. *Shannon-Weaver Diversity Index*, pages 572–573. Springer Netherlands, Dordrecht, 2016.

[48] Hannah Chung, Siyi He, Sharifa Nasreen, Maria E Sundaram, Sarah A Buchan, Sarah E Wilson, Branson Chen, Andrew Calzavara, Deshayne B Fell, Peter C Austin, Kumanan Wilson, Kevin L Schwartz, Kevin A Brown, Jonathan B Gubbay, Nicole E Basta, Salaheddin M Mahmud, Christiaan H Righolt, Lawrence W Svenson, Shannon E MacDonald, Naveed Z Janjua, Mina Tadrous, and Jeffrey C Kwong. Effectiveness of BNT162b2 and mRNA-1273 covid-19 vaccines against symptomatic SARS-CoV-2 infection and severe covid-19 outcomes in Ontario, Canada: Test negative design study. *BMJ*, page n1943, August 2021.

[49] Eleftheria Vasileiou, Colin R Simpson, Ting Shi, Steven Kerr, Utkarsh Agrawal, Ashley Akbari, Stuart Bedston, Jillian Beggs, Declan Bradley, Antony Chuter, Simon De Lusignan, Annemarie B Docherty, David Ford, Fd Richard Hobbs, Mark Joy, Srinivasa Vittal Katikireddi, James Marple, Colin McCowan, Dylan McGagh, Jim McMenamin, Emily Moore, Josephine Lk Murray, Jiafeng Pan, Lewis Ritchie, Syed Ahmar Shah, Sarah Stock, Fatemeh Torabi, Ruby Sm Tsang, Rachael Wood, Mark Woolhouse, Chris Robertson, and Aziz Sheikh. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: A national prospective cohort study. *The Lancet*, 397(10285):1646–1657, May 2021.

[50] Lucy A McNamara, Ryan E Wiegand, Rachel M Burke, Andrea J Sharma, Michael Sheppard, Jennifer Adjemian, Farida B Ahmad, Robert N Anderson, Kamil E Barbour, Alison M Binder, Sharoda Dasgupta, Deborah L Dee, Emma S Jones, Jennifer L Kriss, B Casey Lyons, Meredith McMorrow, Daniel C Payne, Hannah E Reses, Loren E Rodgers, David Walker, Jennifer R Verani, and

Stephanie J Schrag. Estimating the early impact of the US COVID-19 vaccination programme on COVID-19 cases, emergency department visits, hospital admissions, and deaths among adults aged 65 years and older: An ecological analysis of national surveillance data. *The Lancet*, 399(10320):152–160, January 2022.

[51] Demy Dam, Sharifa Merali, Michelle Chen, Cameron Coulby, Brigitte Ho Mi Fane, Felix Bang, Jordan Robson, and Samara David. COVID-19 outcome trends by vaccination status in Canada, December 2020–January 2022. *Canada Communicable Disease Report*, 50(1/2):40–48, February 2024.

[52] Public Health Agency of Canada. Covid-19 vaccination: Vaccination coverage. https://health-infobase.canada.ca/covid-19/vaccination-coverage/.

[53] Jessica A. Plante, Rafael R. G. Machado, Brooke M. Mitchell, Divya P. Shinde, Jordyn Walker, Dionna Scharton, Allan McConnell, Nehad Saada, Jianying Liu, Bilal Khan, Rafael K. Campos, Bryan A. Johnson, Vineet D. Menachery, Corri B. Levine, Ping Ren, Susan L. F. McLellan, Kenneth S. Plante, and Scott C. Weaver. Vaccination Decreases the Infectious Viral Load of Delta Variant SARS-CoV-2 in Asymptomatic Patients. *Viruses*, 14(9):2071, September 2022.

[54] Ali A. Rabaan, Shamsah H. Al-Ahmed, Hawra Albayat, Sara Alwarthan, Mashael Alhajri, Mustafa A. Najim, Bashayer M. AlShehail, Wasl Al-Adsani, Ali Alghadeer, Wesam A. Abduljabbar, Nouf Alotaibi, Jameela Alsalman, Ali H. Gorab, Reem S. Almaghrabi, Ali A. Zaidan, Sahar Aldossary, Mohammed Alissa, Lamees M. Alburaiky, Fatimah Mustafa Alsalim, Nanamika Thakur, Geetika Verma, and Manish Dhawan. Variants of SARS-CoV-2: Influences on the Vaccines' Effectiveness and Possible Strategies to Overcome Their Consequences. *Medicina*, 59(3):507, March 2023.

[55] Kristina M. Babler, Mark E. Sharkey, Samantha Abelson, Ayaaz Amirali, Aymara Benitez, Gabriella A. Cosculluela, George S. Grills, Naresh Kumar, Jennifer Laine, Walter Lamar, Erik D. Lamm, Jiangnan Lyu, Christopher E. Mason, Philip M. McCabe, Joshi Raghavender, Brian D. Reding, Matthew A. Roca, Stephan C. Schürer, Mario Stevenson, Angela Szeto, John J. Tallon, Dusica Vidović, Yalda Zarnegarnia, and Helena M. Solo-Gabriele. Degradation rates influence the ability of composite samples to represent 24-hourly means of SARS-CoV-2 and other microbiological target measures in wastewater. *Science of The Total Environment*, 867:161423, April 2023.

[56] Elisabeth Mercier, Patrick M. D'Aoust, Elizabeth Renouf, Emma Tomalty, Felix G. Addo, Tram Bich Nguyen, Chandler H. Wong, Nathan T. Ramsay, Xin Tian, Nada Hegazy, Md Pervez Kabir, Jian-Jun Jia, Shen Wan, Lakshmi Pisharody, Pawel Szulc, Alex E. MacKenzie, and Robert Delatolla. Effective method to mitigate impact of rain or snowmelt sewer flushing events on wastewater-based surveillance measurements. *Science of The Total Environment*, 956:177351, December 2024.

[57] Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada. Tri-council policy statement: Ethical conduct for research involving humans. https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html, December 2022. Accessed April 30, 2025.

Appendix A Preliminary check of hierarchical model using synthetic data

To check our implementation of the hierarchical model in the programming language R, we simulated data with the same hierarchical structure as expected by the model. Briefly, we chose a value for the universal slope and intercept for the log-linear relationship between wastewater concentration and hospital admissions, then draw variant-level intercept and slopes from distributions centred on these (chosen) universal values and finally drew again intercepts and slopes at the location/variant level (see Methods section in the main text). Table A1 shows the values taken to generate the simulated data. We chose the number of hypothetical variants and geographical locations to be the same as the real data set we used, that is seven variants and nine locations. For each location/variant pair, we imposed the total number of data points to be sampled from integers between 6 and 24. The simulated data is shown in Figure A1.

Table A1 Parameter values for simulated data

Parameter	Value	Description
m	1.0	universal slope
b	1.1	universal intercept
σ_M	0.5	standard deviation of mu, when drawin at variant level
σ_B	0.6	standard deviation of beta, when drawin at variant level
σ_m	1.0	standard deviation of M, when drawin at variant location/level
σ_b	1.1	standard deviation of B, when drawin at variant location/level
σ	30	standard deviation for the drawn paired data points

We ran the same inference model as for the real data in the main analysis, that is using a Hamiltonian MCMC algorithm in the R package `Nimble` with 500,000 iterations with 250,000 used as burn-in on four chains. The posterior distributions for the universal and variant-level parameters are shown in Figure A2. For most of the parameters, the 95% credible interval includes the "true" value of the parameter used to simulate the data (red vertical line, Figure A2), suggesting there is no major problem with the model implementation.

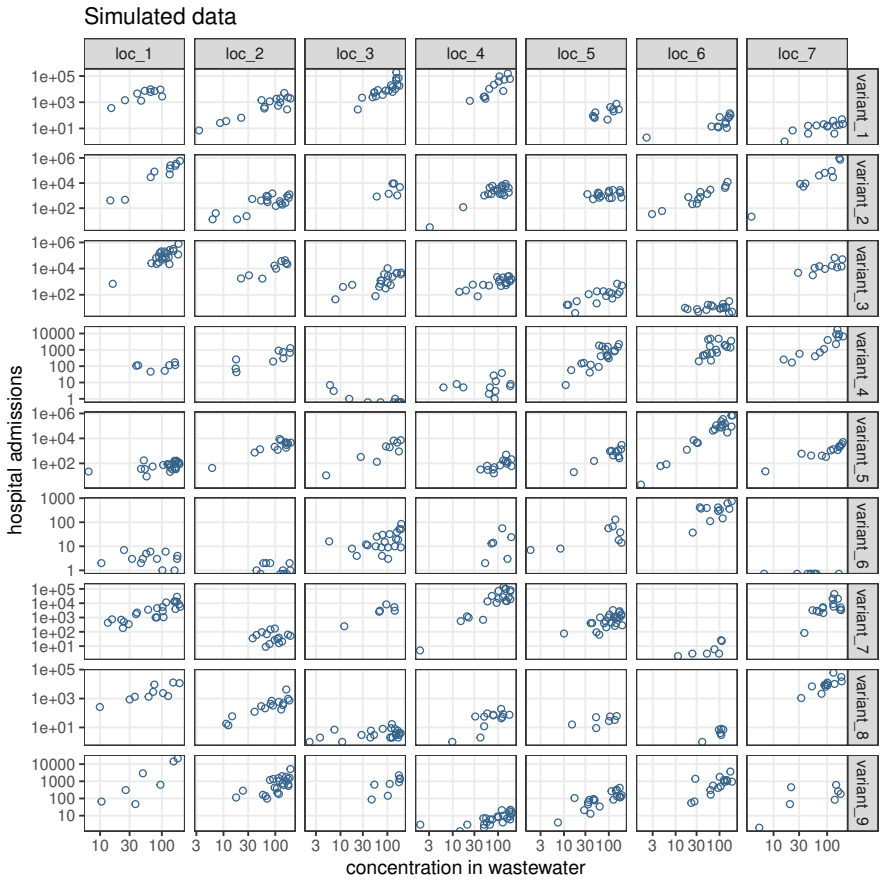

Fig. A1 Simulated data used to check the implementation of the hierarchical model in R.

Appendix B Start and end dates of variant periods in each geographic location

Table B2 shows the start and end dates of variant periods in geographic location used in the our model. Due to the inclusion criteria specified in the methods, the number of locations with a defined start and end date may vary for each variant.

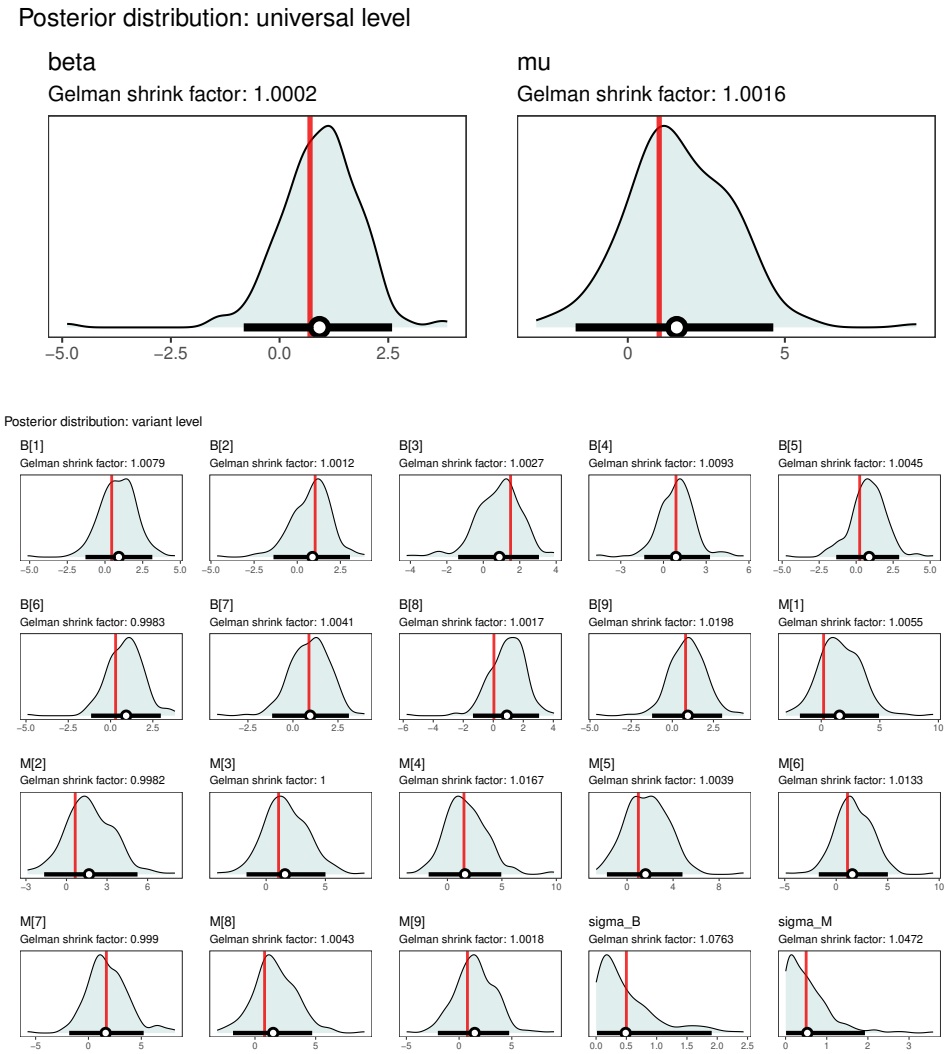

Fig. A2 Posterior distribution for the inference of the universal-level (top panels) and variant-level (lower panels) intercept and slope on simulated data. The vertical red line indicates the value of the parameter used to simulate the data (hence the "true" value that must be inferred). The black open circle shows the mean of the posterior distribution and the black horizontal segment its 95% credible interval.

Appendix C Detailed results of hierarchical model

Here, we present more detailled outputs of the Hamiltonian MCMC algorithm after it ran on the Canadian data sets. As a reminder, for each of the four chains, 500,000 iterations were run of which 250,000 were used as burn-in. The model was implemented in the programming language R and used the packages `nimble` version 1.2.1 and `nimbleHMC` version 0.2.3.

The plots below show the traces of the chains, the Gelman-Rubin statistics and the posterior density distributions for all the parameters at the universal and variant levels. For convenience of the computer implementation, the name of the parameters were slightly changed compared to the notation in the main text. The universal intercept and slope are displayed in the figures of this section as `beta` (b in Equation 1 of the main text) and `mu` (m in Equation 1 of the main text), respectively. The intercepts and slopes at the variant level are shown as `B_XYZ` (b_v as per Equation 2 in the main text for variant XYZ) and `M_XYZ` (m_v as per Equation 2 in the main text for variant XYZ). Moreover, the variance (σ in the main text) is shown as `sigma` in this section.

C.1 Trace plots

A visual inspection of the trace plots in Figure C3 shows a satisfactory mixing of the chains.

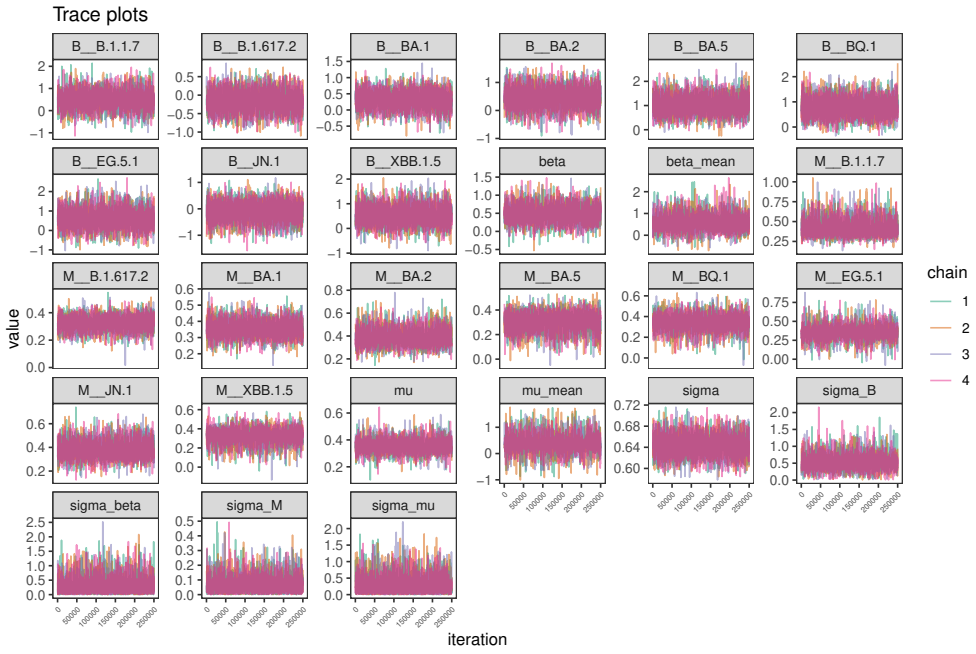

Fig. C3 Posterior trace plots for every parameters. The HMCMC algorithm was run on 4 chains with 500,000 iterations each, including 250,000 used as burn-in.

C.2 Gelman-Rubin convergence statistic

The Gelman-Rubin statistic was below 1.002 for most parameters (Figure C4), confirming a sufficient number of MCMC iterations were performed to reach convergence.

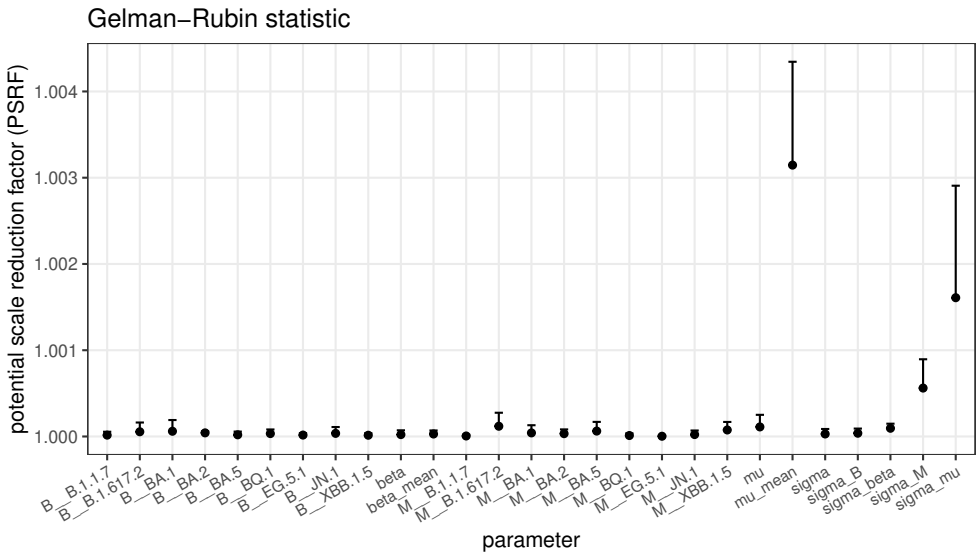

Fig. C4 Gelman-Rubin statistics. The HMCMC algorithm was run on 4 chains with 500,000 iterations each, including 250,000 used as burn-in. Points represent mean estimates and the segment the maximum values.

C.3 Posterior distributions

In Figure C5, we show the posterior distributions for the universal- and variant-level parameters inferred.

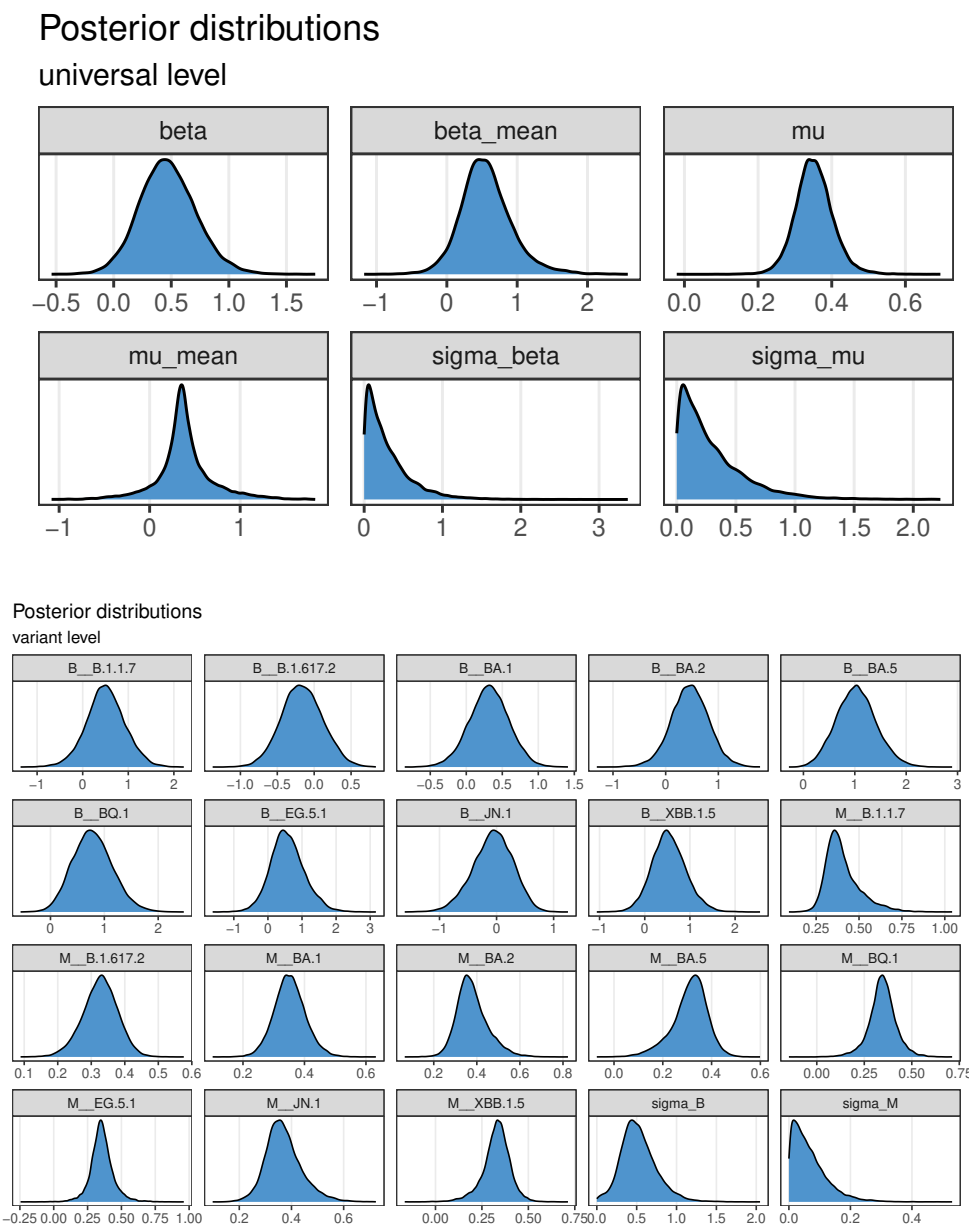

Fig. C5 Posterior distributions.

Appendix D Sensitivity analysis of Hamiltonian MCMC execution

Here, we present the results of our sensitivity analysis to assess the impact of changing the definition of variant dominance on the results of the Hamiltonian MCMC algorithm. To recall, we defined variant dominance by setting the minimum circulation abundance threshold of 60%. This threshold was required to be met for at least 75% within a minimum 8 week data window.

In this sensitivity analysis, the following changes were made to the variant dominance definition:

- Changing the abundance threshold ± 15%, providing alternate thresholds of 45% and 75%.
- Changing the data window ± 2 weeks, providing alternate windows of 6 and 10 weeks.
- Changing the percentage of weeks required to meet the abundance threshold within the window ± 20%, providing alternate percentages of 55% and 95%.

We used the inferred slope (m) and intercept (b) estimates at the universal level to examine the impact of changing the variant dominance definition. In total, 27 distinct Hamiltonian MCMC executions were conducted, including an execution that uses the data that matches the reference dominance definition. Given the number of executions, only 100,000 iterations were run on 4 chains with a burn-in of 50,000. The sensitivity analysis was implemented in the programming language R and used the packages `nimble` version 1.2.1 and `nimbleHMC` version 0.2.3.

Table D3 illustrates that slope and intercept estimates are comparable across different variant dominance definitions.

Appendix E Naive linear regression and its comparison to the Hamiltonian MCMC output

In addition to the hierarchical model, we performed naive linear regression of the log number of hospital admissions and log SARS-CoV-2 wastewater concentrations, independently on each geographic location/variant pair. Despite being two different modelling frameworks, the estimates of the regression intercepts and slopes should not be very different (the hierarchical model is expected to pull variant-level estimates towards the universal-level ones). The comparison of the estimates is shown in E6.

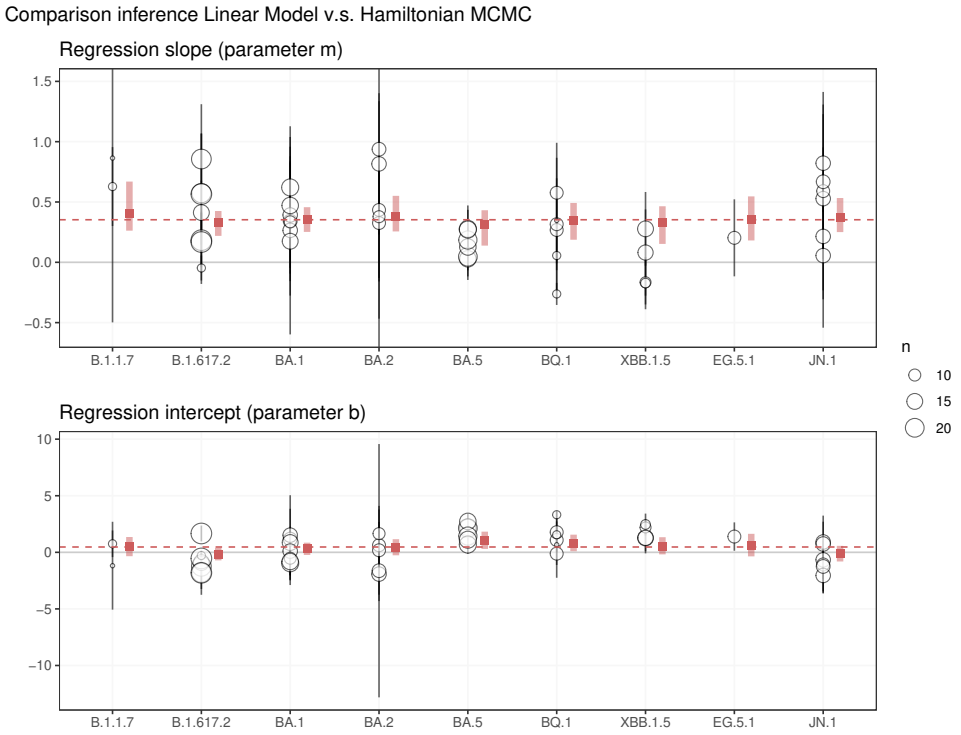

Fig. E6 Comparison of intercepts and slopes obtained from the hierarchical model for each variant with naive linear regression conducted for variants at each geographic location. Red vertical bars represent the HMCMC mean estimates with its respective 95% credible interval. Black circles represent mean estimates from the linear regression for each location, with black vertical lines representing total uncertainty. Size of circles represent the number of data points included.

Table B2 Start and end dates of each variant period used in the model at each location. Start and end dates were derived based on the criteria defined in the methods.

City	Province	Variant	Start	End
Edmonton	AB	B.1.617.2	2021-07-11	2021-12-05
Edmonton	AB	BA.1	2021-12-12	2022-02-27
Edmonton	AB	BA.2	2022-03-20	2022-05-01
Edmonton	AB	BA.5	2022-07-03	2022-10-23
Edmonton	AB	BQ.1	2022-11-27	2023-02-05
Edmonton	AB	JN.1	2024-01-07	2024-03-31
Halifax	NS	B.1.1.7	2021-04-18	2021-05-23
Halifax	NS	B.1.617.2	2021-06-27	2021-11-28
Halifax	NS	BA.1	2021-12-05	2022-03-13
Halifax	NS	BA.2	2022-04-03	2022-05-29
Halifax	NS	BA.5	2022-07-17	2022-11-27
Halifax	NS	BQ.1	2023-01-01	2023-02-12
Halifax	NS	JN.1	2024-01-21	2024-03-31
Regina	SK	B.1.1.7	2021-04-18	2021-05-30
Regina	SK	B.1.617.2	2021-07-04	2021-12-12
Regina	SK	BA.1	2021-12-19	2022-03-06
Regina	SK	BA.2	2022-03-27	2022-05-08
Regina	SK	BA.5	2022-07-03	2022-10-23
Regina	SK	BQ.1	2022-11-20	2023-01-29
Regina	SK	XBB.1.5	2023-02-26	2023-06-04
Regina	SK	JN.1	2024-01-07	2024-03-31
St. John's	NL	B.1.617.2	2021-05-16	2021-11-21
St. John's	NL	BA.1	2021-12-12	2022-03-27
Toronto	ON	B.1.617.2	2021-06-20	2021-12-05
Toronto	ON	BA.1	2021-12-19	2022-03-06
Toronto	ON	BA.2	2022-03-20	2022-05-15
Toronto	ON	BA.5	2022-07-03	2022-11-06
Toronto	ON	BQ.1	2022-12-11	2023-01-15
Toronto	ON	XBB.1.5	2023-03-05	2023-04-30
Toronto	ON	JN.1	2024-01-07	2024-03-31
Vancouver	BC	B.1.617.2	2021-07-18	2021-12-05
Vancouver	BC	BA.1	2021-12-19	2022-02-20
Vancouver	BC	BA.5	2022-07-10	2022-10-30
Vancouver	BC	BQ.1	2022-12-11	2023-01-22
Vancouver	BC	XBB.1.5	2023-03-12	2023-04-30
Vancouver	BC	JN.1	2024-01-07	2024-03-31
Winnipeg	MB	B.1.617.2	2021-06-20	2021-12-05
Winnipeg	MB	BA.1	2021-12-12	2022-03-13
Winnipeg	MB	BA.2	2022-03-27	2022-05-22
Winnipeg	MB	BA.5	2022-07-03	2022-10-23
Winnipeg	MB	BQ.1	2022-11-20	2023-01-29
Winnipeg	MB	XBB.1.5	2023-02-26	2023-05-28
Winnipeg	MB	EG.5.1	2023-10-01	2023-12-10
Winnipeg	MB	JN.1	2024-01-14	2024-03-31

Table D3 Results of the sensitivity analysis of Hamiltonian MCMC execution. Inferred slope (m) and intercept (b) estimates from the Hamiltionian MCMC algorithm at the universal and variant levels were estimated using different variant dominance definitions. Reference refers to the primary dominance definition used in the main model with threshold of 60%, window of 8 weeks, and 75% of window weeks above threshold. Algorithm was executed across 4 chains using 100,000 iterations and 50,000 iterations used as "burn-in". Values in brackets represent the 95% credible interval.

Threshold (%)	**Window (weeks)**	**% above**	**slope** m	**intercept** b
Reference	Reference	Reference	0.354 [0.257; 0.459]	0.466 [-0.004; 0.988]
Reference	Reference	+20	0.343 [0.232; 0.461]	0.53 [0.002; 1.154]
Reference	Reference	-20	0.339 [0.249; 0.434]	0.515 [0.063; 1.017]
Reference	+2	Reference	0.351 [0.251; 0.464]	0.483 [-0.035; 1.062]
Reference	+2	+20	0.333 [0.212; 0.464]	0.606 [-0.002; 1.302]
Reference	+2	-20	0.353 [0.264; 0.451]	0.461 [0.001; 0.955]
Reference	-2	Reference	0.353 [0.257; 0.458]	0.469 [0.002; 0.988]
Reference	-2	+20	0.356 [0.258; 0.466]	0.464 [-0.002; 0.981]
Reference	-2	-20	0.341 [0.249; 0.441]	0.488 [0.026; 0.995]
+15	Reference	Reference	0.313 [0.202; 0.443]	0.636 [0.081; 1.253]
+15	Reference	+20	0.297 [0.165; 0.447]	0.599 [-0.019; 1.285]
+15	Reference	-20	0.307 [0.196; 0.435]	0.688 [0.135; 1.306]
+15	+2	Reference	0.312 [0.19; 0.454]	0.533 [-0.056; 1.184]
+15	+2	+20	0.28 [0.115; 0.472]	0.656 [-0.061; 1.472]
+15	+2	-20	0.306 [0.195; 0.432]	0.655 [0.104; 1.274]
+15	-2	Reference	0.315 [0.206; 0.443]	0.672 [0.121; 1.291]
+15	-2	+20	0.291 [0.171; 0.426]	0.604 [0.042; 1.237]
+15	-2	-20	0.309 [0.199; 0.434]	0.689 [0.14; 1.301]
-15	Reference	Reference	0.367 [0.254; 0.477]	0.448 [-0.044; 0.998]
-15	Reference	+20	0.374 [0.237; 0.504]	0.446 [-0.077; 1.039]
-15	Reference	-20	0.389 [0.295; 0.487]	0.361 [-0.088; 0.851]
-15	+2	Reference	0.375 [0.27; 0.479]	0.398 [-0.068; 0.915]
-15	+2	+20	0.375 [0.257; 0.491]	0.43 [-0.058; 0.985]
-15	+2	-20	0.387 [0.247; 0.484]	0.322 [-0.104; 0.8]
-15	-2	Reference	0.37 [0.253; 0.482]	0.452 [-0.032; 0.999]
-15	-2	+20	0.374 [0.235; 0.505]	0.451 [-0.07; 1.044]
-15	-2	-20	0.39 [0.295; 0.488]	0.361 [-0.091; 0.858]