Science of the Total Environment 940 (2024) 173315

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Encrypted data-sharing for preserving privacy in wastewater-based epidemiology

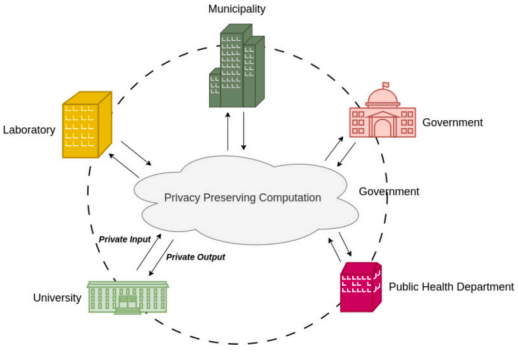
Erin M. Driver [a], Manazir Ahsan [b,c], Lucas Piske [b,c], Heewook Lee [b,c], Stephanie Forrest [b,c,d], Rolf U. Halden [a,e,*], Ni Trieu [b,c,**]

[a] Biodesign Center for Environmental Health Engineering, Arizona State University, Tempe, AZ 85287, USA
[b] Biodesign Center for Biocomputing, Security and Society, Arizona State University, Tempe, AZ 85287, USA
[c] School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85287, USA
[d] College of Health Solutions, Arizona State University, Tempe, AZ 85287, USA
[e] School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85287, USA

HIGHLIGHTS

- Developed homomorphic encryption framework to share wastewater data
- Illustrated encrypted mass balance calculations between municipalities
- Execution time, communication costs, and calculation precision are scalable.
- Framework enables data sharing at neighborhood-level for public health surveillance.
- System designed for data security and responsible data sharing with user flexibility

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Damià Barceló

Keywords:
Wastewater-based surveillance
Privacy
Homomorphic encryption
Ethics

ABSTRACT

The rapidly expanding use of wastewater for public health surveillance requires new strategies to protect privacy rights, while data are collected at increasingly discrete geospatial scales, i.e., city, neighborhood, campus, and building-level. Data collected at high geospatial resolution can inform on labile, short-lived biomarkers, thereby making wastewater-derived data both more actionable and more likely to cause privacy concerns and stigmatization of subpopulations. Additionally, data sharing restrictions among neighboring cities and communities can complicate efforts to balance public health protections with citizens' privacy. Here, we have created an encrypted framework that facilitates the sharing of sensitive population health data among entities that lack trust for one another (e.g., between adjacent municipalities with different governance of health monitoring and data sharing). We demonstrate the utility of this approach with two real-world cases. Our results show the feasibility of sharing encrypted data between two municipalities and a laboratory, while performing secure private computations for wastewater-based epidemiology (WBE) with high precision, fast speeds, and low data costs. This framework is amenable to other computations used by WBE researchers including population normalized mass

* Correspondence to: R.U. Halden, Biodesign Center for Environmental Health Engineering, Arizona State University, Tempe, AZ 85287, USA.
** Correspondence to: N. Trieu, Biodesign Center for Biocomputing, Security and Society, Arizona State University, Tempe, AZ 85287, USA.
E-mail addresses: rolf.halden@asu.edu (R.U. Halden), nitrieu@asu.edu (N. Trieu).

https://doi.org/10.1016/j.scitotenv.2024.173315
Received 18 February 2024; Received in revised form 29 April 2024; Accepted 15 May 2024
Available online 17 May 2024

loads, fecal indicator normalizations, and quality control measures. The Centers for Disease Control and Prevention's National Wastewater Surveillance System shows ~8 % of the records attributed to collection before the wastewater treatment plant, illustrating an opportunity to further expand currently limited community-level sampling and public health surveillance through security and responsible data-sharing as outlined here.

1. Introduction

Wastewater-based epidemiology (WBE) has become a popular public health tool to supplement case-based surveillance at the population-level. Wastewater is typically collected at a centralized wastewater treatment plant (WWTP), where the health and behavior of an entire city population of thousands to millions of people are captured in a single sample. However, sampling at the WWTP scale provides only an average of the infection levels or drug use within the community. For public health surveillance, a more refined geospatial resolution is desirable to divide the city and its total population into sub-sewersheds or sub-populations, to identify hotspots of public health concern that can then be targeted with suitable interventions. This was rare prior to the COVID-19 pandemic, but has since become more common with the goal of identifying vulnerable and at-risk populations.

This work at the sub-sewershed level predominantly includes monitoring at the building-scale including universities, hospitals, and long-term care facilities (Acosta et al., 2021; Davó et al., 2021; Gibas et al., 2021; Wright et al., 2022). Although building-scale sampling is useful, this sampling granularity is more prone to ethical concerns over privacy invasion (fewer people contributing to a sample), and purposefully excludes some locations and their respective subpopulations, creating a potential for inequality (Bowes et al., 2023a). To avoid these limitations, sampling at the neighborhood-level is an attractive alternative. However, the inherent nature of wastewater collection systems often commingles wastewater across neighborhoods. Therefore, it may not be possible to collect a single sample that is representative of a neighborhood, but rather multiple samples must be collected upstream and downstream of areas of interest, and the measurements combined mathematically to account for area-specific impacts using mass balance approaches (Bowes et al., 2023b; Driver et al., 2023; Rainey et al., 2022). Challenges arise when wastewater commingling occurs across different municipalities. This commingling requires cooperation and data-sharing to achieve public health goals. However, one or more municipalities may have reservations regarding data sharing for legal, political or cultural reasons.

To effectively employ neighborhood-level monitoring, novel methods for sharing WBE data are needed to address the concerns discussed above, particularly those related to data security and privacy across multiple stakeholders within this rapidly developing field (Jacobs et al., 2021). One way to facilitate cooperative data sharing in WBE involves the use of data privacy techniques, including encryption. Encryption converts plaintext (the original data) into ciphertext in such a way that the ciphertext is unreadable without access to a secret key to retrieve the original plaintext. Normally, encryption prevents computation on ciphertext, however homomorphic encryption (HE) supports computation over the encrypted data, producing an encrypted output. Homomorphic encryption has been used in healthcare for querying protected patient electronic health records (EHRs) for relevant information to answer pressing research and public health-related questions (Domadiya and Rao, 2022). EHRs contain a wealth of information in centralized databases that may be accessed by many types of institutions with varying levels of security, posing a significant threat to data security. HE has helped to alleviate that risk by allowing for cloud-based computations that are safely shared among researchers and health professionals (Kocabas et al., 2013; Munjal and Bhatia, 2023; Zhang et al., 2023). This same framework could be used for wastewater-derived data.

In this study, HE techniques are used to facilitate data sharing between two municipalities and one analytical laboratory involved in neighborhood-level WBE public health assessments. The objectives of this work were to create a framework that met minimum requirements for execution time, communication costs, and precision in calculations of WBE metrics. This work illustrates how computer science techniques can help WBE mature into a secure and widely accepted public health service that is practical, informative, and safe for its participating stakeholders and entities.

2. Materials and methods

We developed the *cryptWWDB (encrypted wastewater database)* framework to facilitate secure and private data sharing while addressing the following concerns: (1) data availability for legitimate public health surveillance purposes and/or scientific research; (2) support for joint queries and analyses across multiple databases to provide answers to public health questions; (3) supporting the need to combine WBE data securely with other kinds of data, e.g., from Health Departments or health-care providers; and (4) protection from adverse uses including repeated queries used to build an overall picture of a database (illegitimate data mining) when the data are not available for download. In this section, we first define two use case scenarios taken from real-world situations encountered by our team during their 15 years of wastewater-based monitoring.

2.1. Use cases

Use Case 1. In this scenario, an entity who is not interested in obtaining wastewater monitoring data must share information with another entity in order to allow the latter to collect health data of interest. Specifically, Municipality A (Muni A) is collecting a wastewater sample from a single location to learn about health priorities (e.g., presence and daily loads of heroin and its human metabolite in city sewers). Due to the nature of the municipal wastewater system, Muni A's sample also contains wastewater from an upstream community that resides in Municipality B (Muni B), which is not interested nor authorized by its residents to obtain and share such information. To remove the contribution of Muni B from the comingled wastewater of both cities, an additional sample has to be taken upstream of Muni A, within their community, and analyzed to enable subtraction from the commingled sample. The resultant mass in Muni A can be calculated using Eq. (1).

$$\text{MassLoad1} = (\text{Q1}^*\text{C1}) - (\text{Q2}^*\text{C2}) \quad (1)$$

Here, MassLoad1 is the mass load of Muni A, Q1 is the total daily volumetric wastewater flow in Muni A, Q2 is the total daily volumetric flow of wastewater in Muni B, C1 is the concentration of the target chemical of interest in the Muni A sample, and C2 is the concentration of the target chemical of interest in the Muni B sample.

Use Case 2. This is a more advanced scenario involving the addition of a temporal component to scenario 1, specifically, the sample collection time is predetermined, and coincides with changes in population composition or behavior (e.g., a major sporting event or a music festival). Each sample from both municipalities must be collected in the morning on the same day, and the corresponding flow data from each municipality must also be representative of that day. Eq. (2) illustrates the addition of this temporal component where a subscript t denotes a specific predetermined time.

$$\text{MassLoad1t} = (\text{Q1t}^*\text{C1t}) - (\text{Q2t}^*\text{C2t}) \quad (2)$$

In addition to the specifications discussed in the use cases outlined above, other general relationship details between the entities are as follows: (1) the upstream municipality (Muni B) does not want Muni A to know their wastewater test results (and vice versa); (2) Muni B also does not want the third-party laboratory to have access to mass load results and does not want the laboratory to perform any unencrypted data analyses.

2.2. cryptWWDB development

Here, we use HE as a cryptographic building block to implement *cryptWWDB*. While various HE schemes, such as ElGamal (ElGamal, 1986) or Paillier encryptions (Paillier, 1999), are available, we developed *cryptWWDB* using an HE scheme based on the Ring-learning with error (RLWE) assumption (Brakerski and Vaikuntanathan, 2011; Stehlé et al., 2009) which offers improved speed and quantum resistance (protection from classical and quantum computers). Execution time, communication cost, and precision related to the results of the encrypted calculations were all factors considered in the development of this framework. Computational costs are defined as the lump sum of costs incurred from (1) encrypting the data elements; (2) transmitting them to the server; (3) computing a mathematical function without decrypting the data; (4) returning the answer in encrypted form; and (5) decrypting the data by the user. These costs are dominated by the time it takes to compute the encrypted data (step iii), with the other steps being insignificant by comparison. Computational precision was determined by comparing the analytical chemistry result (true value) to the output of the encryption algorithm to be acceptable if the error imparted by the encryption computation was much less than the error imparted by laboratory processing and analysis, which is typically on the order of ±30 %, as frequently determined by spike recovery experiments conducted in multiple replicates by analytical labs. We conducted our experiments on a machine equipped with an Intel(R) Core(TM) i7-11700F 2.50GHz processor and 32GB of RAM. The implementation is written in Python and uses the Tenseal library (Ayoub et al., 2021) for HE.

3. Results

Homomorphic encryption (HE) was used to facilitate data sharing between two municipalities involved in neighborhood-level WBE, including a third-party analytical laboratory generating chemical analyte data (e.g., heroin and 6-acetylmorphine) from wastewater for each city. The protocol uses HE to securely transfer encrypted data (ciphertext), enabling computation on the encrypted data while preserving the privacy of the sensitive information. Basic HE components are outlined in Fig. 1.

3.1. Overview of the cryptWWDB framework

The *cryptWWDB* workflow is outlined in Fig. 2. For Use Case 1, the follow framework satisfies the requirements of each of the participants and the goals of data sharing and security. Muni A is the entity interested in having mass load information generated to support public health decision-making. Muni A contributes their wastewater flow data (Q1) to Eqs. (1) and (2) and generates a private key (secret key [sk] unique to Muni A) to encrypt their data (plaintext converted to ciphertext) using the query interface. This ciphertext is sent to the computation coordinator (a policy checker), which checks whether the query satisfies a system security policy (access controls, repeated queries, etc.). If yes, the coordinator forwards the encrypted query to the third-party lab. Muni A also generates a public (pk), which Muni B uses to encrypt its wastewater flow data (Q2) and the third-party laboratory uses to encrypt the wastewater measurements (e.g., heroin and 6-actylmorphine [metabolite]) for both Muni A (C1) and Muni B (C2). The laboratory then provides the encrypted concentration data (as ciphertexts) to *cryptWWDB*, which uses homomorphic encryption (HE) to complete the computation with the resulting output remaining encrypted. This is accomplished through an evaluation (evk) key also generated by Muni A. The result is then transferred back to Muni A by the laboratory through the system coordinator, and using its private key, Muni A decrypts the result to finish the query without knowledge of Muni B's private data.

In use case 2 with the addition of a time component, *cryptWWD*B was adapted to manage queries involving multiple data fields. Encrypted query time was added to the computation, specifically, each municipality encrypted the time component related to their flow data, as well as their wastewater flow data, and sent to the laboratory for further computation. The laboratory through the *cryptWWD*B framework performs an equality check (ciphertext-ciphertext comparison) that compares the times from each of the two municipalities and records a 1 if the times are equivalent or 0 if they are not. HE mass loads are calculated when the encrypted values conform to equivalent time intervals. Additional details are included in the supplemental information.

3.2. Performance of cryptWWDB

Efficiency and accuracy considerations are critically important for

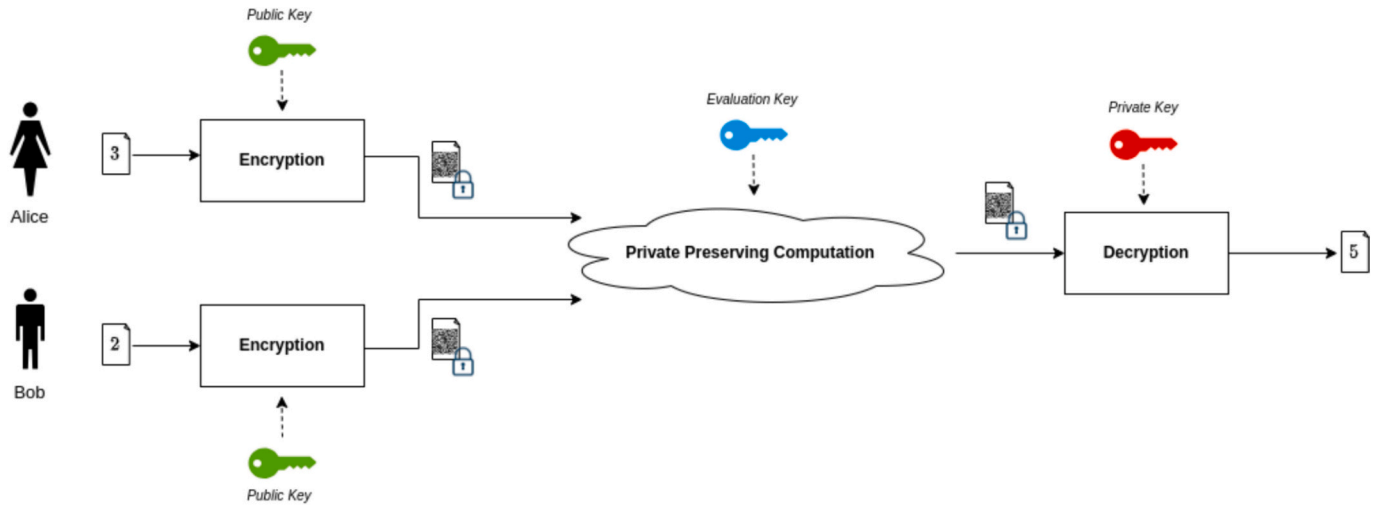


Fig. 1. Homomorphic encryption. Ciphertext from two individuals is encrypted using a public key. An evaluation key then allows computations to be performed on the encrypted data to create an encrypted result. A private key is then used to decrypt the result to ciphertext.

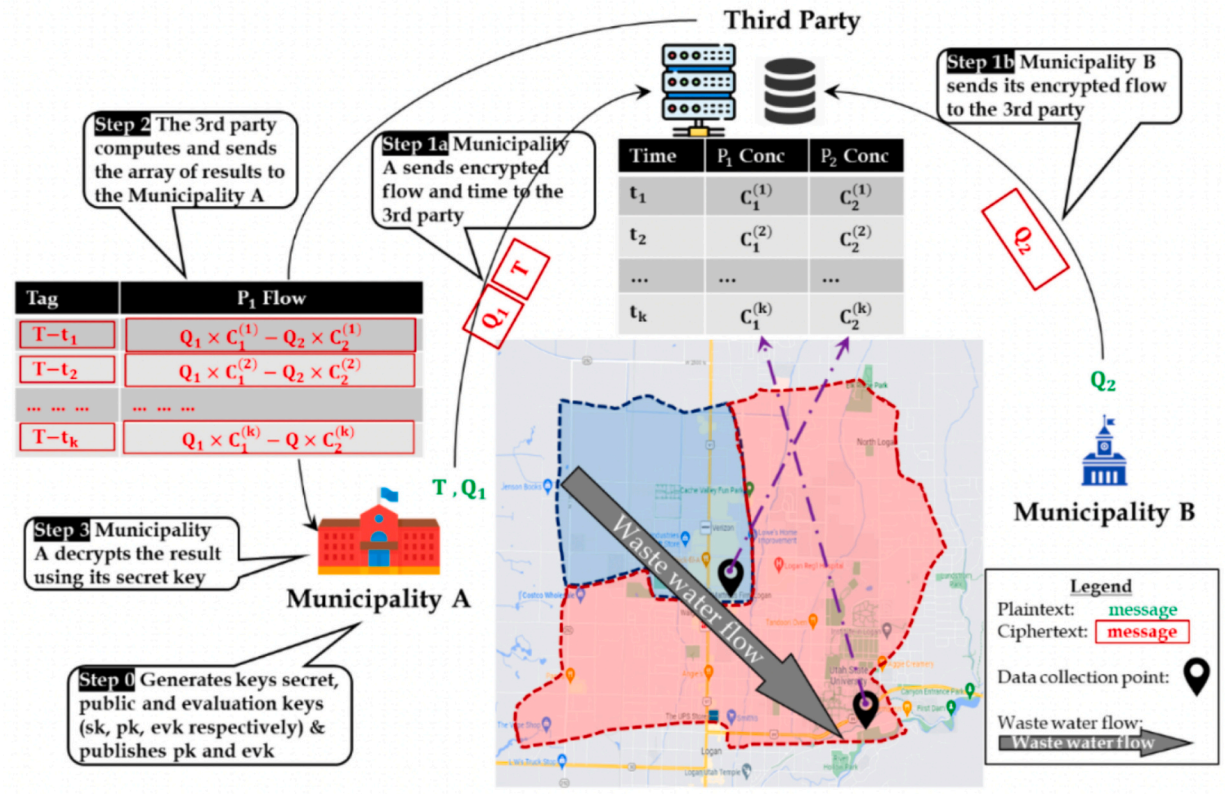


Fig. 2. Overview of *cryptWWDB* protocol. Each municipality has its own wastewater flow data and the third-party laboratory has heroin/6-acetylmorphine concentration data for each of the two municipalities. Municipality A is requesting the computation of mass load and generates a secret (sk), public (pk), and evaluation (evk) keys (pk and evk available for others) [Step 0]. Muni A and the laboratory send encrypted flow data (time-dependent) and chemical analyte concentration data using the pk [Step 1]. The third-party laboratory computes time-dependent mass load directly on encrypted data and sends the result to Muni A [Step 2]. Finally, Muni A obtains the desired mass load by decrypting using the secret key without learning the data of other municipalities.

data encryption, including execution time, communication cost, and precision related to the results of the encrypted calculations. Performance results for Use Cases 1 and 2 are shown in Table 1. Use Case 1 is less complex compared to Use Case 2, because it does not consider collection time, and as expected it exhibits faster running times and lower communication costs. Precision of the calculations can be affected by the encryption process resulting in fewer significant digit pre- versus post- encryption. In both use cases, precision is unaffected by the encrypt/decrypt process, with significant figures to the ten thousandth place (1e-5). This result is more than appropriate for wastewater-derived data.

Table 1
Performance of *cryptWWDB* including execution time (milliseconds) and communication cost (megabytes) for the various components of Use Case 1, 2, and modified 2 (time-efficient).

Characteristics	Use case 1	Use case 2	Use case 2[a]
Execution time (ms)			
Muni A, B	79.78	93.75	39.89
Laboratory	36.90	53.86	31.91
Total	116.68	147.61	71.80
Communication cost (MB)			
Muni A → Muni B	–	–	0.36
Muni B → Muni A	–	–	0.36
Muni A → Laboratory	2.87	3.06	4.44
Muni B → Laboratory	0.72	0.91	0.55
Laboratory → Muni A	0.63	0.31	0.40
Total	4.25	4.29	6.11

Notes: ms - milliseconds; MB – megabytes.
[a] Time-efficient.

3.2.1. *Enhanced performance for use case 2*

In Use Case 2, the primary computational bottleneck is the time required by the laboratory to handle the temporal data in encrypted form. That computation alone accounts for over 50 % of the total running time (Table 1). To enhance efficiency, a time-efficient protocol was designed that leverages computational resources from both Muni A and Muni B rather than relying solely on the laboratory. The use case is identical to the original, except that instead of the laboratory performing the ciphertext-ciphertext comparison of time points, the municipalities provide the equality check. Muni A sends its encrypted time to Muni B and Muni B performs the computation. This computation is a plaintext-ciphertext comparison which is considerably less resource-intensive than the ciphertext-ciphertext comparison. Therefore, the comparison step in this modified Use Case 2 not does not create increased execution times as in its previous iteration. Unfortunately, this time-efficient protocol requires higher communication costs due to the additional data exchanged (Table 1). Systems with high computational power but limited bandwidth may opt for the original protocol, whereas systems with slower processing capabilities and faster networks may favor the time-efficient protocol. In both cases, the additional overhead of *cryptWWDB* is unlikely to be an impediment to adoption, given today's widely available computing resources.

4. Discussion

Unlike the ideal security setting in the cryptographic literature (Oded, 2009), which hides the entire statement, including the function to be computed, *cryptWWDB* hides only the sensitive parameters of the query predicate. In Use Case 1, *cryptWWDB* conceals only the wastewater flow data from the participants, while revealing the formula Eq.

(1) to the third-party laboratory. Although hiding all query information (e.g., including Eqs. (1) and (2)) provides better privacy guarantees for clients, there are two main disadvantages. First, the returned values from the laboratory to Muni A need to be padded with dummy values to hide the distribution of actual output and query type (Pinkas et al., 2015; Pinkas et al., 2018), which adds significant communication and computation cost, especially when the input data are large. Second, implementing a policy checker for a hidden query is challenging. Alternatively, if the query template is revealed, the coordinator can implement a separate policy checker that provides fine-grained control over the query types. For example, it can allow queries only for aggregation values of certain columns and reject all queries that retrieve individual records. However, if the query is completely hidden, one would have to use more expensive techniques like zero-knowledge proofs (Boyle et al., 2021) to achieve the same functionality. These considerations motivated our decision to reveal the query template (Eqs. (1) and (2)) and hide only the sensitive query parameters.

To the best of our knowledge, there is no existing work in the cryptography literature that addresses the same set of challenges as *cryptWWDB*. Further, there is a notable absence of systems designed to handle secure three-party queries using HE. While there are solutions based on multi-party computation (MPC) (Poddar et al., 2020; Volgushev et al., 2019) that also process multiparty queries, they often involve multiple rounds of communication and may have limitations on input size. In contrast, the *cryptWWDB* framework is constructed on HE, which offers a straightforward conceptualization of this cryptographic tool. This accessibility extends to non-technical individuals, including non-expert users and professionals such as wastewater researchers.

The cryptWWDB framework is not without its limitations. As designed, the query initiator (Muni A in our examples) is the only party with access to the decrypted values without ever directly accessing the data of other entities. One critical assumption is that the third-party (the laboratory in our examples) and the query initiator (Muni A) do not collude. If they do, Muni A could simply decrypt data using their secret key because Muni A is the key generator (private, public and encryption). To overcome this limitation a different type of HE, known as Multi-key HE would be used in future iterations (López-Alt et al., 2012). In this method, each entity has its own secret key, and decrypting a message requires that all of the entities participate in a joint computation, each using its own secret key.

We take advantage of the fact that the HE encryption scheme allows direct computations on encrypted data without having to decrypt it first. While the present scenarios involve only a small number of entities, in more realistic scenarios, commingling may occur across sewers of more than two municipalities, and *cryptWWDB* can easily be adapted to these scenarios. Additional multilayered computations are also feasible, which can provide other types of data handling and data analysis relevant to WBE (Table 2). This can include incorporating other data streams into the analysis. Here we illustrated mass load calculations using a measured concentration and flow, however a common data source in WBE is population (the number of people contributing to a sewershed) (Gatidou et al., 2016; Ort et al., 2014). Population can be a constant (US census-derived), quasi-constant (weekday/weekend through use of employment data), and unique daily values (wastewater population biomarkers) (Choi et al., 2018). Other data streams to consider are fecal indicators (e.g., pepper mild mottle virus) for normalizing feces-derived biomarkers like SARS-CoV-2 (Feng et al., 2021), inclusion of pharmacokinetic excretion values to estimate drug consumption versus excretion (Zuccato et al., 2008), and degradation factors to correct for in-sewer biomarker losses (Hart and Halden, 2020). Another type of analysis that may be performed using this framework is quality control measures. This may include performing no calculations when one of the participating municipalities does not collect a sample, or reporting method detection limits in lieu of zero for non-detect measurements or when mass balance subtractions between two communities result in negative values (Bowes et al., 2023b; Tempe, 2023a).

Table 2
Additional parameters that could be included in the *cryptWWDB* framework.

Parameter	Details
Data stream	
Population estimates	WWTP population served (constant), employment data (weekday/weekend differences), chemical measurements or WIFI data (unique) to derive per capita estimates
Fecal normalization	Viral (PMMoV), Bacterial (Bacteroides HF183), Chemical (coprostanol) estimators for changing fecal quantities
Excretion factors	Urinary excretion values & molecular weights to estimate consumption
Degradation factors	Degradation coefficients to correct mass loads for in-sewer degradation
Quality control	
Missing data & non-detects	Upstream sample not collected; handling of non-detects (e.g., MDLs)
Negative mass balances	Identification/response to negative calculations (e.g., negative number changed to MDL or non-detect)
Error estimate inclusion	Error bars calculated from instrument, population, excretion, or other error
Task	
Trigger point calculations	Aggregating data when population thresholds are not met or a specific day of the week triggering different population estimates used
Average calculations	Weekly or rolling averages to assess long-term trends
Percent increases	Week-to-week or month-to-month changes for quick assessment of change (e.g., for public-facing dashboards)

Notes: WWTP – wastewater treatment plant, PMMoV – pepper mild mottle virus, MDLs – method detection limits.

Quality assurance factors may also constitute the addition of error estimates on reported values imparted to the data via sampling, laboratory analysis, or population estimates (Banta-Green et al., 2016). Additional task-based assessments that could be performed using this framework are trigger point calculations. For example, if a minimum threshold population in a community is not met within a single wastewater catchment, then data area aggregated across adjacent catchments to protect privacy, or a specific day of the week triggers use of different population estimates (e.g., weekday versus weekend). Additionally, data are often presented to stakeholders in aggregate form with summary statistics, including weekly averages or percent increase (or decrease) changes from one sampling period to the next (Tempe, 2023b).

The US federal government (Centers for Disease Control and Prevention) in response to successes tracking SARS-CoV-2 in wastewater created the National Wastewater Surveillance System (NWSS) in 2020 (NWSS, 2023). As of December 17, 2023 there were over 730,000 individual records (measurements) of SARS-CoV-2 in wastewater across the country. Of those records, ~8 % are defined as "before the wastewater treatment plant," nomenclature used to signify that the sample was collected from within the collection system. Collection occurred in twelve states (Arizona [AZ], California, Illinois, Florida, Maine, Massachusetts, Michigan [MI], New York, Pennsylvania, Texas, Virginia, Wisconsin) including the District of Columbia, with two states, AZ and MI, accounting for ~65 % of the total number of records. These results suggest WBE is still in the early stages of transitioning from sample collection at wastewater treatment plants to upstream within the collection system due to logistical challenges, including data sharing between communities. Therefore, now is a critical time to begin creating computational infrastructure to securely facilitate community-level data sharing for public health surveillance.

5. Conclusions

The cryptWWDB (encrypted wastewater database) framework created here addresses data sharing and privacy concerns that the authors took directly from their wastewater monitoring experiences. The framework

incorporated fast computing speeds and low data costs with a high precision in wastewater-derived estimates to solve data sharing challenges between different municipalities and a laboratory. The study illustrates that municipalities with different data collection objectives and data sharing preferences can productively cooperate to protect their interests while promoting: (1) public health assessments; (2) identification of communities facing public health challenges; and (3) enabling an assessment of the impact of targeted health and educational interventions in vulnerable subpopulations (e.g., monitoring of schools, hospitals, neighborhoods dominated by demographic minorities, or temporary communities, such as crowds gathering for large events in sporting or entertainment). The principal significance and novelty of this work rest with providing a practical approach of how to use the established, proven encryption strategy of HE and apply it to the field of wastewater monitoring and public health surveillance, as conducted today by thousands of municipalities and communities around the world. Conventional HE cannot serve as a turn-key application to WBE data because the standard HE algorithms are computationally too inefficient to be applied directly to this setting of public health surveillance by cities and analytical labs. Significant innovations presented here include: (1) the reconfiguration of standard HE algorithms such that the computational burden is placed on the server while ensuring that the server gains no knowledge of the sensitive input data; (2) an analysis of the unique data privacy needs and concerns inherent to the use of WBE data by municipalities and public health stakeholders; and (3) a computation framework that is specifically applicable to currently unanswered questions of the wastewater surveillance sector, (i.e., how to share sensitive information that protects the privacy of stakeholders while also maximizing the public health benefit of WBE data, which are acquired post-COVID-19 at a massive scale, at costs exceeding hundreds of millions of US$ per annum). To maximize the visibility, acknowledgment and future use of the provided framework, this analytical approach is published in the general scientific literature rather than in a specialized computer science journal. The *cryptWWDB* framework and associated source code is available to all practitioners by request. It does not require any special computational tools or equipment, and therefore can potentially be used widely across the U.S. and other countries where wastewater monitoring for public health protection is already in use or may be implemented in the future.

CRediT authorship contribution statement

Erin M. Driver: Writing – original draft, Investigation, Data curation. **Manazir Ahsan:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Lucas Piske:** Writing – review & editing, Investigation. **Heewook Lee:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Stephanie Forrest:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Rolf U. Halden:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Ni Trieu:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge support National Science Foundation award 2115075.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.173315.

References

Acosta, N., Bautista, M.A., Hollman, J., McCalder, J., Beaudet, A.B., Man, L., et al., 2021. A multicenter study investigating SARS-CoV-2 in tertiary-care hospital wastewater. Viral burden correlates with increasing hospitalized cases as well as hospital-associated transmissions and outbreaks. Water Res. (Oxford) 201, 117369.

Ayoub, B., Retiat, B., Cebere, B., Belfedhal, A.E., 2021. TenSEAL: A Library for Encrypted Tensor Operations Using Homomorphic Encryption. arXiv.org.

Banta-Green, C.J., Brewer, A.J., Ort, C., Helsel, D.R., Williams, J.R., Field, J.A., 2016. Using wastewater-based epidemiology to estimate drug consumption—statistical analyses and data presentation. Sci. Total Environ. 568, 856–863.

Bowes, D.A., Darling, A., Driver, E.M., Kaya, D., Maal-Bared, R., Lee, L.M., et al., 2023a. Structured Ethical Review for Wastewater-Based Testing. Environ. Sci. Technol. 57 (35), 12969–12980. https://doi.org/10.1021/acs.est.3c04529.

Bowes, D.A., Driver, E.M., Kraberger, S., Fontenele, R.S., Holland, L.A., Wright, J., et al., 2023b. Leveraging an established neighbourhood-level, open access wastewater monitoring network to address public health priorities: a population-based study. Lancet. Microbe 4, e29–e37.

Boyle, E., Gilboa, N., Ishai, Y., Nof, A., 2021. Sublinear GMW-Style Compiler for MPC with Preprocessing. Springer International Publishing, Cham, pp. 457–485.

Brakerski, Z., Vaikuntanathan, V., 2011. Fully Homomorphic Encryption from Ring-LWE and Security for Key Dependent Messages. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 505–524.

Choi, P.M., Tscharke, B.J., Donner, E., Apos, Brien J.W., Grant, S.C., et al., 2018. Wastewater-based epidemiology biomarkers: past, present and future. Trends Anal. Chem. 105, 453–469.

Davó, L., Seguí, R., Botija, P., Beltrán, M.J., Albert, E., Torres, I., et al., 2021. Early detection of SARS-CoV-2 infection cases or outbreaks at nursing homes by targeted wastewater tracking. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases 27 (7), 1061–1063. https://doi.org/10.1016/j.cmi.2021.02.003.

Domadiya, N., Rao, U.P., 2022. ElGamal homomorphic encryption-based privacy preserving association rule mining on horizontally partitioned healthcare data. J. Instit. Eng. (India): Ser. B 103, 817–830.

Driver, E.M., Bowes, D.A., Halden, R.U., 2023. Expansion and diversification of wastewater-based epidemiology strategies in pandemic conditions to serve immediate public health goals. In: Wastewater-Based Epidemiology for the Assessment of Human Exposure to Environmental Pollutants. Elsevier Science & Technology, United States.

ElGamal, T., 1986. On Computing Logarithms Over Finite Fields. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 396–402.

Feng, S., Roguet, A., McClary-Gutierrez, J.S., Newton, R.J., Kloczko, N., Meiman, J.G., et al., 2021. Evaluation of sampling, analysis, and normalization methods for SARS-CoV-2 concentrations in wastewater to assess COVID-19 burdens in Wisconsin communities. ACS ES&T Water 1, 1955–1965.

Gatidou, G., Kinyua, J., van Nuijs, A.L.N., Gracia-Lor, E., Castiglioni, S., Covaci, A., et al., 2016. Drugs of abuse and alcohol consumption among different groups of population on the Greek Island of Lesvos through sewage-based epidemiology. Sci. Total Environ. 563-564, 633–640.

Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Roppolo Brazell, L., et al., 2021. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. Sci. Total Environ. 782, 146749.

Hart, O.E., Halden, R.U., 2020. Modeling wastewater temperature and attenuation of sewage-borne biomarkers globally. Water Res. 172.

Jacobs, D., McDaniel, T., Varsani, A., Halden, R.U., Forrest, S., Lee, H., 2021. Wastewater monitoring raises privacy and ethical considerations. IEEE Trans. Technol. Soc. 1.

Kocabas, O., Soyata, T., Couderc, J.P., Aktas, M., Xia, J., Huang, M., 2013. Assessment of cloud-based health monitoring using homomorphic encryption. In: 2013 IEEE 31st International Conference on Computer Design (ICCD), pp. 443–446.

López-Alt, A., Tromer, E., Vaikuntanathan, V., 2012. On-the-fly multiparty computation on the cloud via multikey fully homomorphic encryption. In: Annual ACM Symposium on Theory of Computing. ACM, pp. 1219–1234.

Munjal, K., Bhatia, R., 2023. A systematic review of homomorphic encryption and its contributions in healthcare industry. Complex Intell. Syst. 9, 3759–3786.

National Wastewater Surveillance System, 2023. COVID data tracker. Ctr. Dis. Control Prev. https://covid.cdc.gov/covid-data-tracker/#wastewater-surveillance.

Oded, G., 2009. Foundations of Cryptography: Volume 2, Basic Applications. Cambridge University Press.

Ort, C., van Nuijs, A.L.N., Berset, J.D., Bijlsma, L., Castiglioni, S., Covaci, A., et al., 2014. Spatial differences and temporal changes in illicit drug use in E urope quantified by wastewater analysis. Addiction 109, 1338–1352.

Paillier, P., 1999. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 223–238.
Pinkas, B., Schneider, T., Segev, G., Zohner, M., 2015. Phasing: Private Set Intersection Using Permutation-based Hashing. USENIX Security 15, Washington, D.C., USA, pp. 515–530.
Pinkas, B., Schneider, T., Zohner, M., 2018. Scalable private set intersection based on OT extension. ACM Trans. Priv. Secur. 21, 1–35.
Poddar, R., Kalra, S., Yanai, A., Deng, R., Popa, R.A., Hellerstein, J.M., 2020. Senate: A Maliciously-Secure MPC Platform for Collaborative Analytics. Cornell University Library, arXiv.org, Ithaca.
Rainey, A.L., Loeb, J.C., Robinson, S.E., Davis, P., Liang, S., Lednicky, J.A., et al., 2022. Assessment of a mass balance equation for estimating community-level prevalence of COVID-19 using wastewater-based epidemiology in a mid-sized city. Sci. Rep. 12, 19085.
Stehlé, D., Steinfeld, R., Tanaka, K., Xagawa, K., 2009. Efficient Public Key Encryption Based on Ideal Lattices. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 617–635.
Tempe Co, 2023a. Biomarker: COVID-19 Wastewater BioIntel Program, 2023. https://wastewater.tempe.gov/pages/biomarker-covid19.
Tempe Co, 2023b. Biomarker: Opioids - Fighting Opioid Misuse by Monitoring Community Health, 2023. https://wastewater.tempe.gove/pages/biomarker-opioids.
Volgushev, N., Schwarzkopf, M., Getchell, B., Varia, M., Lapets, A., Bestavros, A., 2019. Conclave: Secure Multi-party Computation on Big Data (Extended TR). Cornell University Library, arXiv.org, Ithaca.
Wright, J., Driver, E.M., Bowes, D.A., Johnston, B., Halden, R.U., 2022. Comparison of high-frequency in-pipe SARS-CoV-2 wastewater-based surveillance to concurrent COVID-19 random clinical testing on a public U.S. university campus. Sci. Total Environ. 820, 152877.
Zhang, L., Xu, J., Vijayakumar, P., Sharma, P.K., Ghosh, U., 2023. Homomorphic encryption-based privacy-preserving federated learning in IoT-enabled healthcare system. IEEE Trans Netw Sci Eng 10, 2864–2880.
Zuccato, E., Chiabrando, C., Castiglioni, S., Bagnati, R., Fanelli, R., 2008. Estimating community drug abuse by wastewater analysis. Environ. Health Perspect. 116, 1027.