NYC OpenData

Data Dictionary - Dataset Information

General

Dataset Name	SARS-CoV-2 concentrations measured in NYC Wastewater
Agency Name	New York City Department of Environmental Protection, Bureau of Wastewater Treatment
Update Frequency	monthly
Dataset Description	Results of sampling to determine the SARS-CoV-2 N gene levels in NYC DEP Wastewater Resource Recovery Facility (WRRF) influent, disaggregated by the WRRF where the sample was collected, date sample was collected, and date sample was tested
Dataset Keywords	COVID-19, SARS-CoV-2, wastewater, Wastewater Based Epidemiology
Dataset Category	Health
Can Dataset Feasibly Be Automated?	No
Removed Records?	No
Data published on Agency's Website?	No
Update frequency on Agency's Website	N/A

Detailed Description

General Introduction

In March 2020, New York City became an epicenter of the coronavirus disease 2019 (COVID-19) pandemic. In response to this first wave of COVID-19 cases, the New York City Department of Environmental Protection (NYC DEP) – the city agency responsible for wastewater collection and treatment – launched a wastewater monitoring program with the goal of tracking sewershed-level trends in the concentration of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes COVID-19.

In January 2022, data from the wastewater monitoring program became available at NYC Open Data for public use. In May 2022, the methodology to calculate concentrations in wastewater changed to what scientists call a new "gold standard methodology." *While the absolute values of the SARS-CoV-2 concentrations in the more recent data sets are different from earlier versions (V01-V03, for data through March 29, 2022), the application of this new methodology does not modify trends in the SARS-CoV-2 concentrations.*

NYCDEP works closely with the Department of Health and Mental Hygiene on the wastewater monitoring program and it's important to note that public health decision making is based on trends, not on absolute values. Further details are provided below.

Technical Section:

This dataset contains measurements of the N1 gene abundance for SARS-CoV-2, the virus that causes COVID-19, in the influent of NYC DEP Wastewater Resource Recovery Facilities (WRRFs), August 2020 to present. There is no evidence that the SARS-CoV-2 virus, which causes COVID-19, remains infectious in wastewater. However, its RNA (genetic material), can still be detected and quantified.

The data are produced from laboratory analysis of 24-hour, flow-weighted composite samples of wastewater influent. Samples were collected biweekly.

Each record represents a unique sample date and WRRF, and includes (1) the sampling date (date of collection of the first sample that will be used in the composite, at beginning of the 24 hours sample collection period) (2) the test date for the sample (the date of the first day of the 3-day analysis protocol), (3) the WRRF name and two letter abbreviation, (4) a value indicating the measured SARS-CoV-2 concentration in wastewater ("Concentration SARS-CoV-2 gene target", in N1 gene copies/L), and (5) a value indicating the concentration normalized by the population served by the WRRF and the quantity of wastewater influent to the plant on the sample date (flowrate) ("Per capita SARS-CoV-2 load", in units of N1 gene copies/day/capita), which represents the total number of SARS-CoV-2 viruses per capita in wastewater entering a WRRF over a 24 hour period.

Concentrations of the SARS-CoV-2 N1 gene in wastewater samples (N1 gene copies/liter) were determined with reverse transcription quantitative polymerase chain reaction (RT-qPCR), with the calculation of the concentration of the SARS-CoV-2 RNA based on the comparison with a synthetic RNA standard. Calculation of the concentrations of the SARS-CoV-2 N1 gene in wastewater samples (N1 gene copies/liter) requires: 1- reverse transcription quantitative polymerase chain reaction (RT-qPCR) of the samples; 2- comparison of values obtained in 1 with a synthetic SARS-CoV-2 standard. Therefore, the concentration of the synthetic RNA standard will have an impact on the calculated concentration of SARS-CoV-2 in wastewater. N1 gene concentrations in previous versions of this data set (versions 1-3) were calculated using an approximated concentration of the synthetic SARS-CoV-2 RNA standard based on the manufacturer's specifications. N1 gene concentrations in the current version of this dataset (version 4 and subsequent versions) were recalculated using the exact concentration of the synthetic RNA standard measured using reverse transcription droplet digital PCR (RT-ddPCR). This last approach is considered the "gold standard methodology". When the gold standard methodology was applied, the RT-ddPCR measured concentration of the standard was ~ 25 times lower than the one given by the manufacturer. As a result, the measured SARS-CoV-2 concentrations and per capita SARS-CoV-2 loads reported in the current version of this data set (version 4 and subsequent versions) are also ~25 times lower than those previously reported (versions 1-3). Nonetheless, the temporal trends illustrated by the reported N1 gene concentrations remain the same as those observed in previous versions of this data set.

For further details regarding sample collection, sample processing and analysis, and data analysis, please see the following paper: Hoar, C., Chauvin, F., Katehis, D., Clare, A., McGibbon, H., Castro, E., Patinella, S., Dennehy, J.J., Trujillo, M., Smyth, D., Silverman, A.I. 2022. "Monitoring SARS-CoV-2 in wastewater during New York City's second wave of COVID-19: Sewershed-level trends and relationships to publicly-available clinical testing data." In *Environmental Science: Water Research & Technology.*
(accessible at:) https://doi.org/10.1039/D1EW00747E

Data may be used to track trends in SARS-CoV-2 concentrations in NYC WRRF influent. Dataset does not include COVID-19 case rates. Users are referred to DOHMH site for COVID-19 case rate data, see https://www1.nyc.gov/site/doh/covid/covid-19-data-totals.page.