habitants) with a weekly SARS-CoV-2 average aggregate concentration of 967M GC/L, leading health officials to implement new restrictive measures.

Conclusion: Wastewater surveillance could be used as a complimentary tool to estimate the presence and prevalence of COVID-19 in communities and can be used for preventive purposes, as an increasing SARS-CoV-2 trend in wastewater could be a signal of the possible re-emergence of the pandemic.

https://doi.org/10.1016/j.ijid.2021.12.056

OP04.04 (1066)

Estimating SARS-CoV-2 prevalence from large-scale wastewater surveillance: insights from combined analysis of 44 sites in England

M. Morvan [1,2], A. Lojacomo [2,3], C. Souque [2,4], M. Wade [2], T. Hoffmann [2], K. Pouwels [4], A. Singer [5], J. Bunce [2], A. Engeli [2], J. Grimsley [2], K. O'Reilly [2,6,*], L. Danon [2,3]

[1] *University College London, London, United Kingdom*
[2] *Joint Biosecurity Centre, London, United Kingdom*
[3] *Bristol, Bristol, United Kingdom*
[4] *University of Oxford, Oxford, United Kingdom*
[5] *UK Centre for Ecology & Hydrology, Crowmarsh Gifford, United Kingdom*
[6] *London School of Hygiene & Tropical Medicine, London, United Kingdom*

Purpose: Accurate surveillance of the COVID-19 pandemic can be weakened by under-reporting of cases, particularly due to asymptomatic or pre-symptomatic infections, resulting in bias. Quantification of SARS-CoV-2 RNA in wastewater (WW) can be used to infer infection prevalence, but uncertainty in sensitivity and considerable variability has meant that accurate measurement remains elusive.

Methods & Materials: Data from 44 sewage sites in England, covering 31% of the population, are used in this analysis where samples are available from July 2020 to present day. Samples include the raw SARS-CoV-2 gene copy number and associated metadata. To establish the sensitivity and specificity of the WW data, we compare to population representative prevalence surveys available across England (the ONS Covid Infection Survey - CIS). The WW data were mapped to sub-regional data of the CIS and fitted using mathematical modelling. First, a phenomenological model was developed to model how infected individuals shed SARS-CoV-2 into WW and how the markers may degrade in time and compare this to the data. Second, we develop a model to estimate SARS-CoV-2 prevalence directly from WW data which is trained on the CIS data.

Results: Data from 44 sewage sites in England, shows that SARS-CoV-2 prevalence is estimated to within 1.1% of estimates from representative prevalence surveys (with 95% confidence). Using machine learning and phenomenological models, differences between sampled sites, particularly the WW flow rate, influence prevalence estimation and require careful interpretation. SARS-CoV-2 signals in WW appear 4-5 days earlier in comparison to clinical testing data but are coincident with prevalence surveys suggesting that WW surveillance can be a leading indicator for asymptomatic viral infections.

Conclusion: Wastewater-based epidemiology complements and strengthens traditional surveillance, with significant implications for public health. Using WW to quantify infection prevalence requires knowledge of additional meta-data and outbreak detection needs to account for unexplained aberrations in WW data to improve reliability

https://doi.org/10.1016/j.ijid.2021.12.057

PS04.01 (491)

Sub-genomic RNA Expression in SARS-CoV-2 B.1.411 and B.1.1.7 Infections in Sri Lanka

D. Ranasinghe [1,*], C. Jeewandara [1], D. Jayathilaka [1], D. Ariyarathne [1], T. Jayadas [1], D. Madushanka [1], L. Gomes [1], G. Ogg [2], M. Parker [3], T. De Silva [4], N. Malavige [1,2]

[1] *University of Sri Jayewardenepura, Allergy, Immunology and Cell Biology Unit (AICBU), Nugegoda, Sri Lanka*
[2] *University of Oxford, Oxford, United Kingdom*
[3] *The University of Sheffield, Sheffield Bioinformatics Core, Sheffield, United Kingdom*
[4] *The University of Sheffield, Sheffield, United Kingdom*

Purpose: As experienced by many countries, Sri Lanka is currently experiencing a large COVID-19 outbreak, with over 90 cases/one million population. The previous outbreak, which was due to the B.1.411 virus (Sri Lankan lineage) resulted in a significantly fewer number of cases and deaths compared to the current outbreak caused by B.1.1.7. Therefore, we sought to explore if the differences in the transmission rates and higher mortality rates with the introduction of B.1.1.7 is due to an increased expression of sub-genomic RNA, which is an essential step in the virus life cycle.

Methods & Materials: Sputum or nasopharyngeal samples of 472 patients with SARS-CoV-2 infection were included in the analysis. Samples with the cycle threshold <30, were sequenced using 247 amplicons targeting the SARS-CoV-2 genome (MN908947v3). Library preparation was done using AmpliSeq prep kit and sequenced either on illumina iSeq100 or Nextseq550 platforms. Basecalling and demultiplexing were done using the default bcl2fastq (v2.20) pipeline. Raw index-trimmed fastqs were analyzed for sub-genomic RNA using Periscope (https://github.com/sheffield-bioinformatics-core/periscope). Raw reads were aligned and checked for the leader sequence at the start of each open reading frame (ORF). The sgRNA detected reads were counted, classified into ORFs and normalized using the genomic RNA counts at each position. Groups were compared with an unpaired Wilcoxon test using R\rstatix package. Figures were generated in R\ggpubr.

Results: Out of the remaining 434 datasets after the quality control step, 164 were of B.1.1.7 lineage while 237 were B.1.4.11. Means of the normalized sgRNA counts between B.1.411 and B.1.1.7 viruses were significantly different in six ORFs. Viruses of the B.1.411 lineage expressed significantly higher sgRNA for Spike protein ($p = 0.014$), ORF3a ($p = 0.0001$), Membrane protein ($p = 3.62E-10$), ORF8 ($p = 1.81E-05$), and ORF7a ($p = 0.0004$) than those in B.1.1.7 samples. Contrastingly, Nucleocapsid (N) protein had significantly higher sgRNA expression in B.1.1.7 samples ($p=0.0001$).

Conclusion: Our results suggest that increased expression of sgRNA for a particular virus lineage does not necessarily associate with higher transmissibility as higher expression of sgRNA of B.1.1.7 compared to the B.1.411 lineage virus was only seen for the N protein.

https://doi.org/10.1016/j.ijid.2021.12.058