Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Population level SARS-CoV-2 fecal shedding rates determined via wastewater-based epidemiology

Sarah M. Prasek, Ian L. Pepper, Gabriel K. Innes, Stephanie Slinski, Martha Ruedas, Ana Sanchez, Paul Brierley, Walter Q. Betancourt, Erika R. Stark, Aidan R. Foster, Nick D. Betts-Childress, Bradley W. Schmitz

HIGHLIGHTS
- Wastewater data were utilized to calculate 6-day peak shedding rates.
- Fecal shedding rates were estimated for communities with different demographics.
- Age, ethnicity, and socio-economic factors may have influenced shedding rates.
- Populations with median age 30–39 generally had greater fecal shedding rates.
- In some communities, shedding rates increased with rise of Delta variant infections.

GRAPHICAL ABSTRACT

ABSTRACT

Wastewater-based epidemiology (WBE) has been utilized as an early warning tool to anticipate disease outbreaks, especially during the COVID-19 pandemic. However, COVID-19 disease models built from wastewater-collected data have been limited by the complexities involved in estimating SARS-CoV-2 fecal shedding rates. In this study, wastewater from six municipalities in Arizona and Florida with distinct demographics were monitored for SARS-CoV-2 RNA between September 2020 and December 2021. Virus concentrations with corresponding clinical case counts were utilized to estimate community-wide fecal shedding rates that encompassed all infected individuals. Analyses suggest that average SARS-CoV-2 RNA fecal shedding rates typically occurred within a consistent range (7.53–9.29 log_{10} gc/g-feces); and yet, were unique to each community and influenced by population demographics. Age, ethnicity, and socio-economic factors may have influenced shedding rates. Interestingly, populations with median age between 30 and 39 had the greatest fecal shedding rates. Additionally, rates remained relatively constant throughout the pandemic provided conditions related to vaccination and variants were unchanged. Rates significantly increased in some communities when the Delta variant became predominant. Findings in this study suggest that community-specific shedding rates may be appropriate in model development relating wastewater virus concentrations to clinical case counts.

Keywords: Fecal shedding, SARS-CoV-2, Community demographics, COVID-19, Wastewater-based epidemiology

1. Introduction

Wastewater-based epidemiology (WBE) has demonstrated utility during the COVID-19 pandemic to trigger public health responses, such as clinical testing interventions (Betancourt et al., 2021), that aim to stifle disease transmission. Wastewater surveillance relies on SARS-CoV-2 infected...
individuals contributing virus and/or viral RNA into sewage via multiple sources, including feces (Cevik et al., 2021), urine (Brönimann et al., 2020; Kashi et al., 2020), saliva (Wyllie et al., 2020), and sputum (Khiabani and Amirzade-Iranaq, 2021; Li et al., 2022). A recent report suggests sputum may be an important shedding source (Li et al., 2022); however, feces dominate population-level SARS-CoV-2 RNA loading in wastewater (Crank et al., 2022). Knowledge of virus levels in sewage and their contributing sources can subsequently be utilized as an indirect measure to estimate population-level COVID-19 cases (Medema et al., 2020).

Attempts to develop models that correlate wastewater virus concentrations with reported numbers of clinical cases have been elusive for several reasons. One issue has been the underreporting of cases due to large numbers of asymptomatic or mildly symptomatic infections (Petala et al., 2022), as well as at-home diagnostic tests, both of which may result in failure to self-report. There is also limited knowledge regarding how wastewater data align with the timeline of SARS-CoV-2 infection, including the incubation period, the initiation of viral shedding in feces, the duration of shedding and the onset of symptoms. Previous reports have suggested that shedding can last 3–4 weeks at variable rates after symptom-onset (Chen et al., 2020a; Chen et al., 2020b; Wu et al., 2020; Yan et al., 2021; Zheng et al., 2020). Despite limitations, it may be possible to estimate the total number of infections (both symptomatic and asymptomatic) in the community through analysis of SARS-CoV-2 excreted into wastewater. Previous studies have based such estimations on the amount of fecal material excreted per person, the viral shedding rate in feces, and the wastewater flow rate (Ahmed et al., 2020a; Chavarría-Miré et al., 2021; Curtis et al., 2020). Among these parameters, the SARS-CoV-2 shedding rate in feces is the most difficult to define (Cevik et al., 2021).

In this study, community-level fecal shedding rates of SARS-CoV-2 RNA were calculated for six distinct populations that differed in size and demographics. Community-specific fecal shedding rates can then be used to develop estimations for the total number of infected individuals. Ultimately, estimated disease prevalence (EDP) can be used to inform public health officials and key stakeholders (e.g. hospitals, city administration, businesses, schools) to effectively utilize resources for preparedness and response actions that lessen outbreak severity.

2. Methods

2.1. Community characteristics and demographics

Wastewater was monitored in six communities between September 2020 and December 2021 (Table S1). Communities A, B, E, and F were located in Arizona, while Communities C and D were in Florida, USA. Blueprints and maps of the sewer system piping networks were provided by wastewater utilities, and samples were collected from sites that were selected based on wastewater facility or manhole locations corresponding to defined areas and populations. Buildings with unique populations (e.g. hospitals, elderly care facilities, schools) did not have separate treatment systems prior to discharging into the sewage distribution. Wastewater characteristics (i.e. flow rates and service area boundaries) for each community were provided by municipal and/or wastewater utility staff (data not included). Population demographics (Tables 1 and S2) were obtained from the United States Census Bureau (USCB) QuickFacts and Census Reporter (U.S. Census Bureau, 2019a,b). Population counts were adjusted based upon the number of residents living in the Water Reclamation Facility (WRF) service area (Table 2).

2.2. Wastewater sampling and analysis

One liter grab samples were collected from raw sewage influent at WRFs, during peak flows between 7:00 am and 10:00 am for Communities A, B, D, and F. Autocomposite samples were collected over 24-hour periods at Community C (Teledyne ISCO 3700C portable and 5800 refrigerated; Thousand Oaks, CA) and Community E (Hach Sigma 900MAX, Loveland, CO). Wastewater facility staff collected samples twice weekly (i.e. Monday and Thursday) at all communities, except for Community E which typically collected five to seven days apart. In Communities B, C, and F wastewater samples were collected at two or more facilities, each of which served different portions of the community (Table 2). Samples were collected in sterile Nalgene bottles and shipped on ice to laboratories for immediate processing.

Wastewater samples from Communities A, B, F were processed at the University of Arizona’s (UArizona) Yuma Center of Excellence for Desert Agriculture (YCEDA). Samples from Community E were processed at the UArizona Water & Energy Sustainable Technology Center (WEST). Wastewater samples were analyzed using real-time quantitative reverse transcriptase PCR (RT-qPCR) within 4 to 6 h of sampling according to protocols previously described (Betancourt et al., 2021). Samples from Communities C and D were shipped overnight and processed within 24 h by GT Molecular (Fort Collins, CO, USA). Wastewater from Community C was analyzed using digital droplet PCR technology (ddPCR), while samples from Community D were analyzed via RT-qPCR, in adherence with GT Molecular’s protocols. Refer to Section 2.7 for instrument types. Standard curves and detection limits are listed in Table S11.

Samples were tested for SARS-CoV-2 RNA using Centers for Disease Control and Prevention (CDC) RT-PCR assays that targeted regions of the nCoV nucleocapsid gene (N1; Table S3) (Research use only kit, Integrated DNA Technologies, Coralville, IA). Detection of N2 gene was not considered due to low assay efficiency (0.65) (data not shown).

Detected SARS-CoV-2 RNA concentrations were adjusted based on the average recovery efficiency specific to each site/facility at Communities A, B, E, and F (Table S4). Average recovery efficiencies for YCEDA and WEST ranged from 4.61 % to 31.60 %. For Communities C and D, GT Molecular adjusted concentrations based on the recovery for individual samples prior to sharing results. Therefore, mean recovery efficiency data for associated facilities (WRF-C1, WRF-C2, and WRF-D) are not provided.

2.3. Study stages and duration

The study period initiated and concluded at different dates for each community; however, testing overlapped for several months across all communities and data was categorized by stage of the pandemic (Table S1). Specifically, fecal shedding rates were calculated for four time periods:

| Table 1 Community demographics. |
|----------------------------------|
| Community | Persons per household | Median age | >65 years | 18 years | Female | Hispanic or Latino | Poverty | Without health insurance | High school graduate |
|-----------|-----------------------|------------|----------|---------|--------|-------------------|---------|-------------------------|---------------------|
| A         | 3.51                  | 27.5       | 7.4 %    | 33.5 %  | 51.4 % | 96.1 %           | 25.8 %  | 19.4 %                  | 63.0 %              |
| B         | 3.79                  | 30.0       | 7.6 %    | 37.4 %  | 46.0 % | 97.0 %           | 24.2 %  | 21.1 %                  | 49.0 %              |
| C         | 2.85                  | 35.6       | 12.3 %   | 22.0 %  | 50.9 % | 32.7 %           | 12.5 %  | 15.4 %                  | 86.5 %              |
| D         | 2.24                  | 36.4       | 14.4 %   | 18.6 %  | 52.4 % | 28.2 %           | 12.5 %  | 15.5 %                  | 95.1 %              |
| E         | 2.46                  | 38.9       | 20.3 %   | 20.6 %  | 50.8 % | 37.8 %           | 14.0 %  | 13.3 %                  | 88.4 %              |
| F         | 2.19                  | 60.4       | 44.5 %   | 15.6 %  | 48.1 % | 27.3 %           | 11.3 %  | 8.2 %                   | 87.0 %              |

More detailed demographic information is provided in Table S2.

a Information from United States Census Bureau QuickFacts (U.S. Census Bureau, 2019a).
b Information from Census Reporter, 2019 American Community Survey 5-year estimates (U.S. Census Bureau, 2019b).
c Percent population without health insurance is reported for persons under age 65.
Early Pandemic (September–December 2020); Holiday Surge (December 2020–January 2021); Spring Downturn (February–May 2021); and Delta Predominant (mid-July–October/December 2021).

Samples were collected across all four stages of the pandemic for Communities D and E (Table S1). For Communities A, B, C and F samples were only collected from Spring Downturn through Delta Predominant (Table S1). As the names indicate, the Holiday Surge stage corresponded to the 2020–2021 holiday season when COVID-19 cases counts increased; whereas, the Spring Downturn stage aligned with a decrease in reported cases. The Delta Predominant stage was defined as the period in which the Delta variant was responsible for ≥90 % clinical cases in the appropriate region per the CDC Data Tracker (CDC COVID Data Tracker, 2021). However, the exact proportions of SARS-CoV-2 variants in wastewater samples was unknown as whole genome sequencing was not performed.

Additionally, vaccination varied by pandemic stage: vaccination was not yet available during Early Pandemic; was offered to primarily essential workers during Holiday Surge; was accessible to high risk and other groups during Spring Downturn, and was available community-wide by the Delta Predominant stage.

Samples were collected across all four stages of the pandemic for Communities D and E (Table S1). For Communities A, B, C and F samples were only collected from Spring Downturn through Delta Predominant (Table S1).

2.4. Clinical testing data

Clinical data were provided by the Regional Center for Border Health, Inc. (Communities A and B), Local and State Public Health Departments (Communities A, B, D and F), and by USA Facts (Communities C, D, and E) (US COVID-19 Cases and Deaths by State | USAFacts, 2021). For Community D, the source of clinical data varied according to its availability: zip code-level data were obtained from Public Health Departments during Early Pandemic through Spring Downturn, but county-level data were obtained from USA Facts for Delta Predominant. Clinical data were adjusted to enumerate the number of cases that could contribute to the wastewater within the WRF service area. Fig. S1 provides an example for determining overlapping WRF service areas within zip code boundaries.

2.5. Alignment of clinical data with wastewater

To estimate the number of SARS-CoV-2 infected individuals who measurably contributed to a given positive wastewater sample, a six-day sum of clinical data was employed: the day before, day-of, and four days after wastewater sampling (refer to Section 4.1). Reported clinical case numbers within the six-day sum were then adjusted by the CDC estimations for COVID-19 disease burden (reported × 4.3) to approximate the total number of reported and unreported cases (Estimated COVID-19 Burden | CDC, 2021).

2.6. Viral shedding rate estimation

The fecal shedding rate for each community was calculated based on aligned wastewater and clinical data. A modified equation from previous studies that estimated the total number of infected individuals (Ahmed et al., 2020a; Chavarria-Miró et al., 2021; Curtis et al., 2020), was used to calculate community-specific SARS-CoV-2 RNA fecal shedding rates:

\[
FS = \frac{(VC \times Q \times f \times h)}{G \times I}
\]

where, VC is the virus RNA concentration (genome copies/L), Q is the flow rate (GPM), f is the conversion factor between gallons and liters, h is the conversion factor between minutes and days, G is the average mass of stool produced per person per day (Curtis et al., 2020; Rose et al., 2015), and I is total number of infected individuals contributing to the wastewater sample (see Section 2.5). This equation was previously used to estimate fecal shedding rates from infected individuals in defined dormitory communities (Schmitz et al., 2021).

Samples were removed from data analysis when SARS-CoV-2 RNA was not detected in sewage samples and/or there were zero reported clinical cases. Additionally, decay rate was not incorporated into calculations as sample processing took place within 24 h of collection for all locations (Section 2.2). Previous studies report that SARS-CoV-2 RNA persists with minimal to no decay observed within the first three to five days (Ahmed et al., 2020b; Wang et al., 2021).

2.7. Cross-lab control experiments

Cross-lab round-robin experiments were performed to ensure that data could be compared across multiple laboratories. In total, six samples (1 L) were collected from Communities A and B. Aliquots (250 mL) from each sample were processed within 24 h at YCEDA and WEST using the same protocols for virus recovery, nucleic acid extraction, and RT-qPCR, as previously described (Betancourt et al., 2021).

Another cross-lab control experiment was performed to ensure consistency in results across different molecular assays/instruments used at different laboratories. In total, three samples (1 L) were collected from Community B and aliquots (250 mL) were sent to each laboratory (YCEDA, WEST, GT Molecular). Samples from YCEDA and WEST were processed using the same virus concentration and nucleic acid techniques (Betancourt et al., 2021), while samples were processed at GT Molecular according to the company's internal protocols. Each lab utilized a different molecular assay/instrument for quantifying SARS-CoV-2 RNA. RT-qPCR was performed on a CFX Connect Real-Time PCR Detection System (BioRad, Hercules, CA, USA) at YCEDA and a LightCycler 480 Instrument II (Roche Diagnostics, Basel, Switzerland) at the WEST Center, while ddPCR was performed on a QX200 Droplet Generator and QX200 Droplet Reader (BioRad) at GT Molecular.

2.8. Sensitivity analyses

Several sensitivity analyses were performed to investigate how changes in fecal shedding rate parameter values (Eq. (1)) may have impacted results. The first analysis was conducted to determine if fecal shedding rates were influenced by the type of molecular assay or collection method. Shedding rates were calculated using virus RNA concentrations from RT-qPCR and ddPCR on the same samples from Community C between 6/17/21–8/12/21 (n = 11). Also, fecal shedding rates were compared across communities with different sample types (grab vs composite).
The second sensitivity analysis was performed to evaluate how total case count estimations from various day sums (1, 3, 6, and 17-day sums) may have impacted shedding rates (refer to Sections 2.5 and 4.1). A sensitivity analysis was also performed to determine the influence of the CDC multiplier (refer to Section 2.5) on calculated shedding rates. Fecal shedding rates were estimated for Community D throughout all four stages using several multipliers (4.0, 4.1, 4.2, 4.3) for determining the total number of infected persons.

The last sensitivity analysis was performed to determine reliability of adjusting county-wide clinical data when more discreet data (i.e. zip codes) was not available. To do so, clinical data reported at the county and zip-code levels were adjusted for overlapping WRF service area population and compared (Fig. 1).

2.9. Statistical analyses

Kruskal-Wallis H test followed by pairwise Wilcoxon rank sum were conducted to determine whether fecal shedding rates were significantly different across communities, as well as different stages of the pandemic. To account for multiple comparisons, alpha was corrected to either 0.005 or 0.0025 dependent upon the specific analysis (see supplemental for more specifics). Fecal shedding calculations were performed with Microsoft Excel (version 16.47.1, 2021) and statistical comparisons were performed with R studio (R Studio Team, 2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. (http://www.rstudio.com/).

3. Results

3.1. Community characteristics and demographics

Participating communities varied in size from 18,000 to 1,047,279 residents, with monitored service area populations ranging from 4508 to 454,434 (Table 1). Populations in Communities A, B, and F were all <40,000 persons, while Community C was approximately 87,000 persons, Community D was approximately 870,000, and Community E was approximately 1,000,000 (Table 2).

In all communities, the ratio of male to female was split nearly evenly with reported percent female ranging from 46.0 % to 52.4 % (Table 1). With respect to other key demographics factors – ethnicity, median age, and percent poverty – Communities A and B were frequently similar to each other, but different from Communities C, D, E, and F (Tables 1 and S2). Communities A and B are located <25 km apart and are close to the Mexico-Arizona border (<6.5 and <25 km, respectively). Both of these communities are geographically proximate to Community F (<60 and <42 km), despite having much different demographics than Community F.

Populations in Communities A and B were younger (median age 27.5 and 30) than other communities, had lower educational attainment, and were primarily Hispanic or Latino – 96.1 % and 97.0 % respectively (Tables 1 and S2). Communities A and B also had higher poverty rates (25.8 % and 24.2 %), with median household incomes under $40,000 (Table S2). By contrast, Communities C, D, and E had higher median ages (35.6, 36.4, and 38.9 years) and fewer individuals identifying as Hispanic or Latino (ranging between 28.2 % and 37.8 %; Table 1). Meanwhile, Community F had the oldest median age (60.4 years) and the smallest Hispanic or Latino proportion (27.3 %). Communities C, D, E, and F were also more educated and more wealthy than Communities A and B. The median annual household incomes from Communities C–F ranged between $50,749 and $58,254 (Tables 1 and S2), while Community A and B were much lower at $38,315 and $37,255.

3.2. Fecal shedding rates across communities

Mean shedding rates across all communities typically occurred within a consistent range across pandemic periods, between 7.53 and 9.29 log_{10} gc/g-feces (Table 3). However, some variability was observed. In the Spring Downturn period, fecal shedding rates in Community E were greater than all other communities by 0.15–0.90 log_{10} gc/g-feces (Table 3). However, this difference was only significant when compared to Communities A, B, and F, but not C or D (Table S5). Geographically proximate Communities A, B, and F did not have significantly different fecal shedding rates during this stage. Communities C and D, both located in Florida, also did not display significantly different shedding rates from each other. Shedding rates from both C and D were greater than A, B, and F; these differences were significant except for in the case of C and A (Table S5).

During the Delta Predominant stage, fecal shedding rates in Community E were again significantly greater by 0.64–1.26 log_{10} gc/g-feces than all
other communities (Tables 3 and S6). There was no significant difference between the shedding rates in all other communities (Fig. 2 and Table S6). Lastly, no significant difference in shedding rates was observed between Communities D and E during the Early Pandemic and Holiday Surge stages (Table S7).

3.3. Fecal shedding rates across stages of the COVID-19 pandemic

Fecal shedding rates were calculated throughout several stages of the COVID-19 pandemic for each community. Mean shedding rates within the Spring Downturn stage ranged from 7.53 to 8.43 \( \log_{10} \text{gc/g-feces} \), then increased in the Delta Predominant stage to 8.03–9.29 \( \log_{10} \text{gc/g-feces} \) (Table 3). This observed increase was statistically significant in Communities A, B, and F, but not significant in Communities C, D, and E (Table S8). In Community D, shedding rates during the Holiday Surge stage were significantly greater than every other stage, albeit the difference in mean shedding rates was only 0.14–0.24 \( \log_{10} \text{gc/g-feces} \) (Tables 3 and S8). In Community E, median fecal shedding rates were not significantly different across the earlier three stages of the pandemic (Early Pandemic, Holiday Surge, Spring Downturn) prior to high levels of vaccination. Upon the Delta variant becoming predominant, mean shedding rates in this community increased by 0.64–0.88 \( \log_{10} \text{gc/g-feces} \) (Tables 3); however the Delta Predominant rates were only significantly greater than the Early Pandemic stage (Table S8).

3.4. Control experiments

Cross-lab experiments were performed on the same sample sets to ensure that data and results processed from multiple laboratories could be accurately compared. The first cross-lab experiment indicated no significant difference in results when each lab used the same methods for processing samples (Table S9). This corroborates with a recent interlaboratory comparison that showed high reproducibility of results using 36 different methods (Pecson et al., 2021). However, SARS-CoV-2 RNA concentrations were significantly influenced depending on the standard/control utilized to perform molecular assays (Table S10). Thus, it was assumed that results were not influenced by different labs handling and processing of samples. Instead, any differences in concentrations were due to differences in standards used for molecular assays.

Results from WEST were significantly greater than those from YCEDA and GT Molecular, by approximately 1.42 \( \log_{10} \text{ copies/L} \) (Table S10), possibly due to differences in recovery efficiencies (Table S4) and/or standard curves created using different standards (Table S11). Meanwhile, there was no significant difference between results from YCEDA and GT Molecular, despite using different standards (Table S10). Therefore, SARS-CoV-2 concentrations from WEST were decreased by 1.42 \( \log_{10} \text{ copies/L} \) in order to make direct comparisons with results from the other labs.

3.5. Sensitivity analysis for viral fecal shedding estimation

Several sensitivity analyses were performed to investigate how fecal shedding rate parameter alterations (Eq. (1)) may influence results. The first analysis showed that utilizing different molecular assays (RT-qPCR vs ddPCR) to quantify SARS-CoV-2 RNA in wastewater samples did not result in significant changes in the computed fecal shedding rates (Table S12). Also, sample type (grab vs composite) did not influence fecal shedding rates, as results were consistent across communities regardless of collection method (Table 3). This contradicts a previous report and pre-print that suggest composite samples may be more appropriate for wastewater-based epidemiology purposes (Ahmed et al., 2021; Curtis et al., 2020).

Analysis regarding the influence of clinical cases suggested that altering the number of days (1, 3, 6 and 17-day sums) to determine the total infected individuals (\( I \)) does have a significant influence on calculated fecal shedding rates - as the shedding duration was expanded, the shedding rate was significantly decreased (Table S13). The high shedding rates reported here are due to the fact that case counts are based on peak shedding (6-day sum). When longer intervals are considered, more cases accumulate and the shedding rate decreases accordingly (Table S13). Utilizing different multipliers (4.0, 4.1, 4.2, 4.3) suggested by the CDC to adjust for the actual disease burden did not result in significant changes in computed shedding rates (Table S14). Also, no significant difference in the total number of clinical cases was observed when county-wide and zip-code level data were adjusted based on WRF service areas (Fig. 1). Fecal shedding rates were also calculated based on each WRF service area (Table S15), to provide further detail beyond a combined rate for communities with multiple WRFs (Table 3).

4. Discussion

4.1. Fecal shedding rate estimation

A critical factor for calculating fecal shedding rates is the selection of the time interval used to identify the number of cases that resulted in a given virus concentration in the wastewater. In this study, a six-day case summation was utilized, incorporating the day before the wastewater sample was collected, the day of sampling, and four days after the day of sampling. This range was considered optimal given that previous research has documented that: maximum shedding occurs just prior to symptom development, approximately six to eight days following infection (D’Aoust et al., 2021; Cavany et al., 2021; Guan et al., 2020; Lauer et al., 2020; Li et al., 2020a; Peccia et al., 2020; Petala et al., 2022), and an exponential decline occurs in shedding following the onset of symptoms (Chen et al., 2020a; Wu et al., 2020; Zheng et al., 2020). Given that maximum shedding coincides with symptom development, individuals that likely contribute most substantially to viral shedding concentrations are those that recently tested and reported as new cases of COVID-19 during this six-day time period. Individuals who reported illness prior to and after the six-day range were assumed to contribute less substantial SARS-CoV-2 RNA shedding (Chen et al., 2020a; Wu et al., 2020; Zheng et al., 2020). The rationale for this range is similar to, and consistent with, a previous report that used a 6-day range from early-stage infections to estimate viral shedding (Schmitz et al., 2021).

While clinical data may be used as a proxy to determine the number of reported individuals likely to have contributed viral RNA to a wastewater sample (i.e. six-day sum), the largest limitation to clinical case data is the inherent information bias, where individuals may self-select to receive testing and report results. Wastewater measurements, however, incorporate RNA from all infections and does not rely upon self-reporting or self-selection for testing. Therefore, to appropriately align clinical and wastewater data in the study, the clinical data was adjusted to account for the actual disease burden in the community via CDC’s approximation that only 1 in

Table 3

| Community | Stage            | Avg  | Stdev | Med  | Min  | Max  | n   |
|-----------|------------------|------|-------|------|------|------|-----|
| A         | Spring Downturn  | 7.60 | 0.56  | 7.49 | 6.27 | 8.13 | 9   |
| B         | Spring Downturn  | 8.57 | 0.56  | 8.16 | 7.03 | 9.34 | 16  |
| C         | Spring Downturn  | 8.12 | 0.60  | 7.80 | 6.34 | 8.88 | 34  |
| D         | Early Pandemic   | 8.04 | 0.35  | 7.79 | 7.40 | 8.57 | 12  |
| E         | Early Pandemic   | 8.03 | 0.17  | 7.98 | 7.68 | 8.34 | 20  |
| F         | Holiday Surge    | 8.26 | 0.30  | 8.24 | 7.60 | 8.68 | 30  |
| G         | Holiday Surge    | 8.42 | 0.12  | 8.38 | 8.21 | 8.68 | 16  |
| H         | Spring Downturn  | 8.28 | 0.21  | 8.23 | 7.84 | 8.75 | 32  |
| I         | Delta Predominant| 8.18 | 0.21  | 8.09 | 7.89 | 8.62 | 12  |
| J         | Delta Predominant| 8.41 | 0.94  | 7.90 | 6.19 | 9.07 | 13  |
| K         | Holiday Surge    | 8.65 | 0.44  | 8.56 | 7.79 | 9.07 | 8   |
| L         | Spring Downturn  | 8.43 | 0.28  | 8.42 | 7.79 | 8.80 | 12  |
| M         | Delta Predominant| 9.29 | 0.51  | 9.35 | 8.18 | 9.68 | 11  |
| N         | Spring Downturn  | 8.19 | 1.07  | 7.32 | 5.40 | 9.44 | 45  |
| O         | Delta Predominant| 8.65 | 0.76  | 8.17 | 5.48 | 9.79 | 97  |

All values represent SARS-CoV-2 fecal shedding rates (log_{10} gc/g-feces), except for n. Avg, average; Stdev, standard deviation; Med, median; Min, minimum; Max, maximum; n, number of samples.
4.3 COVID-19 cases are reported (Estimated COVID-19 Burden | CDC, 2021). This ratio of underreported clinical case was supported in other studies’ findings (Petala et al., 2022; Schmitz et al., 2021).

Case count adjustments may provide a more accurate estimate to calculate fecal shedding rates; however, limitations exist. First, the CDC estimate for unreported cases was not updated according to specific stages of the pandemic throughout the study period. Thus, the use of the CDC multiplier to adjust case numbers was most appropriate prior to the Delta Pandemic stage when the multiplier remained constant. Second, the proportion between unreported cases and reported cases fluctuates, being lower when disease prevalence is minimal and increasing as the number of reported cases rises (Petala et al., 2022). Overall, sensitivity analysis demonstrated that a difference in selected multipliers (4.0–4.3) did not result in a significant difference in the fecal shedding rates (Table S13).

4.2 Influence of demographic factors on fecal shedding rates

Despite distinct demographic differences among communities, fecal shedding rates occurred within a consistent range (7.53–9.29 log₁₀ gc/g-feces) (Table 3). However, some variability was observed across communities, suggesting possible demographic or geographical influence.

Interestingly, Communities A and B had the lowest fecal shedding rates among all communities and were comprised of the youngest populations. These communities consisted of fewer people over 65 years of age and more people younger than 18 years than other communities (Table S2). Mean fecal shedding rates were typically greater as the median age of populations increased. Aging populations are more likely to exhibit severe COVID-19 symptoms (Cesari and Montero-Odasso, 2020; Chen et al., 2021). This ratio of undereported clinical case was supported in other studies’ findings. This ratio of undereported clinical case was supported in other studies’ findings (Petala et al., 2022; Schmitz et al., 2021).

Socioeconomic factors have also been documented as influencing the incidence of COVID-19, with elevated disease occurrence in low-income communities (McCormack et al., 2020). Once again, Communities A and B differed from C, D, E, and F by having lower median incomes, higher percent of households in poverty without health insurance, and lower educational attainment (Table S2). Importantly, while communities varied with respect to many demographic factors – with some increasing the risk of transmitting and contracting COVID-19 – the factor of age was found to be the most consistent in determining a community’s unique SARS-CoV-2 RNA fecal shedding rate.

There were several limitations of this analysis. This study did not incorporate linear regression analyses which may have better elucidated the magnitude that various factors influenced shedding rates. Rather, the objective of this study was to calculate fecal shedding rates and provide general observations linked to community demographics, geographical location, and stage of pandemic. Additionally, the specific demographics of each infected individual that contributed shedding into wastewater was not known. Thus, fecal shedding rates determined in this study should only be utilized for community-level analyses and not interpreted as representative at an individual level. Also, this study investigated only some of the factors that may have influenced fecal shedding rates; there are unmeasured confounding factors that likely were not considered. Finally, this study did not investigate disparities in wastewater matrices across utilities; however, sewage composition may have significantly influenced results. Future studies should focus more closely on what factors influence community-wide fecal shedding rates, and the extent of their impact.

Fig. 2. Fecal shedding rates across communities. Fecal shedding rates (gc/g-feces) for each community during the Spring Downton (A) and Delta Predominant (B) stages of the pandemic. The boxes represent 50% of the data. The horizontal line inside the box represents the mean. Whiskers represent the minimum to the lower quartile and the upper quartile to the maximum. Black dots represent outliers (>3/2 times of upper quartile or <3/2 times of lower quartile).
4.3. Fecal shedding rates at different stages of the pandemic

Fecal shedding rates within populations were constant while pandemic conditions remained relatively stable. For most communities, there were no significant differences in shedding rates calculated across stages of the pandemic prior to high levels of vaccination and the predominance of the Delta variant. Factors impacting shedding rates during the Delta Predominant stage were many. In particular, vaccination was widespread and available to all adults regardless of age by the beginning of this stage in mid-July. Vaccination rates around the country varied, however, as did the ages of people choosing to vaccinate. Shedding dynamics also changed as a result of the Delta variant, and populations themselves did not necessarily remain steady in each study community due to potential movement of older winter visitors, migrant workers, and younger student populations. This divergent expression of change further illustrates the population specificity of shedding rates.

4.4. Influence of the Delta variant

The influence of the Delta variant appeared to significantly increase fecal shedding rates in some communities located in Arizona (A, B, and F), but did not significantly change rates in Community E or in communities located in Florida (C and D). The increase is likely due to the higher viral loads and longer duration of shedding associated with the Delta variant (Liu and Rocklöv, 2021). The reproductive number R0 represents the average number of new infections generated by one infectious person. If R0 > 1 an epidemic will grow, whereas if R0 < 1 the epidemic will decline. The Delta variant R0 reported range is 3.2 to 8 with a mean of 5.8 (Liu and Rocklöv, 2021). This contrasts with the ancestral SARS-CoV-2 median R0 of 2.79 (Liu and Rocklöv, 2021). The higher R0 values for the Delta variant are indicative of a higher viral load that potentially increased shedding rates for some communities. Specific shedding rates associated with the Delta variant have not been reported to date.

The reason for the lack of an increase in shedding rates for three communities (C, D, and E) is not clear but could be due to geographical location, demographic factors, and/or sample processing. For example, the departure of older people (seasonal migrants) at the same time as the arrival of younger people visiting reopening national theme parks and universities might have moderated an anticipated rise in shedding rates from Delta.

4.5. Shedding rates at estimated disease prevalence

The concentration of SARS-CoV-2 within wastewater can be used to determine the actual disease burden within a community by estimating the total number of infections (including both reported and unreported cases) within the community. This is also known as the EDP. Prior to this study, the major unknown limiting the use of wastewater measurements to generate EDP was an appropriate fecal shedding rate. This study provides a framework for calculating a mean fecal shedding rate for a specific community, which can then be applied to calculate EDP within the population using wastewater concentrations of SARS-CoV-2 RNA. The EDP can be obtained the same day as the wastewater sampling, enhancing the leading indicator status of WBE. It also accounts for the total number of infections, rather than just reported cases, further assisting public health decision makers. In preliminary work, currently being prepared for publication, it has been feasible to use the shedding rates presented in this study to estimate disease prevalence in communities.

4.6. Concluding remarks

One of the key advantages of WBE is that it is a leading indicator of COVID-19 disease and can help predict when an outbreak will occur several days prior to the outbreak. The ability to accurately estimate disease burden within a community from wastewater samples in near real-time can provide public health officials with valuable information for preparedness and response actions. However, current calculations and models that leverage wastewater data to estimate the number of infected individuals within a community are limited due to the lack of understanding of SARS-CoV-2 fecal shedding rates. The data from this study demonstrate that fecal shedding rates occurred within a consistent range, but differed in communities with different demographic and geographical factors. This study also highlights that fecal shedding rates can change during different stages of the pandemic, as shown by the emergence of the Delta variant.

These findings suggest that community-specific fecal shedding rates may be necessary for use in disease prevalence models. This study provides a framework to assist public health officials and researchers align clinical data and wastewater measurements to determine specific fecal shedding rates as a critical step in developing EDP models.

CRediT authorship contribution statement

Sarah M. Prasek: Formal Analysis, Project Administration, Visualization, Investigation, Writing-original draft.
Ian L. Pepper: Project Administration, Visualization, Investigation, Writing-original draft.
Gabriel K. Innes: Visualization, Investigation, Resource, Writing – reviewing & editing.
Stephanie Slinski: Methodology, Resource.
Martina Ruedas: Resource.
Ana Sanchez: Resource.
Paul Brierley: Funding acquisition, Resource.
Walter Q. Betancourt: Methodology, Resource.
Erika R. Stark: Resource.
Aidan R. Foster: Resource.
Nick D. Betts-Childress: Resource.
Bradley W. Schmitz: Formal Analysis, Project Administration, Visualization, Supervision, Funding acquisition, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the six communities in Arizona and Florida for their participation and contributions; the Arizona Department of Health Services (ADHS) and Regional Center for Border Health, Inc. for providing data and public health support; and GT Molecular for assisting with sample processing and analysis. We also thank Ed J. Torres; David K. Ammerman, Franklin W. Martz, and Dr. Jeff Prevatt for providing municipal information. Financial support for the study was provided by the Arizona Department of Health Services (CTR053299). The IRB at the University of Arizona reviewed the study and verified that all data was de-identified and complied with the Human Subjects Protection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.156535.

References

Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tscharke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholz, P., Thomas, K.V., Mueller, J.F., 2020a. 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. Sci. Total Environ. 728, 138764. https://doi.org/10.1016/j.scitotenv.2020.138764.
Ahmed, W., Berish, P.M., Bibby, K., Harzamoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajic, A., Riddel, S., Sherrchan, S.P., Simpson, S.L., Sinikhanchana, K., Symonds, E.M., Verhagen, R., Vasan, S.S., Kitajima, M., Bivins, A., 2020b. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. Environ. Res. 191, 110992. https://doi.org/10.1016/j.envres.2020.110992.
