

1 Improving estimates of epidemiological quantities by combining  
2 reported cases with wastewater data: a statistical framework  
3 with applications to COVID-19 in Aotearoa New Zealand

4 Leighton M. Watson<sup>1</sup>, Michael J. Plank<sup>1</sup>, Bridget A. Armstrong<sup>2</sup>, Joanne R. Chapman<sup>2</sup>,  
5 Joanne Hewitt<sup>2</sup>, Helen Morris<sup>2</sup>, Alvaro Orsi<sup>2</sup>, Michael Bunce<sup>2,3</sup>, Christl A. Donnelly<sup>4,5</sup>,  
6 and Nicholas Steyn<sup>4</sup>

7 <sup>1</sup>School of Mathematics and Statistics, University of Canterbury, New Zealand

8 <sup>2</sup>Institute of Environmental Science and Research Ltd, New Zealand

9 <sup>3</sup>Department of Conservation, New Zealand

10 <sup>4</sup>Department of Statistics, University of Oxford, United Kingdom

11 <sup>5</sup>Pandemic Sciences Institute, University of Oxford, United Kingdom

12

## Abstract

13

**Background:** Timely and informed public health responses to infectious diseases such as COVID-19 necessitate reliable information about infection dynamics. The case ascertainment rate (CAR), the proportion of infections that are reported as cases, is typically much less than one and varies with testing practices and behaviours, making reported cases unreliable as the sole source of data. The concentration of viral RNA in wastewater samples provides an alternate measure of infection prevalence that is not affected by clinical testing, healthcare-seeking behaviour or access to care.

14

15

16

17

18

19

20

**Methods:** We constructed a state-space model with observed data of levels of SARS-CoV-2 in wastewater and reported case incidence and estimated the hidden states of  $R$  and CAR using sequential Monte Carlo methods.

21

22

23

**Results:** Here, we analysed data from 1 January 2022 to 31 March 2023 from Aotearoa New Zealand. Our model estimates that  $R$  peaked at 2.76 (95% CrI 2.20, 3.83) around 18 February 2022 and the CAR peaked around 12 March 2022. We calculate that New Zealand's second Omicron wave in July 2022 was similar in size to the first, despite fewer reported cases. We estimate that the CAR in the BA.5 Omicron wave in July 2022 was approximately 50% lower than in the BA.1/BA.2 Omicron wave in March 2022.

24

25

26

27

28

29

**Conclusions:** Estimating  $R$ , CAR, and cumulative number of infections provides useful information for planning public health responses and understanding the state of immunity in the population. This model is a useful disease surveillance tool, improving situational awareness of infectious disease dynamics in real-time.

30

31

32

### 33 Plain Language Summary

34 To make informed public health decisions about infectious diseases, it is important to under-  
35 stand the number of infections in the community. Reported cases, however, underestimate the  
36 number of infections and the degree of underestimation likely changes with time. Wastewater  
37 data provides an alternative data source that does not depend on testing practices. Here, we  
38 combined wastewater observations of SARS-CoV-2 with reported cases to estimate the repro-  
39 duction number (how quickly infections are increasing or decreasing) and the case ascertainment  
40 rate (the fraction of infections reported as cases). We apply the model to Aotearoa New Zealand  
41 and demonstrate that the second wave of infections in July 2022 had approximately the same  
42 number of infections as the first wave in March 2022 despite reported cases being 50% lower.

## 43 1 Introduction

44 Understanding and predicting the trajectory of infectious diseases is important in planning an  
45 effective public health response. Reported case data depend heavily on testing modalities and  
46 practices which typically change over time, resulting in considerable uncertainty in the case  
47 ascertainment rate (CAR; the fraction of infections that are officially reported). During the  
48 COVID-19 pandemic, many countries relied primarily on symptom-based testing programmes  
49 to inform situational awareness and public health responses. In Aotearoa New Zealand, the CAR  
50 for COVID-19 has been influenced by factors such as access to testing, a shift from healthcare  
51 worker-administered polymerase chain reaction (PCR) tests to self-administered rapid antigen  
52 tests (RATs), reduction in rates of symptomatic and severe disease due to rising population  
53 immunity, relaxation of testing requirements and recommendations, and/or lack of perceived  
54 need to test or ‘pandemic fatigue’ [1–3]. As a result, over time, officially reported cases of  
55 COVID-19 have become a less reliable measure of levels of SARS-CoV-2 infection.

56 Data on hospital admissions and deaths are more consistent and are less affected by testing  
57 practices and behavioural change than reported cases but are subject to additional delays [4]  
58 that limit their usefulness for understanding disease dynamics. Infection prevalence surveys [5]  
59 that aim to regularly test a representative sample of the population are the gold-standard for  
60 tracking the spread of an infectious disease, but these surveys are resource intensive, making  
61 them harder to justify as countries move out of the acute phase of the pandemic. The UK was  
62 the only country to implement regular representative national SARS-CoV-2 prevalence surveys  
63 [6, 7] and there are no current plans for similar surveys in New Zealand.

64 Wastewater surveillance, where levels of SARS-CoV-2 RNA in wastewater samples are mea-  
65 sured, can provide additional data on the prevalence of the virus that are unaffected by individ-  
66 ual testing and self-reporting behaviours. Wastewater surveillance (also known as wastewater-  
67 based epidemiology or WBE) also has the potential to contribute to an integrated global network  
68 for disease surveillance [8–10]. These data, however, can be highly variable and subject to other  
69 biases, such as rainwater dilution, sampling methodologies, and changing locations of selected  
70 sampling sites. To realise this potential, appropriate models and analytical tools are needed to  
71 deliver epidemiological insights from raw data.

72 Two previous studies have presented novel methodology for the real-time estimation of the  
73 effective reproduction number using wastewater data [11, 12], while others have leveraged or  
74 extended these methods [13–16]. One study used reported cases to estimate the reproduction  
75 number and then fitted a model to estimate this quantity from wastewater data [17]. Another  
76 study used wastewater data to fit a mathematical model of multiple viral strains [18] from which

77 estimates of the reproduction number can be derived. Other studies have analysed wastewater  
78 data but did not use it to estimate the reproduction number [19, 20]. Only [11] presented  
79 a model for simultaneously considering clinical and wastewater data, however they assume a  
80 fixed ascertainment rate. No previous work has combined wastewater-based epidemiology with  
81 reported cases to infer changes in case ascertainment over time.

82 Semi-mechanistic models based on the renewal equation are a popular method for epidemic  
83 forecasting and estimation of the instantaneous reproduction number [21–23]. Such methods are  
84 robust to constant under-ascertainment of cases, but may be biased by rapid changes in CAR and  
85 cannot provide any information about the total number of infections. In this paper, we extend  
86 the renewal equation framework [21–23] for reproduction number estimation to incorporate  
87 wastewater time-series data. The model treats the instantaneous reproduction number and  
88 CAR as hidden states and reported cases and quantity of viral RNA in wastewater as observed  
89 states. We use a sequential Monte Carlo approach to infer the hidden states. We apply the model  
90 to national data from Aotearoa New Zealand on reported COVID-19 cases and the average  
91 number of SARS-CoV-2 genome copies per person per day measured in municipal wastewater  
92 samples between January 2022 and March 2023. Because the relationship between infections  
93 and wastewater concentration is only determined in the model up to an overall scaling constant,  
94 it cannot be used to infer the absolute CAR but can be used to estimate relative changes in case  
95 ascertainment over time. The model is designed to be regularly updated as new data become  
96 available, producing real-time estimates of the effective reproduction number and relative change  
97 in CAR. The model has been used to support situational awareness via regular reports to the  
98 New Zealand Ministry of Health from November 2022 to date.

99 From March 2020 until December 2021 New Zealand used strict border controls and intermittent  
100 non-pharmaceutical interventions to suppress and eliminate transmission of SARS-CoV-2. By  
101 the beginning of 2022, there had been a cumulative total of around 3 confirmed cases of COVID-  
102 19 per 1,000 people and around 90% of the population over 12 years old had received at least  
103 two doses of the Pfizer-BioNTech vaccine. From October 2021, interventions were progressively  
104 eased and in January 2022 the B.1.1.529 (Omicron) variant began to spread in the community,  
105 causing the first large wave of infection. Since then community transmission has been sustained,  
106 with multiple further waves of infection being driven by various Omicron subvariants. Between  
107 1 January 2022 and 31 March 2023, there was a cumulative total of around 440 confirmed cases  
108 per 1,000 people, most of which were from self-administered RATs. During this period, SARS-  
109 CoV-2 concentration was regularly measured at various wastewater treatment plants, providing  
110 an additional data source on changes in community prevalence over time.

111 We model the epidemic dynamics and the observed case and wastewater data at the national  
112 level, aggregating over New Zealand’s population of 5.1 million and ignoring regional variations.

113 This is similar to other studies that have aggregated regional case and/or wastewater data to  
114 produce national-level estimates in countries with a comparable population size [24–26]. Our  
115 methodology could, in principle, be applied at a finer geographical scale, although this would  
116 come at the cost of higher levels of noise.

## 117 **2 Materials and Methods**

### 118 **2.1 Data**

119 National daily reported cases of COVID-19 were obtained from the New Zealand Ministry of  
120 Health [27]. Until February 2022, these cases were diagnosed solely by healthcare-administered  
121 PCR testing. From February 2022, in response to the rapid increase in reported cases, RATs  
122 were widely distributed. Since then, the vast majority of reported cases have been from self-  
123 administered RATs, with results reported via an online portal. Hence, data on the number  
124 of tests conducted are not available. Reported cases are shown in Figure 1. As these data  
125 exhibit a clear day-of-the-week effect we remove the weekly trend before fitting the model (see  
126 Supplementary Material section 1.2 for details).

127 SARS-CoV-2 concentration data from wastewater samples tested by the Institute for Environ-  
128 mental Science and Research (ESR) were used for this study [28]. Wastewater samples were  
129 collected every week at municipal wastewater treatment plants located throughout the country,  
130 serving communities with populations ranging from 400 to over 500,000 people. Typically 70-  
131 90% of the national population connected to reticulated wastewater was covered by wastewater  
132 sampling in any given week (60-124 sites, usually sampled twice per week). Each site-level  
133 measurement was normalised to provide an estimate of the number of genome copies per person  
134 per day for that site (see Supplementary Material section 1.1). Typically multiple sites were  
135 sampled per day and, for each day that had at least one sample, we calculated the catchment-  
136 population-weighted average of the genome copies per person (see Figure 1). Because we do  
137 not attempt to model regional variations, we assumed this provided a series of representative  
138 observations of the average concentration of genomic material in the national wastewater (see  
139 also Section 2.2).

### 140 **2.2 Hidden state model**

141 We construct a state-space model (Figure 2) consisting of time-varying hidden states (the in-  
142 stantaneous reproduction number  $R_t$ , daily case ascertainment rate  $CAR_t$ , and daily infection

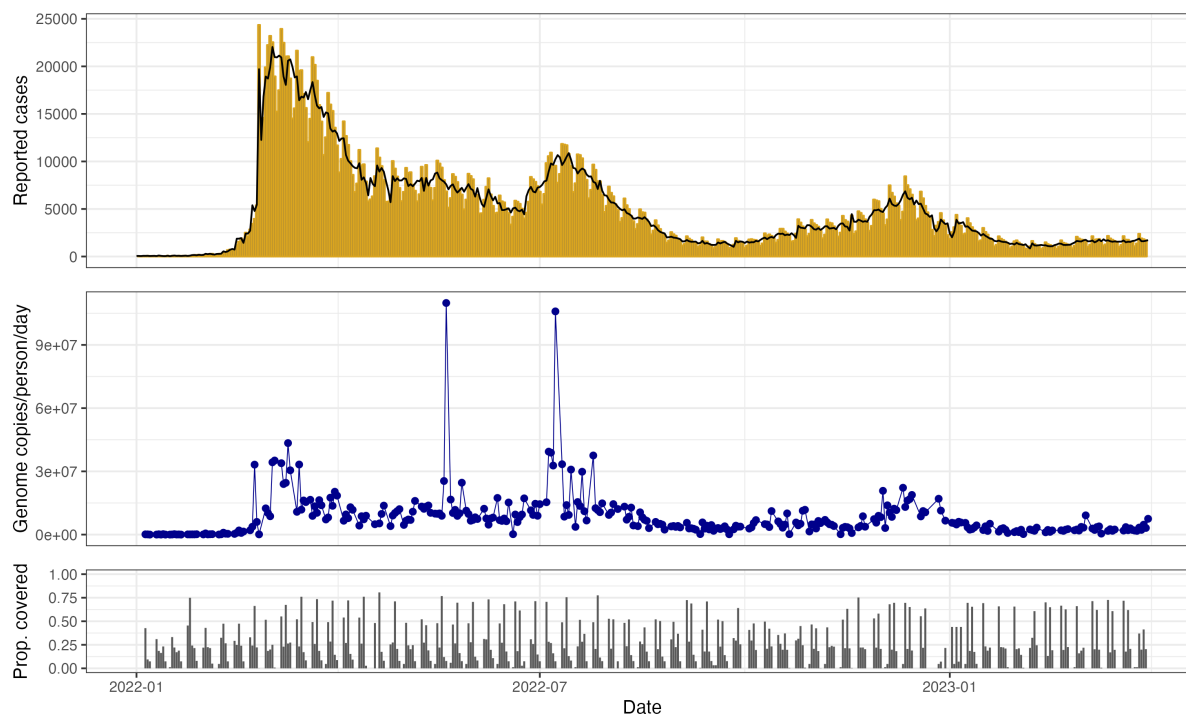


Figure 1: Reported daily cases of COVID-19 (upper), SARS-CoV-2 genome copies per person per day in sampled wastewater (middle), and proportion of the total population covered by sampled wastewater catchments (lower), between 1 January 2022 and 31 March 2023 in Aotearoa New Zealand. The black line in the upper plot shows the adjusted case series with the multiplicative day-of-the-week effect removed (see Supplementary Material section 1.2). The two outliers in wastewater data arise from estimates of a high wastewater flow-rate in Wellington following high rainfall. Since rainfall is a source of noise in wastewater sampling we retain these samples in our analysis. Reported case data were obtained from the New Zealand Ministry of Health [27] and wastewater data were obtained from ESR [28].

143 incidence  $I_t$ ) and time-varying observed states (daily reported cases of COVID-19  $C_t$  and daily  
144 wastewater observations  $W_t$ ). We use subscript  $s : t$  to refer to all values between day  $s$  and  $t$   
145 inclusive.

146 We assume the hidden states  $R_t$  and  $CAR_t$  follow independent Gaussian random walks, encoding  
147 the fact we expect them to vary continuously over time. We also assume that the hidden state  
148  $I_t$  follows a Poisson renewal process, a simple epidemic model commonly used when estimating  
149  $R_t$  [21]. Thus our state-space transitions are governed by:

$$\begin{aligned}(R_t|R_{t-1}) &\sim N_{(0,\infty)}(R_{t-1}, \sigma_R R_{t-1}) \\(CAR_t|CAR_{t-1}) &\sim N_{(0,1)}(CAR_{t-1}, \sigma_{CAR}) \\(I_t|R_t, I_{1:t-1}) &\sim Poisson\left(R_t \sum_{u=1}^{t-1} g_u I_{t-u}\right)\end{aligned}$$

150 Parameters  $\sigma_R$  and  $\sigma_{CAR}$  determine how quickly  $R_t$  and  $CAR_t$  vary. The standard deviation  
151 of the transition distribution for  $R_t \rightarrow R_{t+1}$  is given by  $\sigma_R R_t$ , which means that  $R_t$  varies more  
152 rapidly at larger values. The distribution for  $R_t$  was truncated on  $(0, \infty)$  and for  $CAR_t$  on  
153  $(0, 1)$ . Finally,  $g_u$  is the pre-determined generation time distribution, describing the proportion  
154 of transmission events that occur  $u$  days after infection (see Supplementary Material section  
155 2.7).

156 We assume that the expected number of reported cases  $\mu_t^c$  at time  $t$  is equal to  $CAR_t$  multiplied  
157 by the convolution of past infections with the infection-to-reporting distribution  $L_u$ :

$$\mu_t^c = CAR_t \sum_{u=1}^t I_{t-u} L_u$$

158 Similarly, we assume that the expected number of genome copies  $\mu_t^w$  detected per person at  
159 time  $t$  is equal to the convolution of past infections with the infection-to-shedding distribution  
160  $\omega_u$ , multiplied by a fixed parameter  $a$  representing the average total detectable genome copies  
161 shed into the wastewater by an infectious individual, divided by total national population size  
162  $N$ :

$$\mu_t^w = \frac{a}{N} \sum_{u=1}^t I_{t-u} \omega_u$$

163 We model reported cases using a negative binomial distribution:

$$(C_t|CAR_t, I_{1:t}) \sim NegBin\left(r = k_c, p = \frac{k_c}{k_c + \mu_t^c}\right)$$

164 which has mean  $\mu_t^c$  and variance  $\mu_t^c \left(1 + \frac{\mu_t^c}{k_c}\right)$ . A negative binomial distribution is used to  
165 account for noise in the observations beyond that predicted by a binomial distribution. This is  
166 a common choice in other methods of reproduction number estimation [23, 29].



167 The observed wastewater data  $W_t$  is the total genome copies per person from the wastewater  
168 sites sampled on day  $t$ . We model this using a shape-scale gamma distribution:

$$(W_t|I_{1:t}) \sim \begin{cases} \Gamma\left(k_w \text{pop}_t, \frac{\mu_t^w}{k_w \text{pop}_t}\right) & \text{if } \text{pop}_t > 0 \\ \mathcal{I}(W_t = 0) & \text{if } \text{pop}_t = 0 \end{cases}$$

169 which has mean  $\mu_t^w$  and variance  $\frac{(\mu_t^w)^2}{k_w \text{pop}_t}$ . This assumes that the observed daily data are indepen-  
170 dent draws from the national distribution, which may not hold if there are regional differences  
171 between the subsets of sites that are sampled on different days. In practice, any such differ-  
172 ences will be absorbed into the variance of the daily observation distribution via fitting of the  
173 dispersion parameter  $k_w$ . Since we marginalise out the effect of this parameter when presenting  
174 results, the increased uncertainty associated with regional variability is propagated through to  
175 the credible intervals. The variable  $\text{pop}_t$  refers to the total population in the catchment areas of  
176 the sampled wastewater sites on day  $t$ . Setting the variance of the observation distribution to be  
177 inversely proportional to  $\text{pop}_t$  allows the model to account for increased variability around the  
178 national mean on days when fewer or smaller sites were sampled.  $\mathcal{I}$  is the indicator function,  
179 so on days when no sites were sampled, the probability of observing no wastewater samples is  
180 set to 1, and the model fits to case data alone.

181 Consistent with previous models [30, 31], this formulation assumes that the expected population  
182 shedding rate is proportional to the number of infected individuals, with observations drawn  
183 from a distribution around this mean. We used a gamma distribution, which is a reasonably  
184 flexible choice for a non-negative continuous random variable. However other distributions could  
185 be considered, such as a Weibull or log-normal.

186 In the absence of additional information we are unable to estimate  $\alpha$ , which is proportional to  
187 the average total genome copies shed by an infected individual over the course of their infection.  
188 This means we are unable to estimate the absolute value of  $CAR_t$ . Instead, we run the model  
189 with a range of different values for  $\alpha$ , and estimate the change in  $CAR_t$  relative to its initial  
190 value. This additionally requires the assumption that  $\alpha$  is constant over time, which is unlikely  
191 to be true in general and is a key limitation of our model (see Discussion).

192 In practice, the range of values of  $\alpha$  that we used (see Table 1) was chosen by calibrating model  
193 output for the number of infections with external sources of information. Firstly, we compared  
194 model output to the number of cases in a cohort of around 20,000 border workers who were  
195 tested weekly between January and July 2022 [36]. Secondly, around 40% of all 20-25-year-olds  
196 (an age group unlikely to have a higher CAR than older adults) reported a case of COVID-19 in  
197 the 6 months from 1 February to 31 July 2022 [27]. This suggests that the overall CAR for this  
198 period was likely to be at least 0.4, which translates to an approximate upper bound of 4 million  
199 for the total cumulative number of infections up to 31 July 2022. Neither of these observations

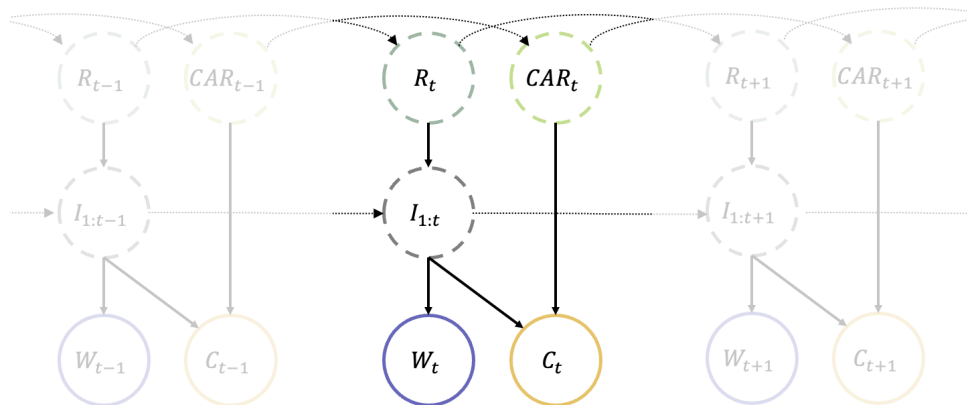


Figure 2: Diagram of the state-space model showing the dependency between hidden-states (dashed circles) and the observed data (solid circles).  $R_t$  is the instantaneous reproduction number on day  $t$ ,  $CAR_t$  is the case ascertainment rate on day  $t$ ,  $I_t$  is the number of new infections on day  $t$ ,  $C_t$  is the number of reported cases on day  $t$ , and  $W_t$  is the observed wastewater, measured as the total genome copies per person per day for the sites that were sampled on day  $t$ .  $I_{1:t}$  denotes the set of states  $\{I_1, I_2, \dots, I_t\}$ . In practice the current infections  $I_t$ , reported cases  $C_t$  and wastewater  $W_t$  depend only on recent values of  $I_t$  as specified by the generation interval distribution, the infection-to-reporting distribution, and infection-to-shedding distribution respectively (see Methods).

Table 1: Parameter values used in the model. The infection-to-reporting and infection-to-shedding distributions are calculated as convolutions of the incubation period distribution [32] and the onset-to-reporting and onset-to-shedding distribution [33] respectively (see Supplementary Material section 2.7).

Parameter	Symbol	Value
Coefficient of variation of $R_t$ transitions	$\sigma_R$	Fitted
Std dev. of $CAR_t$ transitions	$\sigma_{CAR}$	Fitted
Reported cases tuning parameter	$k_c$	Fitted
Wastewater tuning parameter	$k_w$	Fitted
Generation time distribution [34, 35]	$g_u$	Mean = 3.3 days, s.d. = 1.3 days
Infection-to-reporting distribution	$L_u$	Mean = 5.8 days, s.d. = 2.6 days
Infection-to-shedding distribution	$\omega_u$	Mean = 5.2 days, s.d. = 2.9 days
Average total genome copies per infection	$\alpha$	$3 \times 10^9$ [ $2 \times 10^9$ , $4 \times 10^9$ ]
Fixed-lag resampling window	$h$	30 days

200 definitively determines the number of infections as they are subject to approximation, bias and  
201 uncertainty, but they nevertheless serve to bracket the likely range of values for the parameter  
202  $\alpha$ .

203 The infection-to-reporting and infection-to-shedding distributions are calculated as the convo-  
204 lution of the incubation period distribution with the onset-to-reporting and onset-to-shedding  
205 distribution respectively. The incubation period is modelled as a Weibull distribution with mean  
206 2.9 days and standard deviation 2.0 days [32]. The onset-to-reporting distribution is estimated  
207 empirically from New Zealand case data extracted on 16 September 2022, representing over 1.2  
208 million cases, and has mean 1.8 days and standard deviation 1.8 days. The onset-to-shedding  
209 distribution comes from [33] and has mean 0.7 days and standard deviation 2.6 days. The  
210 resulting infection-to-reporting distribution has mean 5.8 days and standard deviation 2.6, and  
211 the resulting infection-to-shedding distribution has mean 5.2 days and standard deviation 2.9  
212 days (see Supplementary Figure S1).

213 The model is solved using a bootstrap filter [37] with fixed-lag resampling. This produces  
214 estimates for the marginal posterior distribution of the hidden states at each time step. The  
215 random walk step variance parameters ( $\sigma_R$  and  $\sigma_{CAR}$ ) and observation variance parameters  
216 ( $k_c$  and  $k_w$ ) are estimated using a particle marginal Metropolis Hastings Markov chain Monte  
217 Carlo method. We use uninformative uniform prior distributions for these parameters, with  
218 the exception of  $\sigma_{CAR}$ , where we use an informative prior distribution to ensure an appropriate  
219 level of smoothness in our estimates of  $CAR_t$ . Different parameter values are fitted in three-  
220 month blocks to allow for some variation over time. See Supplementary Material section 2 for  
221 further details of the numerical method. Code and data to reproduce the results are provided  
222 at <https://github.com/nicsteyn2/NZWastewaterModelling>.

## 223 3 Results

### 224 **Reproduction number, relative case ascertainment, and infection incidence**

225 The estimated value of the reproduction number  $R_t$  (Figure 3a) increased from around 1 at the  
226 beginning of 2022 to a peak of 2.46 (95% CrI 2.04, 3.20) on 18 February 2022 (95% CrI 10 Feb,  
227 23 Feb), corresponding to the sharp increase in cases seen during the first Omicron wave, which  
228 was a mixture of the BA.1 and BA.2 variants [38]. The estimated value of  $R_t$  dropped below 1  
229 on 1 March 2022 (95% CrI 25 Feb, 5 Mar) and infection incidence peaked on 28 February 2022  
230 (95% CrI 23 Feb, 7 Mar), suggesting this is when the wave peaked.

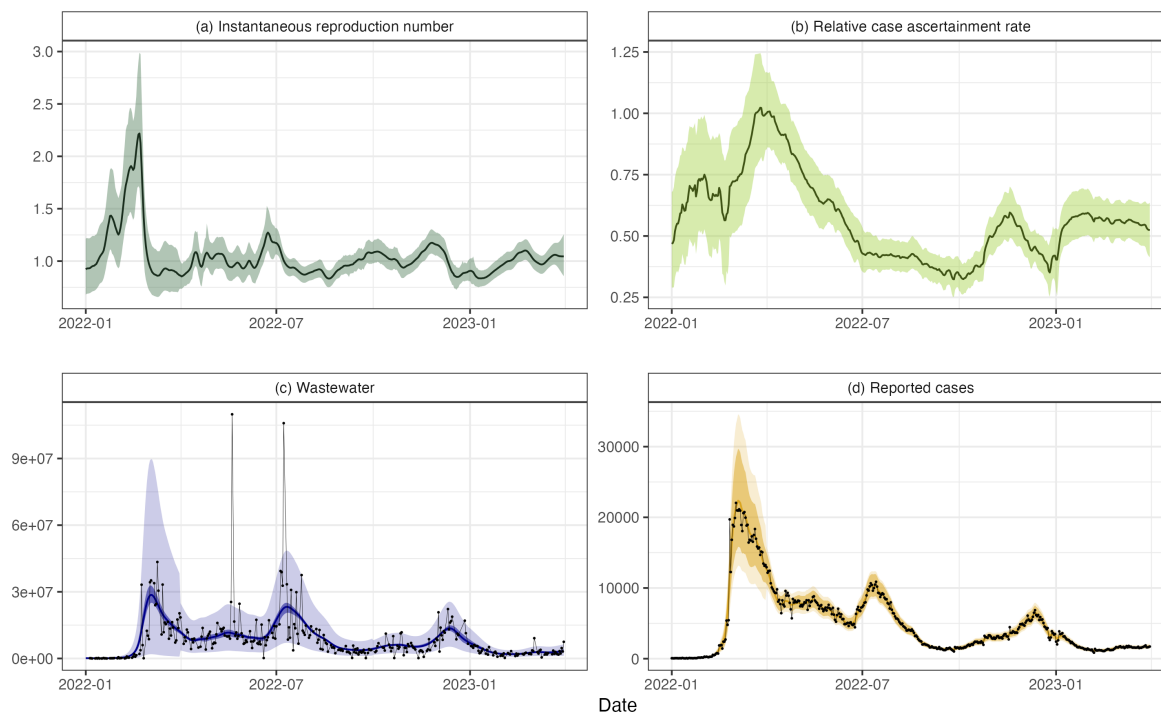


Figure 3: Results for New Zealand data from 1 January 2022 to 31 March 2023. (a) instantaneous reproduction number  $R_t$ , (b) relative case ascertainment rate (compared to the central estimate on 1 April 2022), (c) wastewater data  $W_t$  measured in genome copies per person per day and (d) reported cases  $C_t$ . Results assume the average total shedding per infection does not vary over time ( $\alpha = 3 \times 10^9$ ). Solid lines present central estimates. Shaded regions show 95% credible intervals on the value of the hidden states (subplots a and b), and 95% credible intervals on the expected reported cases and wastewater data (darker shaded regions in subplots c and d) and 95% credible intervals on the prediction distribution for wastewater data and reported cases (lighter shaded regions in subplots c and d). Black dots show the observed data.

231 The estimated CAR (Figure 3b) increased rapidly between mid-February and mid-March 2022.  
232 RATs became widely available for the first time in the last week of February 2022. This likely  
233 led to a significant increase in case ascertainment as the testing system, which had previously  
234 relied solely on laboratory-processed PCR tests, had become overwhelmed [3]. The estimated  
235 CAR approximately halved between April and July 2022, when a second wave of infection  
236 caused by the BA.5 Omicron subvariant [36, 38] occurred. This second wave was visible in both  
237 reported cases and wastewater sampling, with estimated peak infections occurring on 7 July  
238 2022 (95% CrI 3 Jul, 12 Jul). The estimated CAR increased somewhat between mid 2022 and  
239 early 2023, with a noticeable dip in December 2022, possibly reflecting reduced testing during  
240 the Christmas and summer school holiday period (from mid-December to late-January/early-  
241 February). Alternatively, the estimated increase in CAR from mid-2022 could be explained by  
242 a decrease in the average genome copies shed by an infected individual  $\alpha$ , although without  
243 further information we are unable to discern changes in  $\alpha$ . Overall, the model provided a  
244 reasonably good fit to the observed data on cases and wastewater (Figure 3c-d).

245 Figure 4a-b shows the estimated daily incidence and cumulative infections for three values of  
246  $\alpha$ , corresponding to estimated CAR values on 1 April 2022 of 0.42 (95% CrI 0.35, 0.50), 0.61  
247 (95% CrI 0.51, 0.71), and 0.80 (95% CrI 0.67, 0.93), for  $\alpha = 2 \times 10^9$ ,  $3 \times 10^9$ , and  $4 \times 10^9$   
248 respectively. For comparison, the graphs also show the number of cases per capita in a cohort of  
249 approximately 20,000 border workers who were tested weekly between January and July 2022  
250 [36], scaled according to population size. These data were used to help inform the range of  
251 values of  $\alpha$  selected (see Methods).

252 Whilst peak reported cases (adjusted for the day-of-the-week effect) in the second wave were  
253 only 49% of the peak in the first wave (10,879 vs 22,038 respectively), under the assumption  
254 of constant  $\alpha$ , the central estimate from the model suggests that true infections peaked at  
255 approximately 78% of the peak of the initial wave (Figure 4a). Figure 4c-e shows the estimated  
256 absolute and relative CAR and  $R$ . These panels show that, while we are uncertain about the  
257 absolute level of infections and CAR, the relative CAR and reproduction number estimates are  
258 robust to reasonable choices for (constant)  $\alpha$ .

259 Fitting the model to case data alone instead of cases and wastewater (see Supplementary Figure  
260 S7) produced qualitatively similar estimates of  $R_t$ , but with greater temporal fluctuations.  
261 Fitting the model to wastewater data alone led to substantially wider credible intervals, although  
262 the overall trend was similar. Estimates of the relative CAR are only possible when fitting to  
263 case and wastewater data simultaneously.

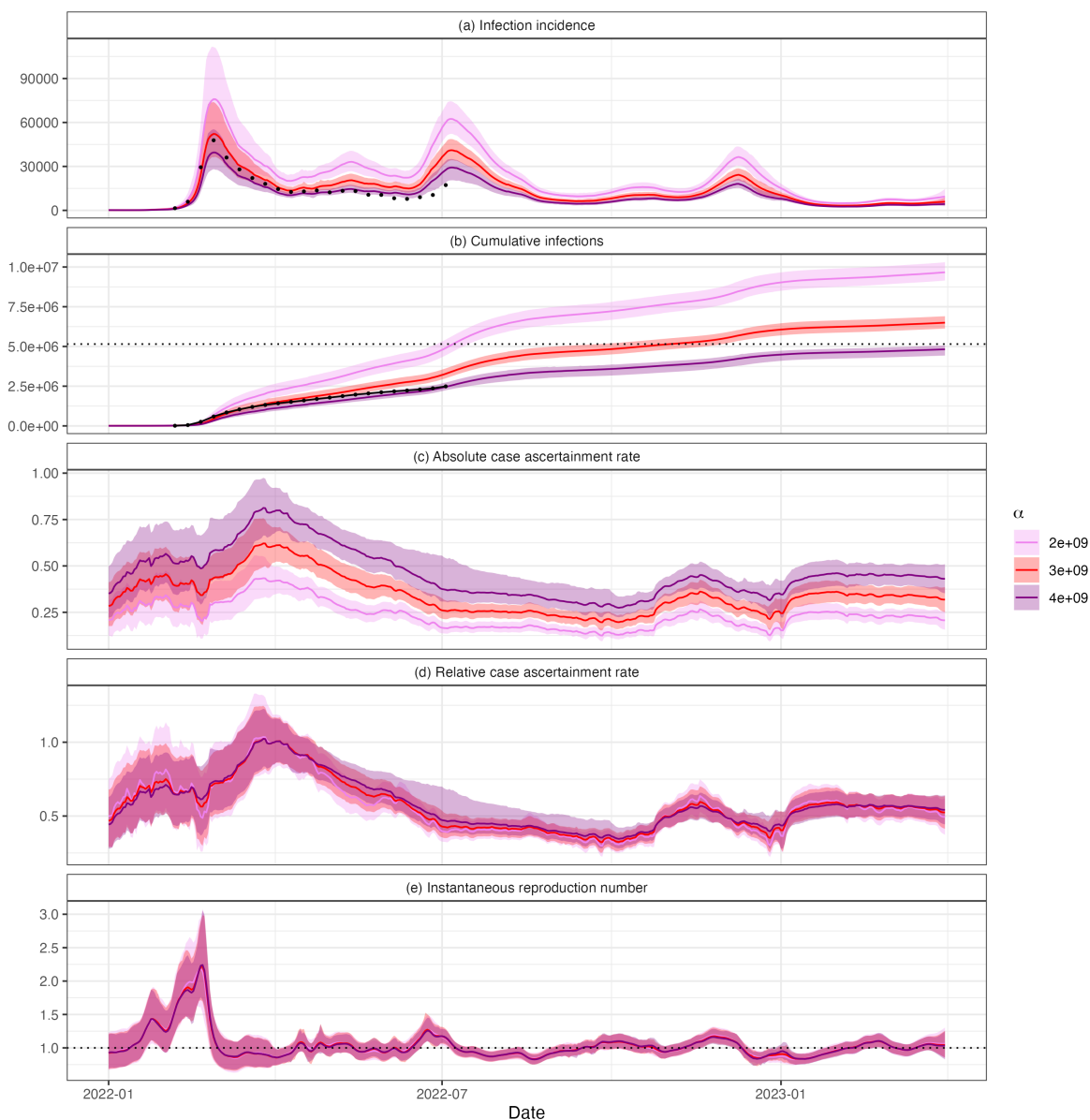


Figure 4: Estimated (a) daily infections  $I_t$ , (b) cumulative infections  $\sum_{s=0}^t I_s$ , (c) case ascertainment rate  $CAR_t$ , (d) relative case ascertainment rate (compared to the central estimate on 1 April 2022), and (e) instantaneous reproduction number,  $R_t$ . Results are presented for three values of  $\alpha$ :  $2 \times 10^9$ ,  $3 \times 10^9$ , and  $4 \times 10^9$ . Solid lines show central estimates and coloured regions are the 95% CrIs. Estimates and credible intervals on cumulative infections are calculated by taking cumulative sums of the estimates and credible intervals in panel (a). Black dots in panels (a) and (b) show the number of per capita cases in a cohort of regularly-tested border workers, scaled according to population size. The horizontal dashed black line in panel (b) shows the New Zealand population at the end of 2022 (5.15 million people) [39]. While changing  $\alpha$  results in different estimates of infections and absolute CAR, the relative CAR and reproduction number estimates are robust to different values, provided  $\alpha$  remains relatively constant.

Table 2: Central estimates and 95% CrIs for estimated model parameters in each time period. Dates in the ‘Period’ column are the start date for the three-month period. All outputs presented to 2 s.f. Higher values of  $\sigma_R$  and  $\sigma_{CAR}$  suggest  $R_t$  and  $CAR_t$  vary faster. Higher values of  $k_c$  and  $k_w$  indicate a lower variance in the corresponding observation distribution. Note a different prior distribution was used for  $\sigma_{CAR}$  in the first period (see Supplementary Material, section 2.5), which may also impact estimates of other parameters in this period.

Period starting	$\sigma_R$	$\sigma_{CAR}$	$k_c$	$k_w (\times 10^{-6})$
1 Jan 2022	0.12 (0.069, 0.21)	0.03 (0.017, 0.043)	31 (20, 49)	1.5 (1.1, 2.0)
1 Apr 2022	0.069 (0.041, 0.12)	0.0099 (0.0053, 0.014)	170 (100, 250)	4.8 (3.2, 6.8)
1 Jul 2022	0.037 (0.02, 0.066)	0.0063 (0.0018, 0.01)	330 (220, 400)	4.8 (3.3, 6.5)
1 Oct 2022	0.038 (0.02, 0.068)	0.011 (0.0073, 0.014)	170 (110, 270)	7.2 (4.7, 10.0)
1 Jan 2023	0.038 (0.018, 0.073)	0.0093 (0.0041, 0.015)	150 (84, 330)	6.8 (4.4, 10.0)

## 264 Parameter estimates

265 The estimated standard deviation  $\sigma_R$  of the random walk on  $R_t$  was greatest in the first time  
266 period (1 Jan – 31 Mar 2022) – see Table 2. This is unsurprising as it coincided with the rapid  
267 increase and then decrease in incidence associated with the first Omicron wave.  $\sigma_R$  decreased  
268 in the second period (1 Apr – 30 Jun 2022) and then remained relatively constant throughout  
269 the remaining periods (1 Jul 2022 – 31 Mar 2023). The estimated standard deviation  $\sigma_{CAR}$  of  
270 the random walk on  $CAR_t$  was also estimated to be greatest in the first time period, although  
271 this is primarily because we applied a prior distribution with a higher mean in this period (see  
272 Supplementary Material section 2.5).

273 The estimated variance parameters,  $k_c$  and  $k_w$ , for cases and wastewater observations, were  
274 lowest in the first time period (1 Jan 2022 – 31 Mar 2022). This implies there is more variability  
275 in the data that is not explained by the model in this time period, possibly as a consequence of  
276 the sharper variations in incidence compared to the later time periods. A less consistent weekly  
277 pattern in reported cases during the first time period, and higher levels of noise in wastewater  
278 observations at the low concentrations seen at the beginning of 2022, could also be contributing  
279 factors.

## 280 4 Discussion

281 Wastewater-based epidemiology has been used globally for COVID-19 surveillance and has been  
282 shown to be a useful public health tool for policy and public health responses [40]. We have

283 presented a semi-mechanistic model that combines reported cases with wastewater data to  
284 estimate the time-varying reproduction number and CAR. This work demonstrates the value of  
285 wastewater-based epidemiology and how the additional data that it provides can be combined  
286 with traditional monitoring (e.g., reported cases) to learn more about the state of an epidemic,  
287 disease dynamics, and the true number of infections in the community. This provides useful  
288 information to inform the public health response.

289 To make reliable estimates of the state of the epidemic from reported cases, it is essential to  
290 understand how case ascertainment changes with time. For example, are there fewer cases  
291 because there are fewer infections or because fewer people are reporting? We applied our  
292 model to national data from Aotearoa New Zealand and derived insights into changes in case  
293 ascertainment that would not be possible using case data alone. Reported cases during the  
294 second wave in July 2022 were significantly lower than in the first wave in February and March  
295 2022. However, the model inferred that there was a substantial drop in case ascertainment  
296 between these waves, and the true number of infections was likely more similar in each wave. The  
297 reduced CAR during the second and subsequent waves may have been due to a higher number  
298 of reinfections with individuals displaying fewer symptoms or due to “pandemic fatigue” and  
299 reduced compliance with public health measures, including testing. This type of insight would  
300 not be possible without regular wastewater surveillance data and without a robust analytical  
301 framework in which to integrate these data with traditional epidemiological data streams.

302 We applied our model to the first period of widespread community transmission of SARS-CoV-  
303 2 in New Zealand. During this time, rapid antigen tests were freely available to everyone,  
304 there was a requirement to report positive results, and a mandatory isolation period for cases  
305 with financial support via employers. Partly as a result of these factors, the CAR, while lower  
306 than in the previous elimination phase, was still reasonably high. The mandatory isolation  
307 period was removed in September 2023, which led to a substantial drop in case ascertainment.  
308 For the datasets we considered, similar (albeit noisier) estimates for the reproduction number  
309 could be obtained from case data alone. However, in a context where case ascertainment is low  
310 and/or unrepresentative, wastewater data are likely to add even greater value compared to using  
311 reported cases. In contrast, in a low-prevalence context (e.g. pre-Omicron in New Zealand),  
312 applicability of the method would be constrained by the amount of noise in the wastewater data.  
313 In this situation, wastewater surveillance may better used for presence/absence monitoring, for  
314 example as an early warning system for the presence of infection in specific catchments, as  
315 opposed to quantitative estimation [41].

316 Strengths of our model include the fact that it has relatively minimal data requirements, re-  
317 quiring only time series for reported cases and wastewater concentrations. The model can be  
318 fitted to datasets in which different sites are sampled on different days and some days have no



319 observed data. This means that it could be readily applied in other jurisdictions with wastewa-  
320 ter surveillance programs, either for SARS-CoV-2 or other pathogens such as influenza viruses  
321 [40, 42]. It is a relatively simple model with minimal mechanistic assumptions and parsimonious  
322 parameterisation. This avoids the need for assumptions about time-varying contact patterns,  
323 transmission rates, and the level of prior immunity that are required by more complex mecha-  
324 nistic models. The model presented here was operationalised by ESR in late 2022 and results  
325 for  $R_t$  and relative CAR are regularly provided to the Ministry of Health to inform situational  
326 awareness and decision-making.

327 There are several limitations to this model and the results. We assumed that the average  
328 number of genome copies shed by an infected individual (represented the parameter  $\alpha$ ) was  
329 constant between January 2022 and March 2023 and did not depend on the infecting variant or  
330 history of prior infection or vaccination. It is possible that some of the inferred changes in CAR  
331 may be partly explained by these factors. For example, some of the inferred increase in case  
332 ascertainment between October and December 2022 may have been due to decreasing  $\alpha$ , caused  
333 by a combination of new immune evasive subvariants displacing the previously dominant BA.5  
334 variant [43] and/or an increase in the proportion of reinfections or asymptomatic infections  
335 [27]. Although estimates of viral shedding rates per infected individual are available [30, 31],  
336 the value of  $\alpha$  may also depend on physical characteristics of the wastewater collection system,  
337 sample collection method, and the method used to quantify concentration of SARS-CoV-2 RNA  
338 in samples. Therefore,  $\alpha$  is likely to vary between jurisdictions and will require recalibration  
339 using local data.

340 As we are unable to estimate the true value of  $\alpha$ , we are unable to estimate the absolute CAR.  
341 Nonetheless, relative CAR is a useful metric and, given an estimated range of values for  $\alpha$ , we  
342 are able to provide plausible bounds on the total number of infections (Figure 4).

343 Wastewater surveillance does not provide any information on how infections are distributed  
344 among population groups (e.g. age groups, ethnicity) and biases in self-administered testing  
345 mean that case counts are not representative either. This information is important for assessing  
346 the clinical burden of disease and addressing health inequities [44]. Thus, other approaches are  
347 needed to determine the distribution of disease burden, such as representative sampling [7, 45],  
348 cohort studies [46] or sentinel surveillance [47, 48]. Although wastewater surveillance could, in  
349 principle, be used to investigate differences in prevalence and case ascertainment between sites  
350 and/or regions, this would require adaptations to our method that are beyond the scope of this  
351 study.

352 As our model is flexible, future work could integrate hospitalisations (such as in [49]) and deaths  
353 data. In principle, this could allow the effects of varying CAR and varying rate of shedding per

354 infection to be separated. However, this would additionally require the effects of age, immunity,  
355 ethnicity, and other variables on clinical severity to be accounted for.

356 Although national-level approaches to situational awareness and reproduction number estima-  
357 tion are common [25, 50, 51], particularly in countries such as New Zealand with a relatively  
358 small population size, this ignores regional variations. Results should therefore be interpreted  
359 as national averages, which could mask demographic and spatial heterogeneity. Our model  
360 could be implemented at a regional level so that local epidemic dynamics can be compared,  
361 although this would be subject to increasing levels of noise in the wastewater data at finer  
362 spatial scales. This paper has focused on modelling for inference: understanding epidemic dy-  
363 namics that have already occurred. However, the state-space transition model coupled with the  
364 estimated parameters provides a natural method for forecasting [23, 52]. Forecasts generated  
365 using this state-space transition model naturally incorporate increasing uncertainty about the  
366 future reproduction number and CAR.

367 While this model has focused on COVID-19, there is a wealth of genetic information within  
368 municipal wastewater that could also benefit from modelling. The detection and concentration  
369 of viral, bacterial and anti-microbial resistance genes within wastewater have the ability to  
370 inform public health decision-making in a number of ways, especially as methodology is refined  
371 allowing more rapid turnaround times. As many jurisdictions seek to retain the wastewater  
372 capabilities they built during the pandemic phase of COVID-19 (and to diversify microbial  
373 targets), there is an ‘opportunity springboard’ to build tools that can predict the trajectories  
374 and spread of pathogens. Modelling has a key role to play in this journey.

## 375 **Data availability**

376 Daily reported case data for Aotearoa New Zealand are available from the Ministry of Health at  
377 <https://github.com/minhealthnz/nz-covid-data> and seven-day average wastewater data  
378 are available from ESR at [https://github.com/ESR-NZ/covid\\_in\\_wastewater](https://github.com/ESR-NZ/covid_in_wastewater).

379 Code to run the model and reproduce the results in this paper are available at <https://github.com/nicsteyn2/NZWastewaterModelling>.

## 381 Acknowledgements

382 The authors acknowledge the role of the New Zealand Ministry of Health in supplying data in  
383 support of this work. The authors thank the wastewater treatment plant staff members who  
384 collected the wastewater samples and the ESR laboratory staff who processed and tested the  
385 samples used in this study. This work was funded by the New Zealand Ministry of Health  
386 and the Department of Prime Minister and Cabinet (DPMC). This work was supported by  
387 the NIHR HPRU in Emerging and Zoonotic Infections, a partnership between PHE, University  
388 of Oxford, University of Liverpool, and Liverpool School of Tropical Medicine (grant number  
389 NIHR200907 supporting C.A.D.). L. M. W. was supported by a Rutherford Foundation Post-  
390 doctoral Fellowship from New Zealand government funding, administered by the Royal Society  
391 Te Apārangi. N.S. acknowledges support from the Oxford-Radcliffe Scholarship from University  
392 College, Oxford, and the Engineering and Physical Sciences Research Council (EPSRC) Centre  
393 for Doctoral Training (CDT) in Modern Statistics and Statistical Machine Learning (Imperial  
394 College London and University of Oxford). We thank A. Maslov for supporting this research  
395 through studentship support for N.S.

## 396 References

- 397 [1] Colman E, Puspitarani GA, Enright J, Kao RR. Ascertainment rate of SARS-CoV-2 in-  
398 fections from healthcare and community testing in the UK. *Journal of Theoretical Biology.*  
399 2023 2;558:111333.
- 400 [2] Eales O, Haw D, Wang H, Atchison C, Ashby D, Cooke GS, et al. Dynamics of SARS-  
401 CoV-2 infection hospitalisation and infection fatality ratios over 23 months in England.  
402 *PLOS Biology.* 2023 5;21(5):e3002118.
- 403 [3] Vattiatio G, Lustig A, Maclaren OJ, Plank MJ. Modelling the dynamics of infection, waning  
404 of immunity and re-infection with the Omicron variant of SARS-CoV-2 in Aotearoa New  
405 Zealand. *Epidemics.* 2022 12;41:100657.
- 406 [4] Parag KV, Donnelly CA, Zarebski AE. Quantifying the information in noisy epidemic  
407 curves. *Nature Computational Science.* 2022 9;2(9):584-94.
- 408 [5] Dawood FS, Porucznik CA, Veguilla V, Stanford JB, Duque J, Rolfes MA, et al. Inci-  
409 dence rates, household infection risk, and clinical characteristics of SARS-CoV-2 infection  
410 among children and adults in Utah and New York City, New York. *JAMA Pediatrics.* 2022  
411 1;176(1):59-67.

- 412 [6] Elliott P, Riley S, Atchison C, Ashby D, Donnelly CA, Barclay W, et al. Real-time As-  
413 sessment of Community Transmission (REACT) of SARS-CoV-2 virus: Study protocol.  
414 Wellcome Open Research. 2020;5:200.
- 415 [7] Pouwels KB, House T, Pritchard E, Robotham JV, Birrell PJ, Gelman A, et al. Community  
416 prevalence of SARS-CoV-2 in England from April to November, 2020: Results from the  
417 ONS Coronavirus Infection Survey. *The Lancet Public Health*. 2021 1;6(1):e30-8.
- 418 [8] Daughton CG. Wastewater surveillance for population-wide Covid-19: The present and  
419 future. *Science of the Total Environment*. 2020;736:139631.
- 420 [9] Dutta H, Kaushik G, Dutta V. Wastewater-based epidemiology: A new frontier for track-  
421 ing environmental persistence and community transmission of COVID-19. *Environmental*  
422 *Science and Pollution Research*. 2022 12;29(57):85688-99.
- 423 [10] Keshaviah A, Diamond MB, Wade MJ, Scarpino SV, Ahmed W, Amman F, et al. Wastew-  
424 ater monitoring can anchor global disease surveillance systems. *The Lancet Global Health*.  
425 2023 6;11(6):e976-81.
- 426 [11] Nourbakhsh S, Fazil A, Li M, Mangat CS, Peterson SW, Daigle J, et al. A wastewater-based  
427 epidemic model for SARS-CoV-2 with application to three Canadian cities. *Epidemics*. 2022  
428 Jun;39:100560.
- 429 [12] Huisman JS, Scire J, Angst DC, Li J, Neher RA, Maathuis MH, et al. Estimation and  
430 worldwide monitoring of the effective reproductive number of SARS-CoV-2. *eLife*. 2022  
431 Aug;11:e71345.
- 432 [13] Huisman JS, Scire J, Caduff L, Fernandez-Cassi X, Ganesanandamoorthy P, Kull A, et al.  
433 Wastewater-based estimation of the effective reproductive number of SARS-CoV-2. *Envi-*  
434 *ronmental Health Perspectives*. 2022 5;130(5):57011.
- 435 [14] Asadi M, Oloye FF, Xie Y, Cantin J, Challis JK, McPhedran KN, et al. A wastewater-  
436 based risk index for SARS-CoV-2 infections among three cities on the Canadian prairie.  
437 *Science of The Total Environment*. 2023 Jun;876:162800.
- 438 [15] Wannigama DL, Amarasiri M, Hongsing P, Hurst C, Modchang C, Chadsuthi S, et al.  
439 COVID-19 monitoring with sparse sampling of sewered and non-sewered wastewater in  
440 urban and rural communities. *iScience*. 2023 Jul;26(7):107019.
- 441 [16] Scire J, Huisman JS, Grosu A, Angst DC, Lison A, Li J, et al. estimateR: An R package  
442 to estimate and monitor the effective reproductive number. *BMC Bioinformatics*. 2023  
443 Aug;24(1):310.

- 444 [17] Jiang G, Wu J, Weidhaas J, Li X, Chen Y, Mueller J, et al. Artificial neural network-based  
445 estimation of COVID-19 case numbers and effective reproduction rate using wastewater-  
446 based epidemiology. *Water Research*. 2022 6;218.
- 447 [18] Pell B, Brozak S, Phan T, Wu F, Kuang Y. The emergence of a virus variant: Dynamics of  
448 a competition model with cross-immunity time-delay validated by wastewater surveillance  
449 Data for COVID-19. *Journal of Mathematical Biology*. 2023 Mar;86(5):63.
- 450 [19] Kisand V, Laas P, Palmik-Das K, Panksep K, Tammert H, Albrecht L, et al. Prediction  
451 of COVID-19 positive cases, a nation-wide SARS-CoV-2 wastewater-based epidemiology  
452 study. *Water Research*. 2023 Mar;231:119617.
- 453 [20] Geubbels ELPE, Backer JA, Bakhshi-Raiez F, van der Beek RFHJ, van Benthem BHB,  
454 van den Boogaard J, et al. The daily updated Dutch national database on COVID-19  
455 epidemiology, vaccination and sewage surveillance. *Scientific Data*. 2023 Jul;10(1):469.
- 456 [21] Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate  
457 time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*.  
458 2013 11;178(9):1505-12.
- 459 [22] Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA,  
460 et al. Improved inference of time-varying reproduction numbers during infectious disease  
461 outbreaks. *Epidemics*. 2019 12;29:100356.
- 462 [23] Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating  
463 the time-varying reproduction number of SARS-CoV-2 using national and subnational case  
464 counts. *Wellcome Open Research*. 2020 12;5:112.
- 465 [24] Fang Z, Roberts AM, Mayer CD, Frantsuzova A, Potts JM, Cameron GJ, et al. Wastewater  
466 monitoring of COVID-19: a perspective from Scotland. *Journal of Water and Health*.  
467 2022;20(12):1688-700.
- 468 [25] McManus O, Christiansen LE, Nauta M, Krogsgaard LW, Bahrenscheer NS, von Kap-  
469 pelgaard L, et al. Predicting COVID-19 incidence using wastewater surveillance data,  
470 Denmark, October 2021–June 2022. *Emerging Infectious Diseases*. 2023;29(8):1589.
- 471 [26] Bertels X, Hanoteaux S, Janssens R, Maloux H, Verhaegen B, Delputte P, et al. Time  
472 series modelling for wastewater-based epidemiology of COVID-19: A nationwide study in  
473 40 wastewater treatment plants of Belgium, February 2021 to June 2022. *Science of The  
474 Total Environment*. 2023;899:165603.
- 475 [27] Ministry of Health. COVID-19 data for New Zealand; 2023. Available from: [https://  
476 github.com/minhealthnz/nz-covid-data](https://github.com/minhealthnz/nz-covid-data).

- 477 [28] ESR. COVID-19 Data Repository by the Institute of Environmental Science and Research;  
478 2023. Available from: [https://github.com/ESR-NZ/covid\\_in\\_wastewater](https://github.com/ESR-NZ/covid_in_wastewater).
- 479 [29] Golding N, Price DJ, Ryan GE, McVernon J, McCaw JM, Shearer FM. A modelling  
480 approach to estimate the transmissibility of SARS-CoV 2 during periods of high, low, and  
481 zero case incidence. *eLife*. 2023 1;12.
- 482 [30] Medema G, Been F, Heijnen L, Petterson S. Implementation of environmental surveillance  
483 for SARS-CoV-2 virus to support public health decisions: Opportunities and challenges.  
484 *Current Opinion in Environmental Science & Health*. 2020;17:49-71.
- 485 [31] Nauta M, McManus O, Franck KT, Marving EL, Rasmussen LD, Richter SR, et al. Early  
486 detection of local SARS-CoV-2 outbreaks by wastewater surveillance: a feasibility study.  
487 *Epidemiology & Infection*. 2023;151:e28.
- 488 [32] Backer JA, Eggink D, Andeweg SP, Veldhuijzen IK, van Maarseveen N, Vermaas K, et al.  
489 Shorter serial intervals in SARS-CoV-2 cases with Omicron BA.1 variant compared with  
490 Delta variant, the Netherlands, 13 to 26 December 2021. *Eurosurveillance*. 2022 2;27(6).
- 491 [33] Hewitt J, Trowsdale S, Armstrong BA, Chapman JR, Carter KM, Croucher DM, et al.  
492 Sensitivity of wastewater-based epidemiology for detection of SARS-CoV-2 RNA in a low  
493 prevalence setting. *Water Research*. 2022 3;211:118032.
- 494 [34] Abbott S, Sherratt K, Gerstung M, Funk S. Estimation of the test to test distribution as  
495 a proxy for generation interval distribution for the Omicron variant in England. *medRxiv*.  
496 2022;10.1101/2022.01.08.22268920.
- 497 [35] Kim D, Ali ST, Kim S, Jo J, Lim JS, Lee S, et al. Estimation of serial interval and  
498 reproduction number to quantify the transmissibility of SARS-CoV-2 Omicron variant in  
499 South Korea. *Viruses*. 2022 3;14(3):533.
- 500 [36] Lustig A, Vattiato G, Maclaren O, Watson LM, Datta S, Plank MJ. Modelling the impact  
501 of the Omicron BA.5 subvariant in New Zealand. *Journal of the Royal Society Interface*.  
502 2023 2;20(199):20220698.
- 503 [37] Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-gaussian Bayesian  
504 state estimation. *IEE Proceedings, Part F: Radar and Signal Processing*. 1993;140(2):107-  
505 13.
- 506 [38] Douglas J, Winter D, McNeill A, Carr S, Bunce M, French N, et al. Tracing the international  
507 arrivals of SARS-CoV-2 Omicron variants after Aotearoa New Zealand reopened its border.  
508 *Nature Communications*. 2022 12;13(1):6484.

- 509 [39] Stats NZ. National population estimates: at 31 December 2022; 2023. Avail-  
510 able from: [https://www.stats.govt.nz/information-releases/  
511 national-population-estimates-at-31-december-2022/](https://www.stats.govt.nz/information-releases/national-population-estimates-at-31-december-2022/).
- 512 [40] Kilaru P, Hill D, Anderson K, Collins MB, Green H, Kmush BL, et al. Wastewater surveil-  
513 lance for infectious disease: a systematic review. *American Journal of Epidemiology*. 2023  
514 2;192(2):305-22.
- 515 [41] Bunce M, Geoghegan JL, Winter D, de Ligt J, Wiles S. Exploring the depth and breadth  
516 of the genomics toolbox during the COVID-19 pandemic: insights from Aotearoa New  
517 Zealand. *BMC Medicine*. 2023;21(1):1-8.
- 518 [42] Toribio-Avedillo D, Gómez-Gómez C, Sala-Comorera L, Rodríguez-Rubio L, Carcereny  
519 A, García-Pedemonte D, et al. Monitoring influenza and respiratory syncytial virus in  
520 wastewater. *Beyond COVID-19. Science of The Total Environment*. 2023 5:164495.
- 521 [43] Prasek SM, Pepper IL, Innes GK, Slinski S, Betancourt WQ, Foster AR, et al. Variant-  
522 specific SARS-CoV-2 shedding rates in wastewater. *Science of the Total Environment*. 2023  
523 1;857.
- 524 [44] Steyn N, Binny RN, Hannah K, Hendy SC, James A, Lustig A, et al. Māori and Pacific  
525 people in New Zealand have a higher risk of hospitalisation for COVID-19. *New Zealand  
526 Medical Journal*. 2021;134:1538.
- 527 [45] Riley S, Ainslie KEC, Eales O, Walters CE, Wang H, Atchison C, et al. Resurgence of  
528 SARS-CoV-2: Detection by community viral surveillance. *Science*. 2021 5;372(6545):990-5.
- 529 [46] Huang QS, Wood T, Jelley L, Jennings T, Jefferies S, Daniells K, et al. Impact of the  
530 COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infec-  
531 tions in New Zealand. *Nature Communications*. 2021 12;12(1):1001.
- 532 [47] Zambon MC, Stockton JD, Clewley JP, Fleming DM. Contribution of influenza and respi-  
533 ratory syncytial virus to community cases of influenza-like illness: An observational study.  
534 *Lancet*. 2001 10;358(9291):1410-6.
- 535 [48] Eales O, Plank MJ, Cowling BJ, Howden BP, Kucharski AJ, Sullivan SG, et al. Key chal-  
536 lenges for respiratory virus surveillance while transitioning out of acute phase of COVID-19  
537 pandemic. *Emerging Infectious Diseases*. 2024;30(2).
- 538 [49] Schenk H, Heidinger P, Insam H, Kreuzinger N, Markt R, Nägele F, et al. Prediction  
539 of hospitalisations based on wastewater-based SARS-CoV-2 epidemiology. *Science of the  
540 Total Environment*. 2023 5;873:162149.

- 541 [50] Brockhaus EK, Wolfram D, Stadler T, Osthege M, Mitra T, Littek JM, et al. Why are  
542 different estimates of the effective reproductive number so different? A case study on  
543 COVID-19 in Germany. *PLOS Computational Biology*. 2023;19(11):e1011653.
- 544 [51] Plank MJ, Watson L, Maclaren OJ. Near-term forecasting of Covid-19 cases and hospital-  
545 isations in Aotearoa New Zealand. *PLOS Computational Biology*. 2024;20(1):e1011752.
- 546 [52] Moss R, Zarebski A, Dawson P, McCaw JM. Retrospective forecasting of the 2010-2014  
547 Melbourne influenza seasons using multiple surveillance systems. *Epidemiology and Infec-  
548 tion*. 2017 1;145(1):156-69.