

Wastewater Metagenomics: A Public-Health Early-Warning System Hidden in the Sewer

Sewage can reveal pathogens, variants, and resistance genes across whole communities, but sampling design and interpretation determine whether the signal is actionable.

Rajnandini Vasantrao Nevhal¹, Khandare Karuna Vinayak¹, Karan Vaijanath Waghmare¹

¹ Department of Biotechnology and Bioinformatics, Mahatma Gandhi Mission's College of Computer Science and Information Technology, Nanded, Maharashtra, India

* Corresponding author: nevhralrajnandini77@gmail.com

PUBLICATION	COLLECTION	REVIEW	LICENSE
22 July 2026	BioFrontiers Launch Collection 2026	Editorially reviewed	CC BY 4.0

ABSTRACT

Wastewater contains biological material shed by thousands or millions of people. By sampling sewage, public-health teams can monitor pathogens and antimicrobial-resistance genes without testing every individual. The COVID-19 pandemic demonstrated that viral RNA trends in wastewater can track community transmission and sometimes reveal variants before clinical reporting catches up. Metagenomic methods extend the concept beyond one virus by sequencing broad mixtures of genetic material. They can support surveillance of enteric pathogens, respiratory viruses, antimicrobial resistance, and emerging biological threats. Yet wastewater data are not direct case counts. Rainfall, industrial discharge, sewer connectivity, travel, sampling time, laboratory recovery, and population behaviour affect the signal. In many Indian settings, incomplete sewer networks and open drains complicate representativeness but also create opportunities for sentinel monitoring. This Article argues that wastewater metagenomics should be built as a calibrated public-health observatory. Its value lies in trends, anomaly detection, and triangulation with clinical and environmental data, not in treating every sequence as evidence of an outbreak.

KEYWORDS wastewater surveillance; metagenomics; public health; antimicrobial resistance; pathogen genomics; India

ARTICLE AT A GLANCE

- Wastewater surveillance measures pooled community signals and does not diagnose individuals.
- Targeted PCR is usually better for frequent tracking of known threats; metagenomics is valuable for broad discovery and AMR profiling.
- Sewer coverage, rainfall, flow, catchment population, and laboratory recovery strongly affect interpretation.

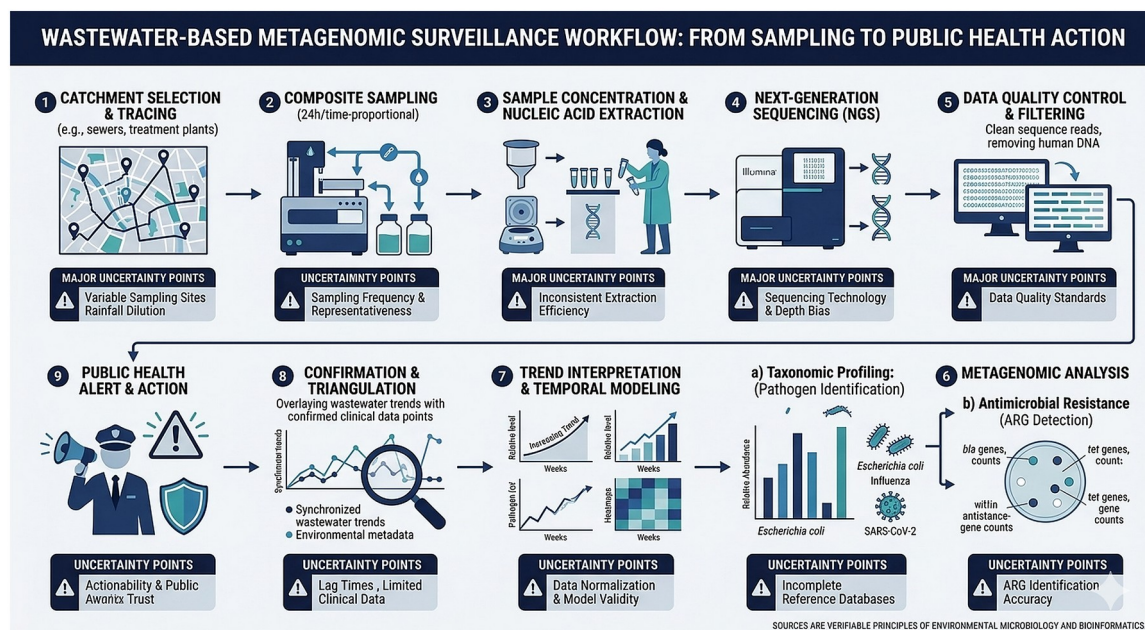


Figure 1. Wastewater-based metagenomic surveillance workflow from catchment sampling and laboratory processing to data quality control, interpretation, and public-health action.

1. A community leaves biological traces in wastewater

People shed viruses, bacteria, parasites, human DNA, metabolites, and antimicrobial-resistance genes through faeces, urine, respiratory secretions, bathing, and household waste. These materials enter sewers, drains, septic systems, and treatment plants. A wastewater sample therefore contains a pooled signal from many individuals, including people who do not seek healthcare or receive diagnostic testing. This makes it attractive for population-level surveillance.

Wastewater surveillance is not new; poliovirus monitoring has long used environmental samples. COVID-19 expanded the approach globally. Studies showed that SARS-CoV-2 RNA in sewage could be detected before large numbers of reported cases and that concentration trends broadly tracked infection prevalence (Medema et al., 2020; Peccia et al., 2020). Indian researchers also demonstrated early that SARS-CoV-2 genetic material could be detected in wastewater, supporting the feasibility of the method under local conditions (Kumar et al., 2020).

The key advantage is coverage. One sample from a treatment-plant inlet can represent a large catchment. The limitation is that the signal is mixed. It usually cannot identify which person is infected, and it may not distinguish residents from commuters or travellers. Wastewater surveillance is therefore best suited to community trends and early warning, not individual diagnosis.

2. Metagenomics broadens the view

Targeted assays such as quantitative PCR look for a known genetic sequence. They are sensitive, relatively fast, and suitable for routine tracking of a defined pathogen. Metagenomics attempts to sequence a wider range of genetic material, allowing researchers to detect many organisms or genes in one sample. Shotgun metagenomics can profile bacterial communities and resistance genes; targeted enrichment can improve recovery of particular viruses or genomic regions.

The broader view is useful when the surveillance question is open. A city may wish to track antimicrobial-resistance genes, identify unexpected enteric viruses, or monitor how pathogen diversity changes after a flood. Hendriksen et al. (2019) used metagenomic analysis of urban sewage to compare AMR patterns internationally. Karthikeyan et al. (2022) showed that deep wastewater sequencing could reveal cryptic SARS-CoV-2 lineages and transmission signals that were difficult to observe through clinical sampling alone.

Breadth creates analytical challenges. Most sequences may come from harmless environmental organisms or human-associated bacteria. Reference databases contain errors and geographic biases. Detection of a resistance gene does not reveal whether it is in a pathogen, a harmless organism, or extracellular DNA. Assembly and host-linking methods can improve interpretation, but uncertainty should be reported explicitly.

3. The signal is shaped by infrastructure

Wastewater concentration depends on how many people contribute, how much water enters the system, how long material travels, and how quickly biological molecules degrade. Rain can dilute samples or wash new material into drains. Industrial chemicals may inhibit laboratory reactions. Temperature, pH, solids, and treatment processes affect recovery. Grab samples capture one moment, while composite samplers integrate several hours and often provide a more stable signal.

In cities with extensive sewer coverage, treatment plants provide clear catchment points. In many Indian settlements, however, sewage may move through open drains, septic tanks, mixed stormwater systems, or decentralised treatment. Populations contributing to a sample may be uncertain. This does not make surveillance impossible, but site selection must be mapped carefully. A sentinel drain near a hostel, market, hospital, or neighbourhood answers a different question from a large treatment-plant inlet.

Normalisation can reduce some variation. Researchers may compare pathogen concentration with flow, faecal indicators, or stable human-associated markers. Repeated sampling is more informative than isolated results. The objective is to detect a consistent rise, unusual sequence, or spatial contrast that justifies investigation.

4. From sequence detection to public-health action

A useful programme begins with a decision question. Public-health authorities may want to know whether influenza is rising, whether a new SARS-CoV-2 variant is present, whether cholera risk is changing after flooding, or whether hospital and municipal wastewater contain high-priority resistance genes. The organism, sampling frequency, laboratory method, and alert threshold should follow from that question.

Wastewater findings should be triangulated with clinical testing, hospital admissions, syndromic surveillance, medicine sales, rainfall, and outbreak reports. A rising signal can prompt targeted diagnostic testing, public communication, vaccination outreach, infection-control review, or additional sampling. The World Health Organization recommends integrating environmental surveillance into broader systems with defined objectives, quality assurance, ethical governance, and communication plans (World Health Organization, 2022).

False certainty is dangerous. Detection does not always mean viable pathogen, local transmission, or clinical risk. Absence does not prove a community is disease-free because recovery may be poor. Laboratories should use controls for sample collection, extraction, inhibition, sequencing, and contamination. Public dashboards should present trends and uncertainty rather than converting concentrations into exact case numbers without validated models.

5. A realistic Indian wastewater observatory

India could build a tiered network. Major treatment plants can provide routine city-level trends. Sentinel drains can cover poorly sewered or high-risk locations. Hospital wastewater can monitor healthcare-associated resistance, with safeguards against stigmatising facilities. Research laboratories can perform periodic metagenomic scans, while public-health laboratories use targeted assays for frequent monitoring of priority threats.

The system should invest in standard protocols, reference materials, data pipelines, and workforce. Sampling metadata must include time, location, flow, weather, catchment description, and processing method. Raw sequence data can be valuable for research, but governance is needed because fine-scale sampling may identify institutions or communities. Communication should focus on protective action rather than blame.

Wastewater metagenomics turns a discarded stream into public-health information. It cannot replace clinics, diagnostic laboratories, or epidemiology. It can reveal what those systems miss: infections among untested people, community trends, hidden variants, and environmental reservoirs of resistance. The sewer becomes an early-warning system only when signals are calibrated, repeated, interpreted cautiously, and connected to authorities prepared to respond.

6. The next surveillance frontier

A mature wastewater programme can expand beyond crisis response. Seasonal monitoring of influenza, respiratory syncytial virus, enteroviruses, norovirus, and selected bacterial pathogens could provide a background map of community circulation. AMR surveillance can track high-priority genes and mobile elements, while chemical

measurements can add information on antimicrobial use or other exposures. Expansion should be gradual because every target requires validated recovery, interpretation, and action pathways.

Metagenomic archives create retrospective value. When a new pathogen or resistance mechanism is identified, stored extracts and sequence data can be reanalysed to determine when and where it first appeared. This requires consistent sample preservation, metadata, and data governance. Reference databases must also be updated, because a sequence classified as unknown today may become informative later.

The strongest measure of success is avoided harm, which is difficult to prove. Intermediate indicators can include lead time over clinical detection, geographic coverage, proportion of alerts confirmed by targeted testing, time from alert to response, and decisions influenced. Cost-effectiveness should compare wastewater surveillance with additional clinical sampling rather than assuming one replaces the other. The system is valuable when it adds earlier or broader information at a cost public-health institutions can sustain.

7. Building public confidence before the next emergency

Wastewater systems are most useful when they operate before an emergency. Routine sampling establishes normal seasonal variation, laboratory recovery rates, and site-specific baselines. Without that background, an unusual signal may reflect a changed sampler, a rain event, or a laboratory batch rather than a biological event. Public communication protocols should also be prepared in advance, including who verifies a finding, who contacts local authorities, and how uncertainty is described. During a crisis, these decisions are harder and more politicised. A standing network can conduct proficiency testing, maintain supplies, and preserve institutional memory between outbreaks. The infrastructure may appear quiet when disease levels are low, but that is when calibration occurs. Early warning depends on knowing what ordinary wastewater looks like.

8. Editorial perspective

The public-health value of wastewater data lies in disciplined interpretation. A positive sequence is a clue, not a diagnosis; a rising trend is a prompt, not proof of cause. Articles and dashboards should state the catchment, method, recovery controls, uncertainty, and action threshold. This makes the system more credible and helps decision-makers use it without converting complex environmental measurements into false precision.

Key takeaways

- Wastewater surveillance measures pooled community signals and does not diagnose individuals.
- Targeted PCR is usually better for frequent tracking of known threats; metagenomics is valuable for broad discovery and AMR profiling.
- Sewer coverage, rainfall, flow, catchment population, and laboratory recovery strongly affect interpretation.
- Trends and anomalies should be triangulated with clinical and environmental data before action.
- India needs a tiered network combining treatment plants, sentinel drains, hospitals, targeted assays, and periodic metagenomics.

Declarations

Author contributions: All listed authors contributed to literature synthesis, manuscript development, revision, and approval of the publication version.

Funding: No specific funding was reported for this article.

Competing interests: The authors declare no competing interests.

Ethics: Not applicable. This article does not report new research involving human participants, animals, or identifiable personal data.

AI-assisted tools: AI-assisted tools were used for language refinement, structural review, and preparation of selected conceptual visuals. All scientific claims, citations, interpretations, and final text were reviewed and approved by the authors and the NiviScholar editorial team.

How to cite this article

Rajnandini Vasantrao Nevhal; Khandare Karuna Vinayak; Karan Vaijanath Waghmare (2026). Wastewater Metagenomics: A Public-Health Early-Warning System Hidden in the Sewer. NiviScholar Insights, NSI-CR-2026-002. <https://insights.nivischolar.com/articles/wastewater-metagenomics-public-health-early-warning>

References

- Aarestrup, F. M., & Woolhouse, M. E. J. (2020). Using sewage for surveillance of antimicrobial resistance. *Science*, 367(6478), 630–632. <https://doi.org/10.1126/science.aba3432>
- Bivins, A., North, D., Ahmad, A., Ahmed, W., Alm, E., Been, F., Bhattacharya, P., Bijlsma, L., Boehm, A. B., Brown, J., Buttiglieri, G., Calabro, V., Carducci, A., Castiglioni, S., Cetecioglu Gurol, Z., Chakraborty, S., Costa, F., Curcio, S., de los Reyes, F. L., ... Bibby, K. (2020). Wastewater-based epidemiology: Global collaborative to maximize contributions in the fight against COVID-19. *Environmental Science & Technology*, 54(13), 7754–7757. <https://doi.org/10.1021/acs.est.0c02388>
- Hendriksen, R. S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S. K., Kjeldgaard, J., Kaas, R. S., Clausen, P. T. L. C., Vogt, J. K., Leekitcharoenphon, P., van de Schans, M. G. M., Zuidema, T., de Roda Husman, A. M., Rasmussen, S., Petersen, B., ... Aarestrup, F. M. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications*, 10, 1124. <https://doi.org/10.1038/s41467-019-08853-3>
- Karthikeyan, S., Levy, J. I., De Hoff, P., Humphrey, G., Birmingham, A., Jepsen, K., Farmer, S., Tubb, H. M., Valles, T., Tribelhorn, C. E., Tsai, R., Sathe, S., Moshiri, N., Henson, B., Mark, A. M., Hakim, A., Baer, N. A., Barber, T., Belda-Ferre, P., ... Knight, R. (2022). Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. *Nature*, 609, 101–108. <https://doi.org/10.1038/s41586-022-05049-6>
- Kumar, M., Patel, A. K., Shah, A. V., Raval, J., Rajpara, N., Joshi, M., & Joshi, C. G. (2020). First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. *Science of the Total Environment*, 746, 141326. <https://doi.org/10.1016/j.scitotenv.2020.141326>
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., & Brouwer, A. (2020). 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*, 7(7), 511–516. <https://doi.org/10.1021/acs.estlett.0c00357>
- Peccia, J., Zulli, A., Brackney, D. E., Grubaugh, N. D., Kaplan, E. H., Casanovas-Massana, A., Ko, A. I., Malik, A. A., Wang, D., Wang, M., Warren, J. L., Weinberger, D. M., Arnold, W., & Omer, S. B. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature Biotechnology*, 38, 1164–1167. <https://doi.org/10.1038/s41587-020-0684-z>
- Pruden, A., Larsson, D. G. J., Amézquita, A., Collignon, P., Brandt, K. K., Graham, D. W., Lazorchak, J. M., Suzuki, S., Silley, P., Snape, J. R., Topp, E., Zhang, T., & Zhu, Y.-G. (2013). Management options for reducing the release of antibiotics and antibiotic resistance genes to the environment. *Environmental Health Perspectives*, 121(8), 878–885. <https://doi.org/10.1289/ehp.1206446>
- World Health Organization. (2022). Environmental surveillance for SARS-CoV-2 to complement public health surveillance: Interim guidance. <https://www.who.int/publications/i/item/WHO-HEP-ECH-WSH-2022.1>

© 2026 The authors. Published by NiviScholar under the Creative Commons Attribution 4.0 International License.