Evaluating the theoretical performance of aircraft wastewater monitoring as a tool for SARS-CoV-2 surveillance

Joseph W. Shingleton, MSc\(^a\); Chris J. Lilley, BA\(^a\), Matthew J. Wade PhD\(^a\)

\(^a\) Analytics & Data Science Directorate, UK Health Security Agency, Nobel House, Smith Square, London, SW1P 3JR, UK
\(^\dagger\) corresponding author

Email address: joseph.shingleton@ukhsa.gov.uk (Joseph Shingleton)

Abstract

Background
Air travel plays an important role in the cross-border spread of infectious diseases. During the SARS-CoV-2 pandemic many countries introduced strict border testing protocols to monitor the incursion of the virus. However, the high implementation cost and significant inconvenience to passengers has led public health authorities to consider alternative methods of disease surveillance at borders. Aircraft wastewater monitoring has been proposed as one such alternative. In this paper we assess the theoretical limits of aircraft wastewater monitoring and compare its performance to post-arrival border screening approaches.

Methods
We use an infectious disease model to simulate an unmitigated SARS-CoV-2 epidemic in a seed country. Seeding of the epidemic into the United Kingdom (UK) is simulated through daily flights between the two countries. We use a probabilistic approach to estimate the time of first detection of the disease in wastewater and respiratory swab screening at the border.

Results
For simulations across a broad range of model parameters, our analysis indicates that the median time between the first incursion of a pathogen and its detection in wastewater would be approximately 17 days (IQR: 7 - 28 days), resulting in a median of 25 cumulative cases (IQR: 6 - 84 cases) in the UK at the point of detection. Comparisons to respiratory swab screening suggest that aircraft wastewater monitoring is as effective as screening of 20% of passengers at the border, using a test with 95% sensitivity. For testing regimes with sensitivity of 85% or less, the required coverage to outperform wastewater monitoring increases to 30%. These results demonstrate the potential use cases...
of aircraft wastewater monitoring and its utility in a wider system of public health surveillance.

1. Introduction

Air travel has acted as a major conduit for international transmission of SARS-CoV-2 throughout the pandemic (Liu et al., 2022). Almost all countries have, at some point, introduced degrees of border control, which have prevented inbound travel from specific locations, added requirements for evidence of a recent negative test or vaccination prior to departure, or required testing of passengers upon arrival (Grépin et al., 2021; Movsisyan et al., 2021; Shiraef et al., 2022). Many of these requirements have been removed in the majority of countries due to a combination of factors, including successful vaccination programmes and a need to limit further economic disruption (Sun et al., 2021; Shiraef et al., 2022).

Prior to removal, border screening programmes provided rich data on the importation of SARS-CoV-2 (Williams et al., 2021), and helped to map the international prevalence of new variants (Douglas et al., 2022). With the reduction in large-scale screening of inbound passengers, many countries have begun to adopt monitoring of aircraft wastewater as a potentially non-intrusive alternative for monitoring pathogens entering a country or region (Ahmed et al., 2022; Le Targa et al., 2022; Farkas et al., 2023).

There exist no standard protocols for aircraft wastewater surveillance and establishing the value of such a function for public health will require extended research and development. Subsequently, several significant limitations on its use remain, not least its acuity in capturing individuals harbouring pathogen(s) of interest, typically not an issue with gross population wastewater sampling. For wastewater monitoring to capture a single infected individual a series of conditions must be met. First, the individual must defecate during the flight - the probability of which varies with the length of the flight. Survey data suggests around 13% of individuals defecate during a short haul (<6 hours) flight, increasing to 36% for long-haul flights (Jones et al., 2023). Additionally, the individual must be shedding genetic markers (i.e., RNA or DNA) of the pathogen(s) in their faeces - the probability of which varies between and within pathogen types; however, it is thought to occur in between 30% to 60% of cases for SARS-CoV-2 (Jones et al., 2023). Finally, the concentration of viral RNA in the aircraft wastewater must exceed the limit of detection (LoD) for Reverse Transcription Polymerase Chain Reaction (RT-PCR) testing. Recent studies show that community wastewater sampling can identify changes in COVID-19 incidence as low as 3.8 new cases per 100,000 (Li et al., 2023), suggesting this limit is likely very low. These limitations mean that the probability of identifying
any one infected individual on a single flight is very low (Jones et al., 2023). As such, wastewater monitoring is not suitable for estimating the on-board prevalence of a disease. Rather, aircraft wastewater monitoring as a binary indicator of the presence of a pathogen on board a flight is more appropriate. One such use case for this might be to estimate the point-in-time of incursion of a new pathogen (or variant of a pathogen) into a country or region - potentially providing key insights for disease modelling and further health protection measures. To assess the feasibility of using aircraft wastewater in such a capacity the likely delay between the first incursion of a pathogen, and when it is likely to be detected in wastewater, must be first understood.

This paper aims to establish the likely length of any delay between incursion and detection, estimate the cumulative cases in the destination country at the point of detection, and to compare the effectiveness of aircraft wastewater monitoring to post-arrival respiratory swab screening. We consider SARS-CoV-2 as a target pathogen for our analysis; however, the results may be applicable to a wide range of diseases.

We use a two-patch infectious disease model to investigate the possible outcomes of aircraft wastewater monitoring. The model simulates flights between a seed (origin) country and the UK and uses a probabilistic approach to identify the first point at which wastewater monitoring identifies the presence of infected individual(s) on board. We then compare these results to simulated respiratory swab screening, in which a proportion of the passengers are tested directly for SARS-CoV-2 at the border.

Our results shed light on some of the practical considerations when using aircraft wastewater monitoring for public health surveillance. By comparing to respiratory swab screening regimes, we are able to highlight how wastewater monitoring can act as a cost effect, non-invasive approach to border pathogen surveillance with minimal impact on airline passengers.

2. Methods

2.1. SEI(F)R model

We use an SEIR (Susceptible, Exposed, Infectious, or Recovered) model to simulate an unmitigated epidemic in a seed country, as detailed in equations 1-5. The model is parameterised to have similar dynamics to SARS-CoV-2, with a latent period of $\alpha^{-1} = 5.2$ days, a recovery period of $\gamma^{-1} = 8$ days and $R_0 = \gamma/\beta = 2.5$, where $\beta$ is the transmissibility constant. Importantly, the infected compartment is split into two stages; an initial infected state, denoted $I_1$, in which individuals can transmit the disease and test positive in both
respiratory swabs and faecal samples, and a faecal positive state, denoted $F$, in which individuals are no longer infectious and test negative in respiratory swabbing but continue to test positive in faecal samples.

\[
S'(t) = -\beta SI/N \tag{1}
\]
\[
E'(t) = \beta SI/N - aE \tag{2}
\]
\[
I'(t) = aE - \gamma I \tag{3}
\]
\[
F'(t) = \gamma I - \delta F \tag{4}
\]
\[
R'(t) = \gamma F \tag{5}
\]

While there is evidence that a post symptomatic faecal shedding state exists for SARS-CoV-2 (Liu et al., 2022; Grépin et al., 2021; Movsisyan et al., 2021), there is little consensus on the mean length of time people spend in this state, denoted $\delta^{-1}$ in the model, with estimates of between 3 and 35 days provided in the literature (Liu et al., 2022; Movsisyan et al., 2021). As such, we will consider how the model evolves over a range of values for $\delta$.

2.2. Air-travel model

We simulate two epidemics, one in a seed country with population $N_1 = S_1 + E_1 + I_1 + F_1 + R_1 = 10^6$, seeded with a single infection at $t = 0$, and an initially unseeded epidemic in the UK with population $N_2 = S_2 + E_2 + I_2 + F_2 + R_2 = 6.7 \times 10^7$. Each day, we simulate a single return trip between the two countries by taking a sample of $n$ people from each population and transferring them to the other population. By doing this we assume that aircraft passengers are a representative sample of the current state of the epidemic in their country of origin.

2.3. Detection of viral RNA in aircraft wastewater

Each flight contains some number of people in each disease state, with the total number of people on the flight given as $n = n_s + n_F + n_I + n_F + n_R$. The probability of a single passenger being in each state is equal to the prevalence of that state in the country of origin. Each passenger has a probability $p_d$ of defecating during the flight. Each defecating passenger in the $I$ and $F$ compartments has a probability $p_s$ of faecal shedding. Hence, the total number of stools produced during the flight is drawn from the binomial distribution $s \sim B(n, p_d)$, of which $s_{IF} \sim B(s, (n_I + n_F)/n)$ come from individuals in states $I$ and $F$, and $s_p \sim B(s_{IF}, p_s)$ test positive for SARS-CoV-2. The model assumes
that the probability of defecation is independent of an individuals’ infection state.

2.4. Limit of detection

There is a theoretical concentration of target RNA below which wastewater testing is unable to reliably identify the presence of the target pathogen. Changes in test-confirmed COVID-19 case rates as low as 3.8 new cases per 100,000 population can be identified in community wastewater with 99% probability (Li et al., 2023). However, this value is likely an under estimation of the true LoD as it relies only on test positive cases (admittedly over a period with very good test coverage), and is a measure of changes in incidence, rather than of prevalence. However, we should also note that aircraft wastewater is much more highly concentrated and specific than community wastewater, so detection is likely easier (Ahmed et al., 2022).

The lowest possible (non-zero) concentration of viral RNA achievable in our model occurs when every passenger on board deposits a stool into the wastewater, with just one of those stools containing viral RNA. In this ‘worst-case’ scenario, the LoD would have to be several orders of magnitude lower than that found previously (Li et al., 2023) for sampling to return a negative result. This is corroborated in a recent study in which aircraft wastewater samples were compared against PCR screening of passengers, finding that positive wastewater samples were returned when a single infected individual was on board (Ahmed et al., 2022).

As such, we will initially assume that wastewater sampling is sensitive enough to return a positive result when a single infected individual successfully transfers viral RNA into an aircrafts’ wastewater. We will also consider the sensitivity of our results to limits of detection of $0.0025 \leq L \leq 0.03$, measured as the ratio of infected stools, $s_p$, to total stools, $s$, in the wastewater system. In such cases, wastewater samples will only test positive if $s_p/s \leq L$.

2.5. Respiratory swab screening

To assess the performance of aircraft wastewater monitoring we will compare our results to the theoretical performance of post-arrival screening via respiratory swabbing of passengers. Respiratory testing is limited to capturing individuals in the $I$ compartment and will not detect those in the $F$ compartment. We will consider scenarios in which the proportion of passengers screened is $0.1 \leq \phi \leq 0.5$, with test sensitivity of $0.65 \leq \eta \leq 0.99$. We will draw the number of infected individuals selected for respiratory swabbing from a binomial distribution such that $n_i \sim B(n_i, \phi)$ infectious passengers are tested,
resulting in $r_p \sim B(r_f, \eta)$ positives. We do not consider the occurrence of false positives in either the respiratory swabbing or wastewater testing.

### 2.6. Initial parameterisation

For the initial results, we simulate the model 10,000 times using the parameterisation described in Table 1. Model parameters $a, \beta, \gamma, L, \phi$ and $\eta$ are kept constant throughout all simulations. We draw the values of $p_d$ and $n$ from uniform distributions such that $0.1 \leq p_d \leq 0.6$ and $100 \leq n \leq 400$, reflecting flights of different lengths and different passenger numbers. Due to the lack of consensus in the literature about the length of the post symptomatic faecal shedding period, $\delta^{-1}$, this value is also drawn from a uniform distribution of integer values such that $3 \leq \delta^{-1} \leq 28$. We draw the faecal shedding probability, $p_s$, from a beta distribution such that $p_s \sim Beta(a, b) \in [0,1]$, where:

$$a = \left(1 - \frac{1}{\sigma^2} - \frac{1}{\mu}\right)\mu^2,$$

and

$$b = a\left(\frac{1}{\mu} - 1\right).$$

where $\mu = 0.459$ and $\sigma = 0.168$ are the mean and standard deviation of the distribution (Jones et al., 2023).

### Table 1: Model parameterisation for results in Figure 1. The post symptomatic faecal shedding period, $\delta^{-1}$, number of people on flight, $n$, and probability of defecation, $p_d$, are drawn from uniform distributions. The faecal shedding probability, $p_s$, is drawn from a Beta distribution with parameters $a$ and $b$ set such that the distribution has mean 0.459 and standard deviation 0.168.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Latency rate</td>
<td>$5.2^{-1} \text{days}^{-1}$</td>
</tr>
<tr>
<td>$R_0$</td>
<td>Reproductive number ($\beta/\gamma$)</td>
<td>3.0</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Recovery rate</td>
<td>$8.0^{-1} \text{days}^{-1}$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Post symptomatic faecal shedding recovery rate</td>
<td>$\delta^{-1} \sim U(3,28) \text{days}$</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of people on flight</td>
<td>$n \sim U(100,400)$</td>
</tr>
<tr>
<td>$p_d$</td>
<td>Probability of defecation</td>
<td>$p_d \sim U(0.1,0.6)$</td>
</tr>
<tr>
<td>$p_s$</td>
<td>Probability of faecal shedding</td>
<td>$p_s \sim Beta(a, b)$</td>
</tr>
<tr>
<td>$L$</td>
<td>Limit of detection (LoD)</td>
<td>0.005</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Proportion screened using respiratory swabs</td>
<td>0.20</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Sensitivity of respiratory swabs</td>
<td>0.85</td>
</tr>
</tbody>
</table>
2.7. Parameterisation for sensitivity analysis

When investigating the sensitivity of the model to specific parameters all other parameters are held constant. For these analyses we set $n = 250$, $\delta^{-1} = 10$ days, $p_d = 0.36$, $p_s = 0.49$ (as used in Jones et al., 2023), and $\eta = 0.3$, all other parameters are as described in Table 1. This parameterisation minimises the delay between the median date of detection in wastewater and respiratory swab screening to zero days.

3. Results

3.1. Monte-Carlo simulation with broad distribution of parameters

Figure 1 shows the time of first seeding into the UK, the time of first detection in aircraft wastewater, and the time of first detection in respiratory swabs from 10,000 simulations of the model with the parameterisation described in Table 1. The analysis suggests that, for the given range of parameters, the median date of detection in wastewater is two days earlier than for respiratory swabs (wastewater median 94 days, inter-quartile (IQ) range 85 - 102 days; respiratory swab median 96 days, IQ range 88 - 103 days). There was a median delay of 17 days (IQ range: 7-28 days) between first seeding in the UK and detection in wastewater, resulting in a median of 25.2 (IQ range: 4.8 - 78.2) cumulative infections in the UK, including imported cases. At the point of detection via respiratory swabbing there were a median of 33.4 (IQ range: 9.3 - 24.2) cumulative infections in the UK.

![Figure 1](https://example.com/figure1.png)

Figure 1: The distribution of results from 10,000 simulations of the model, using the parameter ranges given in Table 1. The black distribution is the time of first incursion in the UK, given as the number of days since the index case in the country of origin. The blue and orange distributions show the time of detection in wastewater and respiratory swabbing, respectively. The dashed lines show the median value for each distribution.
3.2. Effect of screening coverage and sensitivity

Figure 2 shows the delay between detection in wastewater and detection using respiratory swabs across screening coverages and screening sensitivities over 2000 model simulations. All other parameter values and ranges are as indicated in Table 1. A negative value indicates that wastewater detection occurs before detection via respiratory swabs. For test sensitivities $\eta \leq 0.85$, a screening coverage of is required to outperform wastewater testing in more than half of the simulations. As test sensitivity increases (for example by using nasopharyngeal swabbing rather than nasal swabbing) this proportion falls to for .

![Figure 2: Delay between detection in wastewater and detection via respiratory swabs for different values of respiratory swab coverage (\(\phi\)) and sensitivity (\(\eta\)) over 2000 model simulations. Other variables are as indicated in table 1. A negative value indicates detection in wastewater occurs prior to detection in respiratory swabs. The central blue line gives the median value of 2000 simulations of the model, the blue shaded areas give the 25-75% and 5-95% intervals. A comparison of the median delay across each value of \(\eta\) and is shown in the heat-map in the final plot.](image)

3.3. Sensitivity to model parameters

Figure 3 shows the delay between detection in wastewater and in respiratory swab screening across a range of values for LoD, defection probability, shedding probability and post symptomatic faecal shedding period for 2000 model simulations. Other parameter values remain the same as stated in section 2.7.
Analysis of the LoD parameter, suggests that values of have little effect on model output. However, as LoD increases, a significant reduction in performance relative to respiratory swabbing is observed. However, as discussed previously, it is likely that the true LoD is extremely low - with the estimate by Li et al. suggesting (Li et al., 2023).

Defecation probabilities of result in detection in respiratory swabbing earlier than in wastewater in more than half of simulations. For respiratory swabs outperform wastewater monitoring in around 75% of simulations. As is directly correlated with flight duration we can infer that wastewater monitoring is likely to be more effective for longer flights.

The faecal shedding probability, , and post-symptomatic faecal shedding period, , relate to the characteristics of the disease. Analysis of these parameters may give an indication of the effectiveness of wastewater monitoring in the surveillance of other diseases. Wastewater monitoring is less effective than respiratory swab screening when and slightly more effective for . Changes in the post-symptomatic faecal shedding period have very limited influence on the effectiveness of wastewater monitoring. This is likely because the total number of people in the and compartments at the point of detection is very low so the influence of is minimal. It is likely that later in the epidemic would have a more significant effect on detection.

Figure 3: Model outputs for 1000 model simulations across a range of values of (a) LoD, (b) defecation probability, (c) shedding probability and (d) post-symptomatic faecal shedding period. All other parameters are as stated in section 2.6.
1. Discussion

As the appetite for strict border control measures falls, public health authorities are left with limited choices in how to monitor the incursion of SARS-CoV-2, and other infectious diseases, across borders. This paper has investigated the extent to which aircraft wastewater monitoring might be able to bridge the gap left by reduced respiratory screening. We have shown that, for a broad range of parameterisations of our model, respiratory swab screening begins to outperform wastewater monitoring when screening covers around 20% of passengers. However, sensitivity analysis has identified several factors which should be considered when using wastewater monitoring as a surveillance tool.

The length of flight is a clear influence on the viability of aircraft wastewater monitoring since individuals are more likely to defecate during a longer flight (Jones et al., 2023). Unfortunately, there is no reliable data on the specific relationship between flight time and probability of defecation, although one can reasonably assume that the relationship is positive. Further work should aim to estimate the value of $p_d$ for a given flight-time, based on the distribution of times of day individuals tend to defecate (Heaton et al., 1992).

The sensitivity of the wastewater sampling technique is also important. For sampling methods which increase the limit of detection past around 0.01 infected stools per total stools we see a sharp reduction in the performance of wastewater monitoring. Further work should attempt to identify this value in a biologically relevant unit, such as gene copies per ml, however this is outside the scope of this work.

The disease, or variant of disease, may also play a role in the practical application of aircraft wastewater monitoring. The probability of shedding viral RNA into faeces has a significant effect on the viability of wastewater monitoring - a disease with faecal shedding probability of $p_s < 0.4$ is unlikely to be suitable for wastewater monitoring. The sensitivity analysis was unable to identify a strong relationship between post-symptomatic faecal shedding period and performance of wastewater monitoring for the parameter ranges considered in the analysis.

There is typically some delay between the sampling of aircraft wastewater and the sample being tested through RT-qPCR. We have not considered this
delay in our model, instead assuming a test can be returned instantaneously. For cases where aircraft wastewater is being used to retrospectively estimate the point in time of incursion of a disease this assumption is valid, however the delay should be considered for cases where wastewater is being used in a more reactive capacity. Respiratory screening at the border can return (almost) instant indications of traveller positivity - although in practice during the SARS-CoV-2 pandemic many countries opted for testing recently arrived passengers a set number of days after arrival.

This paper has only considered the time to detection as a metric for measuring the value of wastewater monitoring. While we have shown that wastewater can be more effective than limited respiratory screening at detecting the presence of SARS-CoV-2 on board flights, there are other factors which have not been considered. One of the strengths of wastewater monitoring is that sample taking is agnostic to the target pathogen, so a single sample can be tested for multiple diseases. As the same degree of flexibility is not afforded by respiratory swabbing it may be that wastewater monitoring is preferred to border screening even in situations where screening would likely detect the presence of a pathogen more quickly.

As we move away from strict border controls as a means of mitigating the international transmission of pathogens like SARS-CoV-2 it is important that we understand the tools which remain available to us. The results discussed in this paper lay some of the theoretical groundwork in doing that for aircraft wastewater monitoring. We have shown that, under specific circumstances, aircraft wastewater monitoring can be an effective tool in global public health surveillance.

4. Conclusion

The model presented in this paper provides valuable insights into the potential viability of aircraft wastewater monitoring as a tool for public health surveillance. We have shown that, for simulations across a broad parameterisation of our model, there is a delay of around 17 days between the first incursion of a SARS-CoV-2 like pathogen and its detection in aircraft wastewater, and that this delay results in a median of 22 cumulative infections in the UK. Further, we have shown that wastewater screening can detect a new pathogen earlier than respiratory swab screening of 30% of passengers using a test with 85% sensitivity. This result, however, is dependent on a number of factors, with the probability of defecation during a flight (assumed to be correlated to flight time), the probability of shedding pathogen into faeces and
the sensitivity of wastewater testing all having significant influence on the viability of wastewater monitoring.

**Acknowledgements**

We would like to thank the Advanced Analytics Division of the Analytics and Data Science Department at UKHSA for their helpful comments and suggestions during development of this model.

**Author contributions**

Joseph Shingleton: methodology, software, validation, formal analysis, investigation, visualization, writing – original draft. Chris Lilley: supervision, writing – review & editing. Matthew Wade: supervision, conceptualization, writing – review & editing.

**Sources of Funding**

The authors received no specific funding for this research.

**Data availability statement**

No new data were generated or analysed in support of this research. The code developed to perform the analysis is available on reasonable request to the corresponding author.

**Conflicts of Interest**

The authors have declared no conflicts of interest.
References


