COVID-19 epidemiological surveillance using sewage testing – study proposal

Introduction

Sewage testing has been successfully used in the population-wide surveillance of infectious diseases such as poliovirus, hepatitis A and norovirus, particularly in early detection of outbreaks and complementing routine surveillance in cost-effective ways[i, ii]. Recent reports confirmed detection of SARS-CoV-2 viral fragments in human stool with RT-PCR, including among asymptomatic and symptomatic cases and begin to characterise shedding patterns such as the prevalence, timing of onset, duration, viral load and relationship to nasopharyngeal and respiratory shedding and clinical profile.[iii, iv, v, vi, viii, ix] SARS-CoV-2 has been successfully detected in municipal wastewater from large urban centres in the Netherlands[x], France[xi], the USA[xii] and Brisbane in Australia.[xiii]

These emerging data identify sewage testing as a promising tool for environmental surveillance of COVID-19 including community/population level sentinel and ongoing surveillance with wastewater based epidemiology. After refinement and standardization in testing, sampling and analysis methodologies[xiv], this may help track early start, evolution and impact of lockdown/easing and intensified preventive measures in targeted geographic areas. It may also be useful in the Victorian epidemic (with low caseloads) for sentinel surveillance within high risk and/or high priority locations (e.g. abattoirs, prisons or other such institutional populations) to target active case finding efforts.

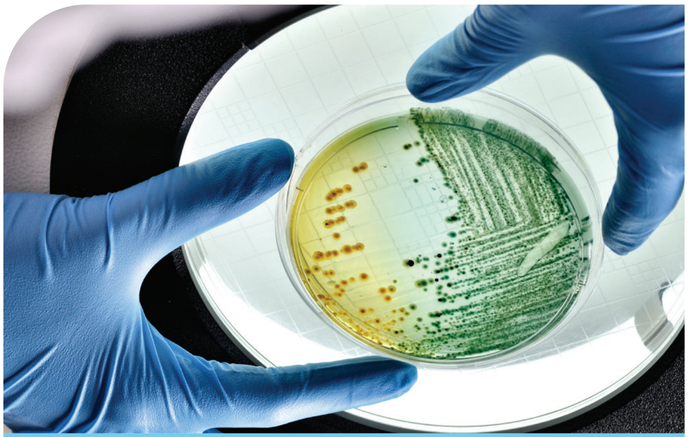
This proposal outlines the epidemiological methods to assess the ability of sewage testing to contribute to innovative surveillance of COVID-19 and associated public health interventions in Victoria, Australia as part of the Victoria node of the national sewage surveillance project ColoSSoS lead by Water Research Australia (see deliverables in box at right) .

Research aims (and related Surveillance Questions of Public Health importance)

Sewage testing may contribute to the surveillance and related public health response of COVID-19 in four ways, which will form the basis for this research work.

1. **Describe trends over time in the qualitative and quantitative detection of SARS-CoV-2 in defined Victorian sentinel sewerage sampling locations and correlate with reported case trends and population level prevention measures.** *Note quantitative results may not be available until validation of testing methodologies so initial analysis may be limited to qualitative RT-PCR (present/absent) results.*

- Is there a correlation between reported diagnosed cases and measurement (qualitative and quantitative) in sewage in municipal and regional catchment areas?



ColoSSoS project deliverables (Vic)

Short term | Development of robust and reliable sampling, analysis, interpretation and integration methods for SARS-CoV-2 sewage surveillance results and health data from across Victoria.

Medium term | Routine sampling and molecular analysis for SARS- CoV-2 sewage surveillance

Long term | Framework for rapidly instituting sewage surveillance for future outbreaks of novel pathogens.

- Using a case-control approach – Is there an optimal case and exposure definition to maximize the odds ratio (utility for surveillance purposes)
- What is the time-lag between community sewage SARS-CoV-2 results and a) case reports and b) change in restrictions?
- Does geographic distance between cases and the sampling point influence the likelihood of a positive sewage result?
- Does community sewage surveillance suggest an impact of community-level preventive measures (e.g. lockdown, easing of restrictions)?

COVID-19 epidemiological surveillance using sewage testing – study proposal

2. **Identify geographic locations where there is a mismatch between caseloads and sewage tests which suggests undiagnosed cases as part of an early sentinel surveillance system and link to surge testing/active case finding as well as other possible preventive public health interventions.**

- Identify any geographic catchment areas which are positive for SARS-CoV-2 in sewage in the absence of known cases? For geographic catchment areas with no recently diagnosed or active cases, any detection of SARS-CoV-2 (using highly specific testing methods) shows undiagnosed case or cases in that catchment area – this is likely to be relevant to the rural/regional centres rather than large municipal areas.
- Identify any geographic catchment areas for which SARS-CoV-2 in sewage is quantitatively higher than that expected based on the known number of cases in the catchment area; for geographic catchment areas with recently diagnosed or active cases (requires quantitative RT-PCR data).

3. **Prospective: Identify high-risk or high-priority populations and associated sewage sampling and establish sentinel sewage surveillance for early detection (may include abattoirs, prisons, hospitals, aged care facilities, schools and the like).**

- Explore feasibility:
    i. from a sewage sampling perspective?
    ii. from a lab testing/cost perspective perhaps using pooled samples as a cost-saving strategy?
- As for 2 above – any positive detection in sewage in the absence of known human cases would trigger expanded testing/intervention at the targeted facility (with potential to prevent/mitigate a large outbreak such as was seen in Victorian and international abattoirs, prisons and hospital settings, and provide reassurance regarding the safety of school openings).

4. **Exploratory/prospective: Explore feasibility of genomic sequencing surveillance among wastewater viral fragments to evaluate genetic relatedness and characterise outbreaks in the Victorian and Australian population**.

- Actual use of sewage testing for surveillance and public health will depend on careful analysis of costs and benefits, but before this assessment can be made it is necessary to discover whether sewage testing would actually be useful in the above ways.

Research methods

1. Descriptive analyses

Descriptive analyses will include frequency distributions, time series, mapping of cases by location by day, overlaid with sewage testing results and catchment area showing key dates for change in population level restrictions. For municipal Melbourne data analyses will also show the geographic distribution of sewage catchment areas with overlay of sub-geographic catchment areas within larger ones.

- Correlation coefficients between the sewage trends and case trends will be defined including best fit adjustment for time-lag.
- Should quantitative data become available - Modelling with defined assumptions (noting methods used in Brisbane and Paris estimations) of quantitative sewage results to estimate no (median and IQR) of expected cases in catchment area and identify mismatch with actual cases.
    - Sensitivity analysis for extremes of key variables (faecal shedding patterns, viral decay etc)

For study AIM 1 and 2

2. Case-control study[1]

A case-control study will be conducted to answer the following initial research questions:

a. To determine the strength of association between a positive sewage test and the presence of active human cases in the same catchment.

b. To define two components of 'exposure': (i) the duration of shedding period for an active human case, and (ii) distance of the cases' residence from the sampling site, which correlate the best to a positive sewage test (i.e. a 'case' catchment-day).

In this case-control approach the 'cases' are days on which at least one positive virus test is registered at a waste-water plant or sewage outflow from a high-risk site. 'Controls' are days when negative test results are received. 'Exposure' means the presence of at least one known active case of COVID-19 in a human in the catchment area of the same waste-water plant.

---

[1] *Inspired by the thunderstorm asthma study done in DHHS Victoria in 2018. That study, now under peer review, defined a case as a hospital-day with very high asthma presentations, and controls were randomly selected from hospital-days without high asthma presentation counts. Exposure was proximity to a meteorological phenomenon called a 'convergence line'. Odds Ratios up to 179 were found, and it was established that exposure to a convergence line is a necessary-but-not-sufficient precursor to a thunderstorm asthma event.*

Case control studies

Odds Ratios are calculated thus:

	Case days (Positive sewage test[s])	Control days (Negative sewage test[s])
Exposed (Human case(s) in catchment)	a	b
Unexposed (No human cases in catchment)	c	d

Odds Ratio = (a/c) / (b/d)

The Odds Ratio[2] (OR) as described in this study can be interpreted as "The odds of there being at least one active human case in the catchment of a waste-water plant are 'X' times higher on a day when a positive sewage test is recorded". If the OR equals 1 there is no association. If it is greater than 1 there is a positive association. If it is less than 1 the odds of finding a human case in the catchment are lower when sewage tests positive (unlikely in this case). The Odds Ratio has a 95% Confidence Interval (CI), and if the CI does not include the value 1.0 then it is considered statistically significant.

Sensitivity analyses will be conducted to define the active case or exposure. For example, an 'active case' could variously be a person with a proven infection +/- symptom onset between -2 and +14 days from the case/control date. Or 0 and 7 days, or various other options. A person being 'in the catchment' could be defined as within 5 km of the waste-water plant, 10 km, 15 km etc. Being in the catchment could refer only to those people whose location was recorded in Victorian PHESS database (or equivalent in other states) or could include people whose residential address is within the catchment even if it is not certain they were there.

A separate Odds Ratio calculation will be done for each combination of potential shedding period and distance from the testing point. Among these the largest Odds Ratios will indicate the optimum definition of 'active case' for surveillance purposes. A subsequent analysis will assess the sensitivity and specificity of a positive sewage sample for the detection of 'active cases' so defined. This will help decide the best way to use sewage testing in COVID-19 surveillance with the initial qualitative testing data (detected/not-detected).

Strengths and limitations of the case control approach

The case control approach is reasonably robust against non-differential misclassification. E.g. false-negative sewage tests, or wrong addresses for infected people. The effect of non-differential misclassification is to draw the OR closer to 1.0 – it may make it difficult to demonstrate an association, but it does not increase the risk of suggesting the association is in the opposite direction.

The result could be that there is no association. For example, it could be that infected people excrete the virus for so long that positive sewage tests do not indicate the presence of active cases. Or it could be that there are seasonal (e.g. temperature) and flow factors that must be considered, e.g. at higher temperatures or with longer transit times, only faecal shedding from infected people very close to the sampling site can be detected. (The latter may restrict the usefulness of testing to samples very close to high-risk targets.)

On the other hand, there may be a suitable set of definitions for case, control, exposure and non-exposure that provides evidence of a strong association. This could be used to quantify the probability that a negative sewage test either rules in or rules out the presence of one or more active cases in the catchment.

Negative sewage tests may mean absences of cases, providing reassurance that the current control measures are working, while the detection of virus in sewage may suggest that at least one active human case in the water catchment, which would then direct intensified case-finding among the population of the same catchment and warrant public health interventions such as social distancing.

Sensitivity and Specificity analysis

If an association can be demonstrated between positive tests in sewage and human cases in the same catchment then it may be possible to use sewage testing as a 'diagnostic test' for the presence or absence of human cases in the catchment. In this case the relevant analysis is the calculation of the sensitivity,

---

[2] It is also possible to use a logistic regression model to calculate Odds Ratios, which includes the possibility of using continuous variables rather than only binary ones, as well as adjustment for potential other contributory factors. This would allow, for example, for exposure to include the number of human cases instead of simple presence/absence. The interpretation of the resulting ORs is very similar to the description above, with the Odds Ratios for continuous variables interpreted as the change in odds for each unit change in the continuous variable.

specificity, positive predictive value and negative predictive value for the presence of virus in sewage. The data needed are very similar to those for the case control study approach, but the calculations are done thus:

	Human case(s) in catchment	No human cases in catchment
Positive sewage test[s]	a	b
Negative sewage test[s]	c	d

Sensitivity = a / a+c

Specificity = d / b+d

Positive Predictive Value = a / a+b

Negative Predictive Value = d / c+d

These can be used directly in a surveillance process to assess, e.g. on the basis of daily sewage testing, the likelihood that there are one or more human cases in the catchment of the relevant waste-water plant. If sewage testing proves highly sensitive, it will be most useful in confidently 'ruling out' the presence of human cases in low-prevalence areas. If it is highly specific it will be good at 'ruling in' the presence of human cases. (The ideal test is both highly sensitive and highly specific, but it does not need to be both in order to be useful in practice.)

The possibility of undetected cases in a catchment, and the high specificity of PCR testing, mean that we should interpret the apparent 'false positives' with caution. These may well be true positives.

There are, again, multiple potential definitions of 'positive sewage test' and 'human case in the catchment', and each combination creates a separate research question, separate sample size calculation and separate analysis.

For study AIM 2

3. Mathematical relationship between amount of virus in sewage and disease prevalence

For study aim 3, a mathematical calculation can be used to accurately estimate the number of active human cases in the catchment of the same treatment plant based on the amount of virus in the sewage sample. In this case a regression model would be fitted along these lines:

$$\text{Human cases} = \beta_0 + \beta_1(\text{Amount of virus in sewage}) + \in$$

where, β_0 - Intercept

β_1 - Slope

$\in$ - Error

Other factors can be included in such models and there are many ways to make the raw data suitable for inclusion (such as taking the log of the original counts etc).

IF such a model can be fitted with sufficient precision, then the amount of virus in sewage samples at the waste-water plant could be used to estimate the number of human cases in the catchment.

Both types of study are feasible but the case control and 'diagnostic test' approach is proposed given this is likely to provide useful results quickly and there may only be availability of qualitative data initially.

Data needed

Initially data will be used as it becomes available (as defined in the separate data request) from already collected ColoSSoS samples in metropolitan and regional Victoria sewage sampling locations. These include weekly samples from most sites and twice weekly automated 24 hour collection samples from the inlet of the two main metropolitan sewage treatment plants.

It is noted that test results will also be made available progressively with most recent (last week samples) being prioritised. Having the complete initial testing data from ColoSSoS testing April – May as soon as possible will be very helpful for timely initial analysis.

If prior samples are available from the 2-3 week period prior to the 1st reported case in Victoria (i.e. from 1st January 2020 and collected/stored for any other purpose) these could be very informative.

Prospectively, after preliminary review of the data and considering surveillance priorities the group will discuss any possible refinement in sampling locations (including high priority institutions/locations) and type and frequency of samples. The testing group is also working to refine and standardize the testing methodology and maximize sensitivity. Quantitative or semi-quantitative results would expand the potential epidemiologic analysis options greatly.

COVID-19 epidemiological surveillance using sewage testing – study proposal

Core data needs (non-exhaustive)

'Case' or 'control' status: regarding each sewage sample tested
Date and time
Location from which sample taken (Latitude and Longitude)
Location where tested
Type of sample
Type of test
Test result
Final assessment ('positive' vs 'negative')
Quantity of virus detected (*not available initially)

Regarding each sampling site
Type of site
Location of site (Latitude and Longitude)
Flow rate at time of sampling
Average dry weather flow/other flow characteristics
The catchment area covered (exact Latitude and Longitude of the boundaries), and any variation during the study period.

'Exposure' status: for each known human case
Date of onset of symptoms
Date considered no longer an active case
Residence address during each day between onset day -1 and onset +28 (Latitude and Longitude)
Home address (if previous unknown) (Latitude and Longitude)
Age
Sex
Severity of disease (asymptomatic, sick but treated at home, admitted to hospital, admitted to ICU, died)
Symptoms (incl. diarrhoea)

For each test on each human case
Date of test
Test performed
Result of test

Collection and Interpretation of data:

Key considerations: for further input from literature review and inter-disciplinary colleagues (water/lab etc)

1. **Virological/Clinical: Only individuals shedding viral fragments/virus will be potentially detectable in sewerage.**

- The proportion of COVID-19 cases with faecal shedding *and variation by disease severity (from asymptomatic through mild, moderate and severe disease and those with or without diarrhoea/GI symptoms).*
- The timing of onset of faecal shedding in relation to infection or symptom onset noting literature suggests wide inter-individual variability but that generally this follows after nasopharyngeal swabs being positive.
- The duration and inter-individual variability of faecal shedding noting literature shows this may persist for 2-3 weeks, including after nasopharyngeal swabs are negative and may also follow a pattern of intermittent shedding.

  The quantitative level and inter-individual variability of faecal shedding when present.
- The frequency and inter-individual variation of an infected individual passing stools and related probability it would then be captured in the sewage sampling.

  (Note there is also some literature on Genito-urinary shedding and upregulation of target receptors ACE-2 but very few reports of positive urinary RT-PCR for SARS-CoV-2).

2. **Individual location vis a vis sewage sampling locations**

- Accuracy of case reporting including location (from PHESS).
- The location and movement of population including known cases (hospital, quarantine, other) and other population (and its relationship to passing stool).

3. **Sewage sampling and catchment are**a

- The sewage sampling point (sewage treatment plant, main line, other) and its geographic catchment area. How well defined is the geographic area drained by the sewers being tested?
- The sewage samples (automated/24 hour, grab, other and frequency) and their relationship to the period of sewage generation by the population in the catchment area – stated simply what is the probability an individual person's stool being included in the sample.
- Can water authorities sample targeted individual priority institutions – feasibility considerations/which – prisons, abattoirs etc

4. Testing of the sewage samples using PCR

- Optimal recovery, testing and/or testing algorithm protocols and properties: What constitutes 'positivity'? Number of specimens positive in a time period? Number of cycles in RT PCR? Confirmatory test/absence of inhibitors/other?
- Sensitivity – including qualitative and at lower limits of detection. (How much virus has to be in the sample in order to reliably detect it?)
- Specificity
- Noting that qualitative testing is proposed initially – can we get to a quantitative test? Limitations / timeframe?
- Are there strategies for more cost effective and/or timely testing? E.g. Pooled testing – use of algorithms with confirmatory tests to improve specificity for SARS-CoV-2 RT-PCR results, use of digital RT-PCR etc
- What is the rate of degradation of viral particles within the wastewater management system and key influencing factors (e.g. temperature, distance/time in system, other) which affects detection (qualitative) or quantitative results?

What is known about the flow rates in the relevant sewers? How far upstream might a positive stool have entered the sewer to be detectable at the given time? How does this vary by time of day? Other factors? E.g. sewage flow etc.

References

[i] Asghar, H., Diop, O.M., Weldegebriel, G., Malik, F., Shetty, S., Bassioni, L. El, Akande, A.O.,Maamoun, E. Al, Zaidi, S., Adeniji, A.J., Burns, C.C., Deshpande, J., Oberste, M.S.,Lowther, S.A., 2014. Environmental surveillance for polioviruses in the global polioeradication initiative. J. Infect. Dis. https://doi.org/10.1093/infdis/jiu384.

[ii] Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T.,Norder, H., 2014. Detection of pathogenic viruses in sewage provided early warnings of hepatitis A virus and norovirus outbreaks. Appl. Environ.Microbiol. https://doi.org/10.1128/AEM.01981-14

[iii] Wu Y, Guo C, Tang L, Hong Z, Zhou J, Dong X, et al. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. The Lancet Gastroenterology & Hepatology. 2020;5(5):434-5

[iv] Jiehao, C., Jing, X., Daojiong, L., Lei, X.,et al. A case series of children with 2019 novel coronavirus infection: clinical and epidemiological features. Clin. Infect. Dis. https://doi.org/10.1093/cid/ciaa198 [Epub ahead of print]..

[v] Tang A, Tong Z, Wang H, Dai Y, Li K, Liu J, et al. Detection of Novel Coronavirus by RT-PCR in Stool Specimen from Asymptomatic Child, China. Emerg Infect Dis. 2020;26(6):1337-1339. https://dx.doi.org/10.3201/eid2606.200301

[vi] Zhang, N., Gong, Y., Meng, F., Bi, Y., Yang, P., Wang, F., 2020b. Virus Shedding Patterns in Nasopharyngeal and Fecal Specimens of COVID-19 Patients. Infectious Diseases (except HIV/AIDS).

[vii] Wölfel, R., Corman, V.M., Guggemos et al. 2020. Virological assessment of hospitalized patients with COVID-2019. Nature, 1–10 https://doi.org/10.1038/s41586-020-2196-x.

[viii] Xu, Y. et al. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. Nat. Med. https://doi.org/10.1038/s41591-020- 0817-4 (2020)

[ix] Xiao F, Tang M, Zheng X. et al. Evidence for Gastrointestinal Infection of SARS-CoV-2. Gastroenterology https://doi.org/10.1053/j.gastro.2020.02.055

[x] Medema G, Heijnin L, Elsinga G, Italiaander R and Brouwer A. Presence of SARS-Coronavirus-2 in sewage. doi: https://doi.org/10.1101/2020.03.29.20045880 (*Preprint – not yet peer reviewed)

Medema G, Heijnen L, Elsinga G, Italiaander R, and Brouwer A. 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 Just Accepted Manuscript. DOI: 10.1021/acs.estlett.0c00357

[xi] Wurtzer S, Marechal V, Mouchel JM et al. Evaluation of lockdown impact on SARS-CoV-2 dynamics through viral genome quantification in Paris wastewaters. (*Preprint – not yet peer reviewed) MedRxiv 2020.04.12.20062679; doi: https://doi.org/10.1101/2020.04.12.20062679

[xii] Fuqing Wu, Amy Xiao, Jianbo Zhang, Xiaoqiong Gu,Wei Lin Lee, Kathryn Kauffman, Hanage, Mariana Matus, Newsha Ghaeli, Noriko Endo, Claire Duvallet, Katya Moniz, Timothy Erickson, Peter Chai, Janelle Thompson, Eric Alm. SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases. doi: https://doi.org/10.1101/2020.04.05.20051540.

[xiii] Ahmed W, Angel N, Edson J, Bibby K, Bivins A, O'Brien JW, et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. Science of The Total Environment. 2020;728:138764

[xiv] Hart OE, Halden RU. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. Science of The Total Environment. 2020;730:138875

[xv] Wu z, Zhang Z and Wu S. Focus on the Crosstalk between COVID-19 and Urogenital Systems. The J. of Urology. https://doi.org/10.1097/JU.0000000000001068.