



Université du Québec
à Rimouski

Développement d'un outil de normalisation populationnel pour l'analyse du SRAS-CoV-2 dans les eaux usées

Mémoire présenté

dans le cadre du programme de maîtrise en chimie de l'environnement et modélisation
environnementale

en vue de l'obtention du grade de maître ès sciences

PAR

© ELLIOTT OMER HOLLAND CHARTRAND

Mai 2025

Composition du jury :

Souhir Marsit, présidente du jury, Université du Québec à Rimouski

Richard St-Louis, directeur de recherche, Université du Québec à Rimouski

Marc-Denis Rioux, codirecteur de recherche, Université du Québec à Rimouski

Hugues Charest, examinateur externe, Laboratoire de santé publique du Québec

Dépôt initial le 12 février 2025

Dépôt final le 20 mai 2025

UNIVERSITÉ DU QUÉBEC À RIMOUSKI
Service de la bibliothèque

Avertissement

La diffusion de ce mémoire ou de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire « *Autorisation de reproduire et de diffuser un rapport, un mémoire ou une thèse* ». En signant ce formulaire, l'auteur concède à l'Université du Québec à Rimouski une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de son travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, l'auteur autorise l'Université du Québec à Rimouski à reproduire, diffuser, prêter, distribuer ou vendre des copies de son travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de la part de l'auteur à ses droits moraux ni à ses droits de propriété intellectuelle. Sauf entente contraire, l'auteur conserve la liberté de diffuser et de commercialiser ou non ce travail dont il possède un exemplaire.

“I will not say the day is done,
nor bid the stars farewell.” – J.R.R.
Tolkien

REMERCIEMENTS

Je me tiens de remercier d'abord Richard St-Louis et Marc-Denis Rioux, mes deux directeurs de maîtrise, qui m'ont encadré et guidé tout au long de mon parcours et du projet de recherche présenté dans mon mémoire. Leurs connaissances, leurs expertises, leurs conseils et leurs commentaires m'ont permis d'avancer et de m'épanouir dans ma recherche. Merci de votre patience et d'avoir cru en moi.

J'aimerais ensuite remercier toutes les personnes à l'Université du Québec à Rimouski qui m'ont porté aide, de prêt ou de loin. Merci, Laury-Ann Dumoulin, de m'avoir guidé au début de mon parcours et de m'avoir accompagné dans mes réflexions et mes premières expériences. Merci, Marie-Josée Ouellet, pour tes efforts sur le développement des méthodes en génétiques liées au projet, pour tes conseils dans le laboratoire et surtout pour ton aide lors des analyses de mes échantillons. Merci Steeven Ouellet de m'avoir accompagné dans mes analyses de chromatographie et de ton temps pour m'aider à résoudre les problèmes techniques.

Je remercie également Éric Gosselin et toute son équipe aux travaux publique de la ville de Rimouski pour l'échantillonnage rapide et efficace de l'eau usée de la ville.

Je remercie aussi CentrEau, le ministère de la Santé et des Services sociaux, le ministère de l'Environnement et des Changements climatiques et les scientifiques affiliés.es au projet pilote de vigie des eaux usées dont les efforts ont rendu ce projet possible. Sans eux, les connaissances et les subventions nécessaires pour cette étude ne seraient pas suffisantes.

Finalement, je veux remercier mes amis, ma famille et ma partenaire pour leur encouragement et leur soutien, morale et autre. Ce fut d'un grand aide et m'a permis d'en arriver à bout de mon parcours.

AVANT-PROPOS

La pandémie de la COVID-19 a eu plusieurs impacts négatifs à l'échelle mondiale. Cette crise majeure a favorisé une mobilisation collective sans précédent, en particulier au sein de la communauté scientifique, amenant une concentration significative des efforts dans le domaine de la recherche. Le présent projet s'inscrit dans ces efforts mondiaux visant à approfondir la compréhension des mécanismes de transmission du virus de la COVID-19 (SRAS-CoV-2) au sein des communautés et à identifier et optimiser les meilleures stratégies de gestion en contexte pandémique. Ces connaissances contribueront à renforcer notre capacité de réponse face aux crises sanitaires futures, nous permettant ainsi d'être mieux préparés à d'éventuelles pandémies.

Dès le début de la pandémie de la COVID-19, l'eau usée a rapidement été exploitée comme méthode de suivi du virus à travers les villes. Le potentiel des eaux usées comme outil de suivi de la progression du virus et de gestion a rapidement été observé et mis à l'épreuve. Plusieurs obstacles se sont présentés, tels que les méthodes d'analyses, les interprétations, l'accès et le partage des données, ainsi que l'utilisation optimale des données générées par l'eau usée. Avec l'effort d'innombrables groupes de recherche, la majorité de ces obstacles ont été surmontés. Le plus gros défi restant repose dans le lien entre les données d'eaux usées et les données cliniques dans la population ainsi que l'utilisation efficace des données d'eaux usées. Ce projet vise à répondre à ces besoins en développant une méthode pour mieux estimer les cas de COVID-19 dans une population à travers l'eau usée.

Quoique cette étude possède plusieurs limitations, notamment l'estimation des cas actifs de COVID-19 dans la population de Rimouski lors de la période de l'étude, les retombées sont toutes aussi importantes. Malgré le contexte et le cadre précis, soit une période relativement courte dans une seule ville, l'objectif est entre autres de démontrer l'efficacité de marqueurs populationnels universels non spécifiques au cadre du projet. La

validation adéquate est toujours nécessaire, mais les marqueurs choisis ont le même potentiel de succès à Rimouski que dans la majorité des autres régions capables de telles analyses.

RÉSUMÉ

L'épidémiologie basée sur les eaux usées (ÉBEU) peut être utilisée, entre autres, pour estimer les nouveaux cas cliniques de COVID-19 dans des villes ou des municipalités. Quoique cette approche comporte plusieurs avantages comparés au dépistage clinique, sa performance d'estimation dépend grandement des méthodes utilisées, particulièrement la méthode de normalisation des concentrations du SRAS-CoV-2 dans les eaux usées. Le virus de la marbrure légère du poivron (PMMoV) fut populaire comme biomarqueur pour normaliser les concentrations du SRAS-CoV-2, mais produit des résultats mitigés. Notre objectif est donc d'identifier et quantifier différents biomarqueurs pouvant avoir une meilleure performance dans la normalisation des concentrations du SRAS-CoV-2 dans les eaux usées.

Six critères ont été utilisés pour la sélection des biomarqueurs : la voie d'excrétion, la détectabilité, la stabilité, l'unicité, la variation spatio-temporelle et la méthode d'analyse. Les biomarqueurs sélectionnés sont le gène de la sous unité 5 du NADH déshydrogénase de l'ADN mitochondrial humain (mth-ND5), le gène CPQ_056 du virus bactériophage *Carjivirus communis* (CrAss), la caféine et l'acide hydroxyindoleacétique (5-HIAA). Ces biomarqueurs ont le potentiel d'avoir une meilleure performance que le PMMoV.

Les biomarqueurs ainsi que le SRAS-CoV-2 ont été analysés par réaction en polymérase en chaîne quantitative en temps réel et par chromatographie gazeuse couplée avec un spectromètre de masse. Les cas cliniques utilisés pour vérifier la performance des biomarqueurs ont été fournis par l'hôpital régional de Rimouski. En utilisant les concentrations de SRAS-CoV-2 dans les eaux usées, les biomarqueurs, le débit journalier de la station d'épuration d'eau de Rimouski et une estimation du taux d'excrétion virale, 11 modèles d'estimation de nouveaux cas actifs ont été créés avec des régressions linéaires. La performance des modèles est calculée et comparée en utilisant des corrélations de Spearman. Ces modèles démontrent que le CrAss, le mth-ND5, le 5-HIAA et le débit journalier sont tous capables de mieux normaliser les concentrations du SRAS-CoV-2 lorsque comparé au PMMoV.

L'utilisation de ces nouveaux biomarqueurs en ÉBEU mènerait à une estimation plus exacte des cas actifs de COVID-19 dans la population. Ceci aiderait les autorités de santé publique dans la prise de décision rapide et la bonne gestion de pandémies futures.

Mots clés : SRAS-CoV-2, COVID-19, Épidémiologie basée sur les eaux usées, Modélisation, RT-qPCR, Traceurs chimiques

ABSTRACT

Wastewater-based epidemiology (WBE) can be used, among other things, to estimate new active cases of COVID-19 in cities and smaller municipalities. While this approach does have many advantages compared to clinical testing, its performance in estimating cases varies greatly based on methods used, specifically which normalisation method is used to model the SARS-CoV-2 concentrations in wastewater. The pepper mild mottle virus (PMMoV) has been a popular biomarker to normalise SARS-CoV-2 concentrations, with mitigated success. Our objective is to identify and quantify different biomarkers with a more promising performance regarding normalising the SARS-CoV-2 concentrations in wastewater.

Six criteria were used for the selection of these biomarkers: excretion pathway, detectability, stability, uniqueness, spatial-temporal variability, and method of analysis. The chosen biomarkers are the gene for the human mitochondrial DNA sub-unit 5 of the NADH dehydrogenase (mth-ND5), the CPQ_056 gene from the bacteriophage *Carjivirus communis* (CrAss), caffeine, and hydroxyindoleacetic acid (5-HIAA). These biomarkers have the potential to outperform PMMoV.

The biomarkers and SARS-CoV-2 were analysed using real-time polymerase chain reaction (RT-qPCR) and gas chromatography in tandem with a mass spectrometer (GC-MS). Clinical cases used to verify the performance of the biomarkers were provided by the Rimouski regional hospital. Using the biomarker and SARS-CoV-2 concentrations, daily flowrate from the Rimouski wastewater treatment plant, and a basic model to estimate viral shedding rate, 11 models were constructed using linear regressions. The performance of each model was checked using Spearman correlations and compared between each other. These models demonstrate that CrAss, mth-ND5, 5-HIAA, and daily flowrate all outperform PMMoV in normalising the SARS-CoV-2 viral concentrations in wastewater.

The use of these new biomarkers in WBE would lead to a more accurate estimate of new active cases of COVID-19 in the population. This would help health authorities in their timely decision making and management of all present and future pandemics.

Keywords: SARS-CoV-2, COVID-19, Wastewater-based epidemiology, Modelling, RT-qPCR, Chemical markers

TABLE DES MATIÈRES

REMERCIEMENTS.....	v
AVANT-PROPOS	vii
RÉSUMÉ	ix
ABSTRACT.....	xi
TABLE DES MATIÈRES	xiii
LISTE DES TABLEAUX	xvi
LISTE DES FIGURES	xvii
LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES.....	xix
INTRODUCTION GÉNÉRALE	1
CHAPITRE 1 IDENTIFICATION DES MARQUEURS POPULATIONNELS	22
1.1 BUT ET FONCTIONNEMENT DES MARQUEURS POPULATIONNELS	22
1.2 CRITÈRES DE SÉLECTION	28
1.2.1 L'excrétion par l'urine ou les fèces	29
1.2.2 La détectabilité dans les eaux usées.....	29
1.2.3 La stabilité dans les eaux usées.....	30
1.2.4 L'unicité de la provenance humaine	30
1.2.5 La variabilité temporelle, spatiale et individuelle.....	31
1.2.6 La facilité et rapidité d'analyse.....	32
1.2.7 La pharmacocinétique	33
1.3 IDENTIFICATION DES NOUVEAUX MARQUEURS POPULATIONNELS	34
1.3.1 Matériel génétique	34
1.3.2 Bactéries et parasites.....	36

1.3.3	Édulcorants	37
1.3.4	Drogues	39
1.3.5	Métabolites naturels	40
1.3.6	Produits pharmaceutiques et produits de soins personnels	42
1.3.7	Polluants et substances toxiques.....	44
1.3.8	Autres marqueurs	45
1.4	SÉLECTION	46
CHAPITRE 2 Normalisation du SRAS-CoV-2 dans les eaux usées		51
2.1	RÉSUMÉ EN FRANÇAIS	51
2.2	BIOMARKERS AS A NORMALISATION TOOL FOR WASTEWATER-BASED EPIDEMIOLOGY TO CORRELATE WITH COVID-19 CASE DATA	53
2.3	ABSTRACT	53
2.4	INTRODUCTION	54
2.5	MATERIALS AND METHODS.....	56
2.5.1	Biomarker selection process.....	56
2.5.2	Sampling and treatment.....	58
2.5.3	Quantification.....	59
2.5.4	Statistical analysis and modelling	61
2.6	RESULTS	62
2.6.1	Marker concentrations	62
2.6.2	Case Data.....	64
2.6.3	Modelling	64
2.7	DISCUSSION	68
2.8	CONCLUSION	74
CONCLUSION GÉNÉRALE		76
ANNEXES		80
RÉFÉRENCES BIBLIOGRAPHIQUES		85

LISTE DES TABLEAUX

Table 1 List of potential biomarkers considered, and their respect of the criteria used.....	57
Table 2 Genetic biomarker target sequences.....	60

LISTE DES FIGURES

Figure 1. Structure générale du virion du SRAS-CoV-2	8
Figure 2. Structure du spicule de glycoprotéine du SRAS-CoV-2, modifié à partir de PBD:6VSD (Wrapp <i>et al.</i> , 2020). Les trois protéines formant l'homotrimère sont représentées avec des couleurs différentes : vert, mauve et orange. Les composantes encadrées sont identifiées uniquement sur la structure en vert. La portion rouge au bas représente la membrane du virion du SRAS-CoV-2.	10
Figure 3. Cycle de réplication du SRAS-CoV-2. Modifiée de BioRender.com (2024). http://app.biorender.com/biorender-templates	12
Figure 4. Exemple d'une STEP et ses étapes principales. Modifié de Genesis Water Technologies (2024). genesiswatertech.com	16
Figure 5. Schéma de système d'égout. Modifié de USAID (1982).	18
Figure 6. Principe de la normalisation basée sur un marqueur populationnel. [SRAS] : Concentration du SRAS-CoV-2 dans les eaux usées. [SRAS normalisé] : Concentration du SRAS-CoV-2 normalisé dans les eaux usées. [marqueur] : Concentration d'un marqueur populationnel dans les eaux usées.	24
Figure 7. Exemples de relations et interprétations entre les concentrations d'un virus cible (noir) et d'un marqueur populationnel (bleu). A. Une situation idéale, le marqueur est stable dans le temps et permet de bien voir les variations dans la concentration du virus tout en permettant la normalisation adéquate des données. B. Une situation problématique, le marqueur varie grandement dans le temps alors que le virus est stable, ce qui augmenterait grandement la variabilité de la normalisation et pourrait produire des données qui sont non-représentative de la situation épidémiologique de cette région à ce moment donné. C. Une situation particulière, où le marqueur et le virus ont les mêmes variations de concentrations, démontrant soit un mauvais marqueur populationnel (qui varient trop dans le temps) ou des facteurs externes hors du commun affectant les deux concentrations simultanément <i>e.g.</i> des variations dans le nombre de personnes dans la population contribuant. D. Une situation médiocre, le marqueur est	

plus ou moins stable dans le temps et permet de voir les variations dans la concentration du virus, mais pourrait négativement affecter les données lors de la normalisation..... 26

Figure 8. Diagramme des principaux critères de sélection pour les marqueurs populationnels (biomarqueurs)..... 29

Figure 9. SARS-CoV-2 N2 gene and biomarker concentrations (A-F1) and daily loads over time (A-F2). A1-2: SARS-CoV-2 N2 gene concentrations and daily load. B1-2: CrAss concentrations and daily load. C1-2: mth-ND5 concentrations and daily load. D1-2: PMMoV concentrations and daily load. E1-2: Caffeine concentrations and daily load. F1-2: 5-HIAA concentrations and daily load. 63

Figure 10. Case data from the study period. Dotted line: HR, number of patients who tested positive for COVID-19 in the Rimouski Regional Hospital per day. Line: HROS, shedding rate modelling based on the number of patients who tested positive per day. 64

Figure 11. Spearman Correlation Matrix for all variables used for modelling, including case data. HR (Hospital Reported cases), HROS (Hospital Reported cases with Shedding dynamics), Flow (daily flowrate from the WWTP), N2 (SARS-CoV-2 N2 gene target). Significant correlations are noted with * <0.05, ** <0.01, and *** <0.001. 65

Figure 12. Spearman Correlation Matrix for all models and case data sets. Significant correlations are noted with * (<0.05), ** (<0.01), and *** (<0.001)..... 67

Figure 13. WBE Models with Case Data (HROS) over time. Full lines are models excluding Flow; dotted lines are models including Flow. Colours are based on the biomarkers used. A: N2, B: N2_F, C: N2_PMMoV, D: N2_F_PMMoV, E: N2_CrAss, F: N2_F_CrAss, G: N2_CAF, H: N2_F_CAF, I: N2_mthND5, J: N2_F_mthND5, K: N2_HIAA, L: N2_F_HIAA. 68

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

5-HIAA	Acide 5-Hydroxyindole-3-acétique
ACE2	<i>Angiotensine Conversion Enzyme II</i> , enzyme de conversion de l'angiotensine II
ADN	Acide désoxyribonucléique
ADNmth	ADN mitochondrial humain
ARN	Acide ribonucléique
BRSV	<i>Bovine Respiratory Syncytial Virus</i> , virus respiratoire syncytial bovin
BSTFA	N,O-Bis(trimethylsilyl)trifluoroacétamide
CAF	Caféine
CISSS	Centres intégrés de santé et de services sociaux
COVID-19	<i>Coronavirus Disease 19</i> , maladie à coronavirus 2019
CrAss	<i>Cross-Assembly Phage</i> , CrAssphage
Ct	<i>Cycle Threshold</i> , seuil de cycle
CTD	<i>C-terminal Domain</i> , domaine C terminal
ddPCR	<i>Digital Droplet PCR</i> , PCR digitale à gouttelettes
dPCR	<i>Digital PCR</i> , PCR digitale
ÉBEU	Épidémiologie basée sur les eaux usées (<i>WBE : Wastewater-based epidemiology</i>)

gc	<i>Gene copies</i> , Nombre de copies de gène
GC-MS	<i>Gas Chromatography tandem Mass Spectrometry</i> , chromatographie gazeuse en tandem avec un spectromètre de masse
HPLC	<i>High Performance Liquid Chromatography</i> , chromatographie liquide haute performance
HR1	<i>Heptad Repeat 1</i> , sous-domaine répétitif heptade 1
HR2	<i>Heptad Repeat 2</i> , sous-domaine répétitif heptade 2
HR	Variable de cas de COVID-19 actif dans l'hôpital régional de Rimouski
HROS	Variable ajusté de cas de COVID-19 actif dans l'hôpital régional de Rimouski
INSPQ	Institut national de la santé publique du Québec
IPC	<i>Internal Positive Control</i> , contrôle interne positif
LC-MS	<i>Liquid Chromatography tandem Mass Spectrometry</i> , chromatographie liquide en tandem avec un spectromètre de masse
LOD	<i>Limit of Detection</i> , limite de détection
LOQ	<i>Limit of Quantification</i> , limite de quantification
MeOH	Méthanol
MPS	Matière particulaire en suspension
Mth-ND5	Gène de la sous unité 5 du NADH déshydrogénase de l'ADNmth
NTD	<i>N-terminal Domain</i> , domaine N-terminal
PCI	Phénol-chloroforme-isoamyle

PCP	<i>Personal Care products</i> , produits de soins personnels
PCR	<i>Polymerase Chain Reaction</i> , réaction de polymérase en chaîne
PMMoV	<i>Pepper mild mottle virus</i> , virus de la marbrure légère du poivron
qPCR	<i>Real-time quantitative PCR</i> , PCR quantitative en temps réel
RT-qPCR	<i>Reverse Transcription qPCR</i> , qPCR par transcription inverse
Rxn	Réaction
SPE	<i>Solid Phase Extraction</i> , extraction par phase solide
SRAS-CoV-2	Syndrome Respiratoire Aigüe Sévère Coronavirus 2 (<i>SARS-CoV-2 : Severe acute respiratory syndrome coronavirus 2</i>)
STEP	Station d'épuration d'eau (<i>WWTP: wastewater treatment plant</i>)
TMCS	Chlorotriméthylsilane
TMS	Triméthylsilyl

INTRODUCTION GÉNÉRALE

Lors de la pandémie de la COVID-19, plusieurs opportunités de recherche ont été créées, particulièrement dans le domaine de l'épidémiologie basée sur les eaux usées (ÉBEU) (Hrudey *et al.*, 2020). Le besoin urgent d'être en mesure d'utiliser une approche de surveillance rapide, à grande échelle et peu dispendieuse a propulsé l'ÉBEU au front de la recherche liée à la COVID-19 dès 2020 (Bivins *et al.*, 2020 ; Hrudey *et al.*, 2020). L'utilisation des eaux usées pour la détection de pathogènes humains n'est pas nouvelle ; le poliovirus, virus responsable de la poliomyélite, a été détecté dans les eaux usées de la Caroline du Nord en 1939 (Paul *et al.*, 1939). Dans les dernières décennies, avec la réduction des coûts associés aux techniques moléculaires avancées et la meilleure accessibilité de celles-ci, l'eau usée est utilisée au-delà des pathogènes humains (Singer *et al.* 2023). Les drogues illicites sont rapidement devenues une cible d'intérêt via l'ÉBEU par plusieurs groupes de recherches pour suivre et quantifier la consommation de substances dangereuses dans les villes (Palardy *et al.*, 2016 ; Singer *et al.*, 2023 ; Verenitch *et al.*, 2006 ; Gao *et al.*, 2022). Actuellement, plusieurs méthodes qualitatives et quantitatives sont utilisées en ÉBEU pour détecter, mesurer et suivre une multitude de cibles d'intérêts telles que des virus, des bactéries, du matériel génétique, des substances illicites, des produits pharmaceutiques, des métabolites, des polluants ou autres substances chimiques.

La popularité récente de cette approche dans le contexte de la pandémie de la COVID-19 provient de ses multiples avantages lorsque comparée au dépistage clinique typiquement utilisé par les autorités de santé publique (Hrudey *et al.*, 2020 ; Patey *et al.*, 2021). L'avantage principal de l'ÉBEU est la vitesse de production de données. D'abord, les personnes infectées par le SRAS-CoV-2 excrètent des particules virales du virus indépendamment des symptômes respiratoires (Patey *et al.*, 2021). Cela signifie qu'une personne infectée peut excréter des particules virales avant de reconnaître les symptômes de la COVID-19 ou même avant d'avoir des symptômes. Les personnes asymptomatiques excrètent aussi des particules

virales du SRAS-CoV-2 (Patey *et al.*, 2021). C'est-à-dire, les particules virales vont se retrouver dans le système d'eau usée avant d'être détectées ou mesurées par les autorités de santé publique lors du dépistage clinique. Ceci permet d'obtenir des données quantitatives d'eau usée qui peuvent devancer les données de dépistage clinique en moyenne de quelques jours à un maximum de 14 jours (Hutchison *et al.*, 2022 ; Daughton *et al.*, 2020 ; D'Aoust *et al.*, 2021 ; Kumar *et al.*, 2021). Un autre facteur contribuant à la vitesse de production de données est la rapidité des analyses dans le laboratoire. Puisque les échantillons d'eau usée sont déjà à l'échelle populationnelle, soit d'une ville ou d'une section d'une ville, moins d'échantillons sont nécessaires pour avoir un portrait global que lors du dépistage clinique. Ceci réduit le temps nécessaire pour traiter les échantillons et obtenir les données liées au virus de la COVID-19. De plus, le traitement des échantillons d'eau usée peut être fait en une seule journée par une seule personne. Donc, les données d'eau usée peuvent être produites, partagées et utilisées la même journée que l'échantillon est récolté et transporté au laboratoire. Le tout démontre le potentiel immense de l'ÉBEU et de l'eau usée en tant que système d'alerte précoce dans un contexte de pandémie. Plus les données sont produites rapidement, plus cela donne du temps aux autorités de santé publique à se préparer, réagir, traiter des patients et assurer une bonne gestion d'une pandémie.

Mise à part la vitesse de production de données, un autre avantage significatif de l'ÉBEU par rapport au dépistage clinique est la réduction des ressources nécessaires, soit humaines, matérielles et financières. Premièrement, le traitement de l'eau usée peut être effectué par une seule personne qualifiée, et ce pour plusieurs villes ou municipalités dans la même journée. Au contraire, le dépistage clinique nécessite plusieurs employés à tous les points de dépistage, autant les points de prélèvement que les centres d'analyses. Pour une même population, que ce soit une ville entière, une portion d'une ville ou une municipalité, les ressources humaines nécessaires pour l'analyse de l'eau usée est significativement plus faible, même en considérant le personnel des stations d'épuration d'eau qui coordonne et gère l'échantillonnage de l'eau usée. Deuxièmement, l'ÉBEU requiert un seul centre d'analyse pour une ville où les échantillons sont traités. En comparaison, le dépistage clinique requiert souvent plusieurs centres, chacun ayant leur propre équipement et leur propre espace

de travail, augmentant le nombre d'appareils nécessaire à un dépistage efficace. Aussi, l'ÉBEU utilise moins d'équipement de protection personnel et moins de matériel à utilisation unique, contribuant à la grande réduction de ressources matérielles nécessaires pour l'analyse de l'eau usée comparée au dépistage clinique. Troisièmement, selon le ministère de la Santé et des Services sociaux du Québec (MSSS), les coûts d'une journée de dépistage clinique peuvent excéder un million de dollars dépendant du nombre de tests effectués à travers la province. Pour une ville de petite taille, comme Rimouski, les coûts peuvent varier entre 1000 \$ et 30 000 \$ par jour pour subvenir aux besoins du dépistage clinique, basé sur les données de l'INSPQ. Cependant, l'ÉBEU, peu importe le nombre de personnes contribuant au réseau d'eau usée choisi, peut avoir des coûts d'opération inférieurs à 500 \$ par jour d'analyse d'une STEP. Ceci est significativement plus bas que les coûts du dépistage clinique, ce qui permettrait de rediriger des fonds vers les travailleuses et travailleurs de la santé ainsi que les patients en besoin.

D'autres avantages importants de l'ÉBEU proviennent de la nature même de l'eau usée et l'échantillonnage. Notamment, cette approche est non invasive, représentative et confidentielle. Aucun écouvillonnage oronasopharyngé, ni de type de prélèvement similaire, n'est nécessaire à l'échantillonnage. Cela fait en sorte que l'ÉBEU ne dépend pas d'une bonne technique d'écouvillonnage effectuée par les travailleurs et travailleuses de la santé. De plus, cela évite une étape du dépistage clinique qui est souvent perçue comme désagréable et parfois douloureuse, menant à moins de personnes volontaires à se porter au dépistage clinique (INSPQ, 2023). Donc, l'exactitude et la représentativité des données d'ÉBEU ne dépendent pas de la participation volontaire de la population à se soumettre à ce type de procédure. Au contraire, toutes personnes rattachées à la STEP contribuent aux échantillons d'eau usée. Cette contribution est systématique, qu'une personne soit infectée ou non, symptomatique ou non. C'est-à-dire que les échantillons d'eau usée sont plus représentatifs de la population que le dépistage clinique (Polo *et al.*, 2020). Les données d'eau usée ne discriminent pas basé sur les symptômes, sur la situation sociale ou financière, sur l'horaire d'une personne à se présenter à un centre de dépistage, sur les opinions ou alignements politiques ou autre facteur affectant le dépistage clinique. Ceci réduit le biais associé aux

données d'ÉBEU par rapport au dépistage clinique, permettant une meilleure représentativité des personnes rattachées à la STEP échantillonnée. Aussi, les données tirées de l'ÉBEU ne sont pas associées à une personne ou un groupe ; elles sont uniquement retraçables à la STEP d'où provient l'échantillon. Ceci évite le besoin de protocole de confidentialité de données qui entoure normalement les données de dépistage clinique. Donc, les données d'ÉBEU sont plus facilement publiables, partageables, et utilisables par une multitude de groupe de recherche ou d'autorités de santé dans le but commun de la meilleure gestion d'une pandémie.

Malheureusement, l'ÉBEU n'est pas une approche parfaite. Il existe plusieurs défis et inconvénients à son utilisation, surtout dans le contexte d'une pandémie. En premier, l'échantillonnage doit être fait de manière à empêcher le plus possible la dégradation des différents composés dans l'eau usée et assurer une bonne représentativité. Évidemment, l'échantillonnage de l'eau usée doit être fait en amont du traitement de l'eau usée (D'Aoust *et al.*, 2021). Ceci peut être fait à différent niveau du système pour avoir un échantillon qui représente une portion plus ou moins grande de la STEP. Il existe deux types d'échantillonnage couramment utilisés : composite et instantanés (« grab ») (Ahmed *et al.*, 2021 ; D'Aoust *et al.*, 2021). Les échantillons instantanés sont rapides et faciles à récolter puisqu'ils sont constitués uniquement d'un volume donné pris à un endroit à un moment précis. Toutefois, ils donnent un portrait des eaux usées uniquement pour ce moment précis dans le temps, ce qui peut apporter un certain biais dans les données produites dépendant de l'heure exacte de l'échantillonnage (Ahmed *et al.*, 2021). Les échantillons composites sont une combinaison de plusieurs petits volumes pris sur une certaine période, souvent 24 heures, puis recombinaison ensemble pour former un seul plus gros volume. Ce type d'échantillonnage est plus représentatif et permet d'obtenir une meilleure idée de l'état du réseau sur la période déterminée en agissant comme une moyenne des conditions de l'eau usée (Ahmed *et al.*, 2021). Ce dernier est souvent plus utilisé avec l'eau usée due à la nature changeante des éléments d'intérêts qui sont analysés. Toutefois, ce type d'échantillonnage requiert davantage de matériel, de temps et surtout de coordination avec les personnes responsables de la STEP qui ont la charge de la manipulation de l'eau usée lors de l'échantillonnage. Ce sont tous des

facteurs à prendre en compte avant d'entreprendre un projet de recherche portant sur l'analyse de l'eau usée.

En deuxième, sans une méthode robuste d'analyse en laboratoire, l'eau usée est peu utile en contexte de pandémie. Le traitement des échantillons au laboratoire nécessite plusieurs contrôles pour assurer la qualité, l'exactitude et la précision des données produites. D'abord, il est nécessaire de concentrer l'eau usée. Plusieurs techniques différentes peuvent être employées telles que l'ultrafiltration, la filtration sur membrane électronégative, la précipitation par polyéthylène glycol, la floculation par lait écrémé, l'utilisation de microbilles magnétiques, l'ultracentrifugation et d'autres (Patey *et al.*, 2021 ; Ahmed *et al.*, 2020 ; North *et al.*, 2023 ; Dehghan Banadaki *et al.*, 2024). Toutes ces méthodes possèdent leurs propres avantages et inconvénients. Plusieurs groupes de recherche se sont attardés sur la détermination de la meilleure méthode, mais la conclusion générale stipule que la meilleure méthode dépend de la situation et des ressources financières et temporelles (Patey *et al.*, 2021 ; Ahmed *et al.*, 2020 ; Abd-Elshafy *et al.*, 2022 ; Cutrupi *et al.*, 2023 ; North *et al.*, 2023 ; Dehghan Banadaki *et al.*, 2024). Suivant la concentration de l'eau usée, l'extraction et la purification sont les étapes qui suivent pour l'analyse du matériel génétique liée aux pathogènes tel que le SRAS-CoV-2. Plusieurs entreprises produisent et fournissent des trousseaux commerciales à petites et grandes échelles pour cette étape. Avec quelques différences notables dans les protocoles expressément fournis par les compagnies, le principe reste le même. La lyse cellulaire permet d'abord de relâcher le contenu intracellulaire, comme l'ADN et l'ARN. Ensuite, le lavage des inhibiteurs qui se retrouvent dans les eaux usées est effectué, ce qui augmente la stabilité du matériel génétique et permet une bonne quantification. Après, le lavage des sels de purification permet de purifier l'extrait, pour assurer qu'uniquement le matériel génétique y reste. Pour la quantification même, le choix est fait parmi les multiples techniques de qPCR, que ce soit la RT-qPCR, la dPCR ou bien la ddPCR (D'Aoust *et al.*, 2021 ; West *et al.*, 2022 ; Rainey *et al.*, 2023). Encore une fois, malgré les différences, ces trois techniques de PCR sont amplement adéquates dans un contexte d'ÉBEU (D'Aoust *et al.*, 2021 ; Rainey *et al.*, 2023). En tout, il y a une multitude d'étapes complexes au bon traitement de l'eau usée dans un contexte d'ÉBEU.

Heureusement, la littérature et la recherche sur les différentes approches, méthodes et techniques sont exhaustives et assez complètes pour assurer le bon déroulement d'un projet tel que celui-ci. Quoiqu'un défi, ceci n'est pas considéré comme une problématique.

En troisième, une fois produites, les données sont difficilement interprétables et utilisables telles quelles. Ici est où se trouve la problématique actuelle à laquelle nous voulons contribuer. Seules, les données de concentration d'un virus dans les eaux usées d'une ville ne sont pas une mesure d'infection à travers une ville. Elles sont simplement une mesure de présence de particules virales dans les eaux usées. Ces données doivent souvent être ajustées, modélisées et avant tout interprétées de façon à ne pas tirer des conclusions inexactes. Il existe une bonne corrélation entre les concentrations du SRAS-CoV-2 dans les eaux usées et l'indice des cas dans la population (Greenwald *et al.*, 2021 ; Polo *et al.*, 2020). Toutefois, cette corrélation est insuffisante pour estimer le nombre de cas dans la population sans ajustements (Greenwald *et al.*, 2021 ; Polo *et al.*, 2020). Une augmentation des cas dans la population ne mène pas automatiquement à une augmentation de la concentration du SRAS-CoV-2 dans les eaux usées et vice versa. Il existe plusieurs facteurs ayant un effet significatif sur la concentration du virus dans les eaux usées outre les cas dans une ville. Par exemple, s'il y a un gros événement social dans une ville, comme un festival ou une compétition sportive qui attirent un grand nombre de personnes n'habitant pas dans la ville, ceci peut augmenter de manière temporaire les concentrations du virus dans les eaux usées. Puisque ces personnes n'habitent pas dans la ville surveillée, les nouvelles personnes contribuant à la hausse de concentration dans les eaux usées ne seront pas nécessairement identifiées par le dépistage clinique de la ville. Ceci peut mener à une surestimation du nombre de cas dans la ville si les concentrations du virus sont utilisées seules. Autrement, s'il y a de fortes averses ou une fonte de neige et de glace rapide, cela peut grandement augmenter la quantité d'eau qui entre dans la STEP. Cela dilue le virus dans les eaux usées, menant à une plus faible concentration, sans réel changement dans le nombre de cas dans la ville. Pour ces raisons, il est important de regarder plus loin que simplement les concentrations du SRAS-CoV-2 dans les eaux usées lorsqu'on veut estimer le nombre de cas dans la population.

Il existe pour l'instant deux approches principales pour contourner ces entraves. D'abord, il est possible d'utiliser le taux d'excrétion des particules virales des personnes infectées pour calculer le nombre de cas positif dans la population (Wu *et al.*, 2020 ; Lowry *et al.*, 2023). Ceci est fait avec la charge virale du SRAS-CoV-2 dans les eaux usées, ce qui est calculé à l'aide du débit d'eau journalier de la STEP. Le débit journalier représente la quantité totale d'eau qui passe à travers la STEP. Avec cette métrique, nous pouvons calculer la charge virale à partir des concentrations du SRAS-CoV-2, la charge virale étant la quantité (en copies de gènes) du virus qui est présente dans les eaux usées. C'est-à-dire, au lieu d'avoir une donnée étant une quantité du virus par unité de volume, la charge virale est simplement la quantité du virus totale qui circule dans la STEP cette journée. Avec cette nouvelle donnée, l'estimation du nombre de cas à partir du taux d'excrétion des particules virales est rapide. Toutefois, le taux d'excrétion des particules virales du SRAS-CoV-2 n'est pas si facilement obtenu (Wu *et al.*, 2020 ; Lowry *et al.*, 2023). Il est particulièrement difficile à modéliser dû aux grandes différences entre les individus c.-à-d. une personne peut avoir un taux d'excrétion très élevé alors qu'une autre peut ne pas en excréter du tout (Wu *et al.*, 2020 ; Lowry *et al.*, 2023). Cette incertitude réduit l'efficacité de cette méthode pour relier les concentrations dans les eaux usées avec les cas dans la population. Pour cette raison, l'approche la plus répandue à ce jour est la deuxième, soit l'utilisation de marqueurs populationnels pour normaliser les concentrations du virus. Cette approche est expliquée plus en détail dans la section 1.1. C'est sur cette technique que nous basons le présent projet de recherche.

Ce mémoire de maîtrise présente en intégralité les méthodes, les résultats et les conclusions de la recherche portant sur le développement d'un outil de normalisation populationnel pour l'analyse du SRAS-CoV-2 dans les eaux usées à Rimouski pendant la période hivernale et printanière de 2022. Pour assurer la bonne mise en contexte et compréhension des aspects connexes importants au projet, les sections suivantes présentent et discutent des mécanismes d'infection et de transmissions du SRAS-CoV-2, les notions importantes des STEPs et finalement les objectifs de la présente recherche.

Infection et transmission du SRAS-CoV-2

Le SRAS-CoV-2 est un coronavirus appartenant à la même espèce que le SRAS-CoV-1 (ou simplement SRAS) lié à la pandémie de 2002-2004 (Zhou *et al.*, 2020). C'est un virus à enveloppe à ARN simple brin qui a été découvert en 2019 dans la région de Wuhan en Chine (WHO, 2020). Quoique soupçonné d'être d'origine zoonique, spécifiquement provenant des chauves-souris du genre *Rhinolophus*, le prédécesseur direct du SRAS-CoV-2 n'est pas encore identifié (Zhou *et al.*, 2020 ; Temmam *et al.*, 2022).

Le virion est formé de cinq composantes principales : l'ARN simple brin, les protéines membranaires, les spicules de glycoprotéines, les protéines de l'enveloppe et les protéines nucléocapsidiques (Figure 1) (Wu *et al.*, 2020). Chacune de ces composantes joue un rôle spécifique. L'ARN est la portion du virus qui est répliqué par la cellule hôte infectée permettant le transfert de toute l'information génétique du virus. Les protéines membranaires sont responsables de la bonne formation des virions lors de la réplication du virus et coordonnent le positionnement des autres protéines structurales. Les spicules de glycoprotéines sont les structures responsables du mécanisme d'infection du virus. Elles permettent de s'attacher aux parois cellulaires et d'infiltrer les cellules humaines pour amorcer l'infection et la réplication du virus. Les protéines de l'enveloppe aident à la

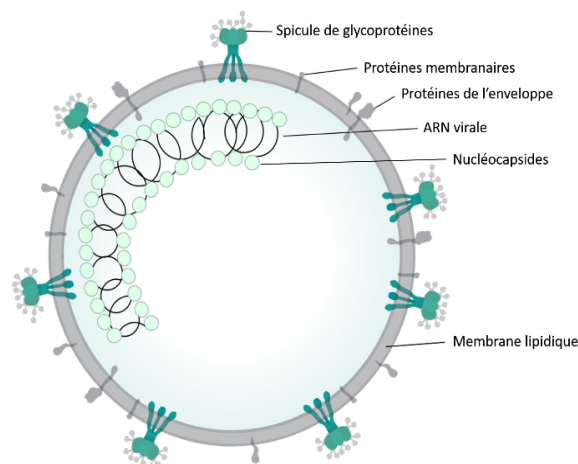


Figure 1. Structure générale du virion du SRAS-CoV-2

formation de l'enveloppe virale lors de la réplication, assurant une bonne structure de l'enveloppe et de la nucléocapside. Les protéines nucléocapsidiques forment la structure interne qui contient l'ARN viral au centre de l'enveloppe. Toutes ces composantes sont essentielles à un aspect du cycle d'infection du SRAS-CoV-2 et sans la bonne formation et le bon fonctionnement de toutes ces structures ensemble, le virus est soit incapable d'être répliquer, incapable d'infecter de nouvelles cellules ou est grandement entravé dans ces étapes.

De manière générale, un virus dépend d'une cellule hôte, comme une cellule humaine, pour être en mesure de se répliquer. Les virions ne sont pas des cellules, ils n'ont pas toutes les composantes nécessaires au bon fonctionnement de leur cycle de réplication pour l'effectuer de manière indépendante. Les virions doivent entrer dans une cellule hôte, puis utiliser les organelles de réplication d'ARN de la cellule hôte pour être en mesure de créer des copies du virus (Jackson *et al.*, 2022). Une fois l'ARN copié et les protéines formées, l'assemblage des nouveaux virions se fait à l'aide du réticulum endoplasmique et de l'appareil de Golgi (Jackson *et al.*, 2022).

Pour les coronavirus comme le SRAS-CoV-2, les spicules de glycoprotéines (Figure 2) sont des protéines homotrimère de fusion de classe I avec deux régions : S1 et S2 (Li, 2016). La région S1 est formée de deux sous-structures, le domaine N-terminal (NTD) et le domaine C-terminal (CTD) (Li, 2016). Le CTD est responsable de la liaison entre le virion et les enzymes de conversion de l'angiotensine 2 (ACE2) qui se retrouvent sur la face externe de certaines cellules humaines (Li, 2016). Une fois le virion lié entre le CTD et le ACE2, la fusion du virion à la membrane cellulaire débute. La région S2 est responsable d'assurer la fusion entre ces deux acteurs pour mener à la pénétration du virion dans la cellule hôte (Zhu *et al.*, 2021). La région S2 est formée du peptide de fusion ainsi que les sous-domaines répétitifs HR1 et HR2 (« heptad repeat »). Le peptide de fusion a pour but de pénétrer la membrane cellulaire et de perturber sa structure. Les sous-domaines HR1 et HR2 changent de conformation pendant le processus de fusion, en passant de structures d'hélices à un

faisceau de six hélices permettant de mettre en contact le peptide de fusion et la membrane cellulaire (Zhu *et al.*, 2021).

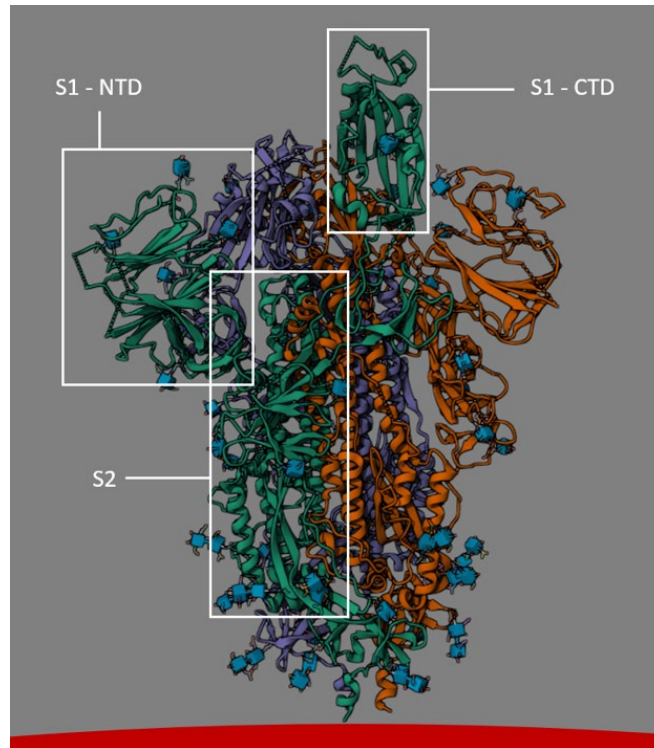


Figure 2. Structure du spicule de glycoprotéine du SRAS-CoV-2, modifié à partir de PBD:6VSD (Wrapp *et al.*, 2020). Les trois protéines formant l'homotrimère sont représentées avec des couleurs différentes : vert, mauve et orange. Les composantes encadrées sont identifiées uniquement sur la structure en vert. La portion rouge au bas représente la membrane du virion du SRAS-CoV-2.

Après la fusion, il y a deux voies possibles (Jackson *et al.*, 2021). Soit un endosome, une organelle intracellulaire, englobe le virion et l'emmène dans l'espace intracellulaire, un processus appelé endocytose. Le virion ensuite se libère de l'endosome et libère l'ARN viral dans le cytosol de la cellule hôte (Jackson *et al.*, 2021). Soit la région S2 du spicule de glycoprotéine est scindée par la protéase transmembranaire à sérine 2 (TMPRSS2), ce qui permet au virion d'injecter l'ARN viral directement dans le cytosol à partir de la membrane cellulaire (Jackson *et al.*, 2021). Dans les deux cas, l'ARN viral se retrouve dans le cytosol

de la cellule hôte et peut commencer à être répliqué par la machinerie de la cellule hôte (Figure 3). La réplication peut se produire à deux endroits différents : soit directement dans le cytosol à l'aide d'un ribosome ou à l'intérieur du réticulum endoplasmique (Jackson *et al.*, 2021). La réplication implique les mêmes étapes peu importe l'endroit précis où elle se déroule. D'abord, le brin d'ARN viral positif est traduit et ensuite produit une polymérase d'ARN qui est unique au virus, c'est-à-dire qui réplique uniquement l'ARN virale, assurant une réplication exclusive et efficace (Jackson *et al.*, 2021). Ensuite, la polymérase d'ARN procède à la réplication d'un brin d'ARN viral négatif qui est utilisé pour créer des copies du brin d'ARN positif original (Jackson *et al.*, 2021). Ces copies sont ensuite utilisées pour la production des protéines structurales du virion (Jackson *et al.*, 2021). La protéine nucléocapsidique se lie au génome viral et la protéine membranaire utilise la membrane du réticulum endoplasmique ainsi que le compartiment intermédiaire du ER-Golgi (ERGIC) pour commencer la formation d'un nouveau virion avec les autres protéines structurales (Jackson *et al.*, 2021). Depuis l'appareil de Golgi, le nouveau virion est envoyé vers la membrane cellulaire par des vésicules pour être expulsé de la cellule ; ceci est l'exocytose (Jackson *et al.*, 2021). Une fois à l'extérieur, le nouveau virion peut voyager vers d'autres cellules pour les infecter pendant que la cellule hôte de base continue de produire des virions.

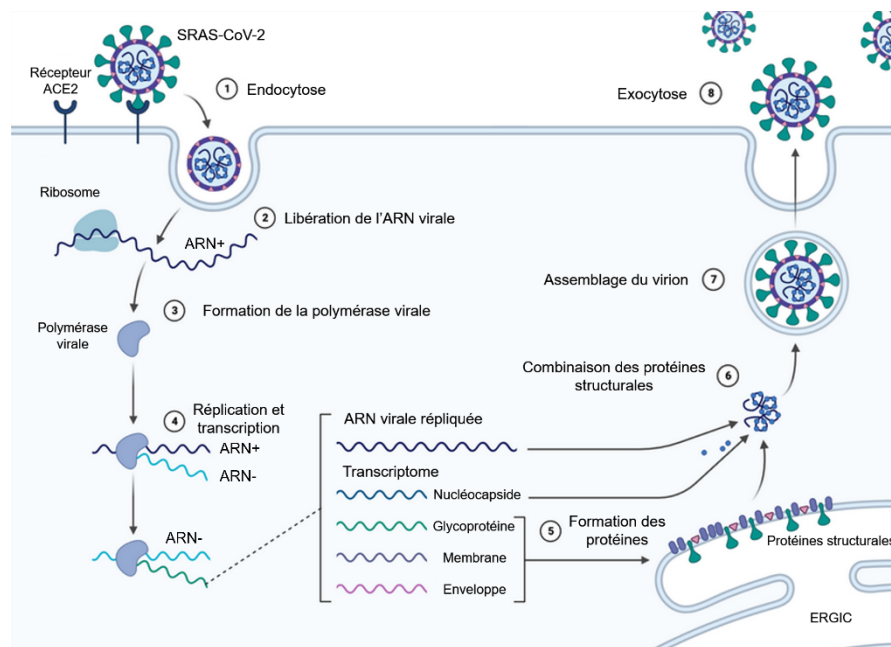


Figure 3. Cycle de réplication du SRAS-CoV-2. Modifiée de BioRender.com (2024). <http://app.biorender.com/biorender-templates>

Les récepteurs ACE2 jouent un rôle important dans le corps humain. Ces récepteurs sont responsables de l'homéostasie et de la régulation de l'hormone angiotensine II (Turner *et al.*, 2015). L'angiotensine II est une hormone dans les voies métaboliques liées au système cardiovasculaire et endocrinien qui, entre autres, régule la pression vasculaire avec ses précurseurs et ses métabolites (Turner *et al.*, 2015). Les récepteurs ACE2 se retrouvent sur les cellules endothéliales de plusieurs tissus différents, majoritairement le petit intestin, les reins, les testicules, les ovaires, le cœur, les cellules adipeuses et la thyroïde (Hamming *et al.*, 2004 ; Li *et al.*, 2020). Toutefois, ce récepteur est aussi exprimé dans les cellules du système respiratoire en petite quantité, notamment les poumons et l'œsophage (Li *et al.*, 2020). La distribution du récepteur ACE2 peut expliquer la pathologie du virus *i.e.* ses voies de transmission (orales et respiratoires) et l'expression des symptômes chez les personnes infectées (Khreefa *et al.*, 2023 ; Li *et al.*, 2020). La forte présence du récepteur ACE2 dans le système digestif est la raison pour laquelle il est possible de détecter et subséquemment quantifier les particules virales du SRAS-CoV-2 dans les selles humaines (Khreefa *et al.*, 2023). Puisque la finalité du processus de réplication du virus est de produire une grande

quantité de virions pour infecter d'autres cellules, plusieurs virions sont relâchés dans le système digestif directement (Khreefa *et al.*, 2023). Donc, à chaque évacuation intestinale, s'il y a présence du SRAS-CoV-2 dans le système digestif dû à la forte présence de récepteur ACE2, plusieurs particules virales sont évacuées simultanément aux selles (Patey *et al.*, 2021 ; Khreefa *et al.*, 2023 ; Lowry *et al.*, 2023).

Les concentrations de particules virales excrétées par les personnes infectées sont méconnues et variables (Lowry *et al.*, 2023 ; Soller *et al.*, 2022). De plus, quoique la majorité des particules virales qui sont acheminées dans le système d'eau usée provient des selles, il y a plusieurs sources autres possibles : le mucus, la salive, les expectorations et l'urine (Zhang *et al.*, 2020 ; Patey *et al.*, 2021 ; Lowry *et al.*, 2023 ; Soller *et al.*, 2022). Toutefois, l'excrétion des particules virales par les fèces est faite en plus grande quantité et perdure bien plus que l'expulsion par les autres voies (Zhang *et al.*, 2020 ; Patey *et al.*, 2021). Cela souligne non seulement l'importance de l'ÉBEU comme outil de surveillance et de gestion d'une pandémie, mais aussi le risque de la transmission par voie fécale-orale.

L'excrétion de particules virale du SRAS-CoV-2 par la matière fécale n'est pas constante toute au long de l'infection et des stades de la maladie (Zhang *et al.*, 2020 ; Rioux *et al.*, 2023). La dynamique d'excrétion virale (« viral shedding dynamics ») module la quantité de particules virales excrétée tout au long de l'infection. Plusieurs facteurs affectent la dynamique d'excrétion virale dont l'âge de la personne infectée et la sévérité des symptômes, mais le facteur le plus important est le temps (Cevik *et al.*, 2020 ; Rioux *et al.*, 2023). Une personne infectée va excréter plus de particules virales au début de l'infection, lors de la période où le virus se réplique davantage. Ensuite, la quantité diminue avec le temps puisque le système immunitaire entre en jeu, réduisant la réplication du virus. Avec la diminution de la réplication et subséquemment les symptômes vient la réduction des particules virales excrétées. Toutefois, le tout est aussi observé chez les personnes asymptomatiques et les tendances sont les mêmes (Cevik *et al.*, 2020). Donc non seulement une personne infectée excrète des particules virales tout au long de la période d'infection, tout au long de la maladie, mais la quantité de particules virales est plus élevée au début et

diminue avec le temps. La dynamique d'excrétion virale est donc importante de prendre en compte dans les applications de modélisations qui utilisent en majeure partie des données d'eaux usées.

Le SRAS-CoV-2 possède plusieurs modes de transmission différents. D'abord, la voie de transmission la plus connue et la plus efficace pour la transmission du virus est la voie respiratoire, soit l'inhalation du virus lorsqu'il se retrouve en suspension dans l'air (Leo *et al.*, 2023). Le virus est souvent expulsé du corps par les personnes infectées par la salive, la toux, le mucus et les expectorations (Leo *et al.*, 2023). Les expulsions comme les éternuements produisent des aérosols et des gouttelettes de taille microscopique, soit d'environ moins de 5 μ m et de 5-10 μ m respectivement (Leo *et al.*, 2023). Ces particules contiennent une grande quantité de virions étant capables d'infecter une autre personne dans les alentours, pouvant aller à plusieurs mètres de distance pour les gouttelettes ou à l'échelle d'une pièce dans le cas des aérosols (Leo *et al.*, 2023). Les virions sont encore viables et capables d'infection jusqu'à 48 heures après l'expulsion initiale, ce qui rend le SRAS-CoV-2 très contagieux et souligne l'importance du port du couvre-visage et d'une bonne ventilation et filtration de l'air (Leo *et al.*, 2023).

Un autre mode de transmission courant est la transmission par les surfaces infectées (Leo *et al.*, 2023). Ces mêmes aérosols et gouttelettes se déposent éventuellement sur les surfaces telles que des meubles, les murs, les planchers, les poignées de portes, les divers objets aux alentours, etc. Puisque le virus est encore actif une longue période à l'extérieur du corps humain, le virus peut se transférer de surfaces à surfaces en utilisant différents vecteurs ; un tel exemple est une main. Si le virus se retrouve sur une surface, une personne vient en contact avec cette surface et par la suite touche une autre surface, cette deuxième surface est aussi contaminée avec le SRAS-CoV-2 (Leo *et al.*, 2023). De plus, une fois que ces surfaces contaminées sont touchées, un contact avec une partie du visage comme les yeux, la bouche ou le nez, peut facilement apporter le virus dans le corps et permettre le début de l'infection (Leo *et al.*, 2023). Les surfaces peuvent prendre la forme d'animaux de compagnie, de linge, ou d'autres objets qui sont transportés ou qui sont mobiles, augmentant

les zones possibles où peut se trouver le SRAS-CoV-2 et du même fait augmentant aussi le risque de transmission du virus. Ceci souligne l'importance de nettoyer les surfaces et de se laver les mains pour réduire le risque d'infection (Leo *et al.*, 2023).

Système d'eau usée et station de traitement d'épuration d'eau

Un système d'eau usée, ou système d'assainissement est défini comme étant un travail ou une structure, dont une partie est située sur terre, qui est utilisée pour la récolte et le dépôt d'eau usée, que cette eau usée soit traitée ou non, et qui inclut un étang d'eau usée (gouvernement du Canada, SOR/2012-139, 2012). Un tel système est typiquement composé d'une multitude de structures sur l'étendue d'une ville ou d'une municipalité tous portant un rôle spécifique. Ceci inclut le système d'égouts qui achemine l'eau des usagers à la station d'épuration, les bassins de collections d'eau usée, l'équipement de filtration, les bassins de sédimentations, les bassins d'aérations, l'équipement de désinfection et le système de rejet d'eau traitée (Figure 4).

Près de la moitié des systèmes d'eau usée au Canada ont deux composantes de traitement principales, soit le traitement primaire et le traitement secondaire (ECCC, 2023). Le traitement primaire consiste à la filtration de la matière solide grossière et les débris à l'aide d'écrans de filtration suivis des bassins de sédimentation pour éliminer les sédiments comme le sable et le gravier de l'eau usée (EPA, 1998). Ensuite, le traitement secondaire a pour but d'accélérer le processus naturel de décomposition de la matière organique par la communauté microbienne (EPA, 1998). Ceci est fait dans des bassins ou des étangs d'aération où se trouve la boue activée et l'apport en air qui aide à la dégradation de la matière présente dans l'eau usée. Certaines STEP's incluent aussi un traitement tertiaire étant la

désinfection de l'eau traitée. Celle-ci est souvent faite avec des produits chimiques tels le chlore ou par des processus physiques tels les rayons UV (EPA, 2003).

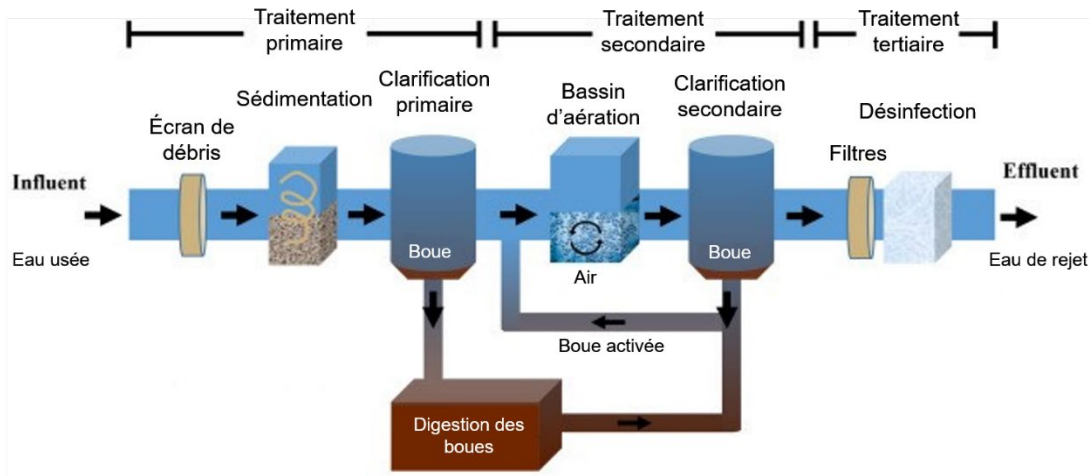


Figure 4. Exemple d'une STEP et ses étapes principales. Modifié de Genesis Water Technologies (2024). genesiswatertech.com

Au Canada, il existe deux configurations de système d'eau usée : un système séparé (fermé) et un système combiné (ouvert) (ECCC, 2023). Les systèmes séparés possèdent deux systèmes d'égouts distincts : un pour l'eau sanitaire et un pour l'eau pluviale. L'eau sanitaire est toute l'eau usée produite par les usagers, que cela soit d'une source résidentielle commerciale, industrielle ou institutionnelle (ECCC, 2023). Les eaux pluviales sont constituées de toutes les eaux naturelles qui sont récoltées par le système, soit provenant des précipitations, de la fonte de la neige et de la glace ou des débordements fluviaux et riverains. Dans les systèmes séparés, les eaux pluviales et sanitaires sont traitées séparément et ne sont jamais mélangées avant le rejet. Les systèmes combinés possèdent un seul système d'égout où s'achemine l'entièreté de l'eau usée, sanitaire et pluviale. Les deux types d'eaux usées sont traités simultanément dans un système combiné.

Dans un contexte d'ÉBEU, les systèmes séparés possèdent des avantages particuliers. D'abord, en empêchant l'eau pluviale de se mélanger aux eaux sanitaires, tous les composés

présents dans les eaux sanitaires ne sont pas dilués. Ceci facilite l'analyse et la quantification de ces composés puisqu'ils sont plus concentrés, rendant leur détection plus fiable. Pour de potentiels marqueurs populationnels qui sont trouvés en faible concentration, comme les produits pharmaceutiques ou les drogues, ceci peut être très bénéfique (Polo *et al.*, 2020). Un autre avantage des systèmes séparés est le débit constant à la STEP. La quantité d'eau pluviale varie grandement dans le temps. Dans un système combiné, cela peut créer une source de variation de plus dans la concentration des marqueurs dans les eaux usées analysées. Les systèmes séparés n'ont pas ce problème. Puisque les eaux pluviales n'affectent pas le débit des eaux sanitaires, les concentrations des composés sont beaucoup plus stables dans le temps. Lorsque l'ÉBEU est effectuée dans une ville ayant un système combiné, il est particulièrement important de prendre en compte le débit de la STEP puisque le débit variable peut avoir un grand impact sur les concentrations de pathogènes et des marqueurs analysés (Polo *et al.*, 2020).

L'échantillonnage de l'eau usée peut être effectué à plusieurs endroits différents au long du système d'eau usée. La zone d'échantillonnage de prédilection pour l'ÉBEU est l'influent précédant les bassins d'aérations (Kumar *et al.*, 2021 ; North *et al.*, 2023 ; Curupi *et al.*, 2023 ; Dehghan Banadaki *et al.*, 2024 ; Ahmed *et al.*, 2020b ; Greenwald *et al.*, 2021). Quoique l'échantillonnage des boues primaires est parfois fait également, il est moins commun (D'Aoust *et al.*, 2021 ; Boehm *et al.*, 2023 ; Graham *et al.*, 2020). Idéalement, pour ne pas compromettre l'intégrité du virus ou des marqueurs analysés, l'échantillonnage est fait avant le traitement, soit biologique ou chimique, de l'eau usée, donc avant le traitement secondaire à la STEP (D'Aoust *et al.*, 2021). L'échantillonnage directement à la STEP permet d'avoir un échantillon représentant l'entièreté de la population contribuant au système. Ceci est pratique pour avoir une vision globale de l'état de la ville dans le contexte d'ÉBEU. Toutefois, il est possible et intéressant d'échantillonner également en amont.

Les systèmes d'égouts sont typiquement divisés en secteurs différents, avec un réseau pouvant s'apparenter à un système de rivière (Figure 5). D'abord, l'eau usée est acheminée des différents bâtiments et infrastructures vers des conduits primaires qui se trouvent typiquement sous les rues les plus proches. Ensuite, ces conduits primaires se regroupent dans des conduits secondaires plus volumineux. Ces conduits secondaires contiennent l'eau usée d'une certaine zone d'une ville ou d'une municipalité. Les conduits secondaires se combinent en conduits tertiaires qui dirigent l'eau usée de la ville entière vers la STEP. Dans le cas de grandes métropoles, il existe souvent plusieurs STEPs, c'est-à-dire que les conduits tertiaires ne contiennent pas l'eau usée de la ville entière, mais plutôt d'une grande section de la métropole. Ce modèle pour le réseau d'égouts d'une ville permet d'isoler certaines zones d'intérêt lors de l'échantillonnage de l'eau usée pour l'ÉBEU. Par exemple, un certain quartier résidentiel ayant une plus grande proportion de personnes à risque ou de personnes âgées pourrait être surveillé de plus près que les autres secteurs de la même ville. Le même principe peut s'appliquer à un bâtiment également. Donc un hôpital ou une résidence pour personnes âgées peut être surveillé à travers l'eau usée plus facilement pour en faire une gestion de plus près. Ceci permet une gestion plus nuancée et adaptée aux besoins précis des personnes plus à risque ainsi que le reste de la population contribuant à la STEP.

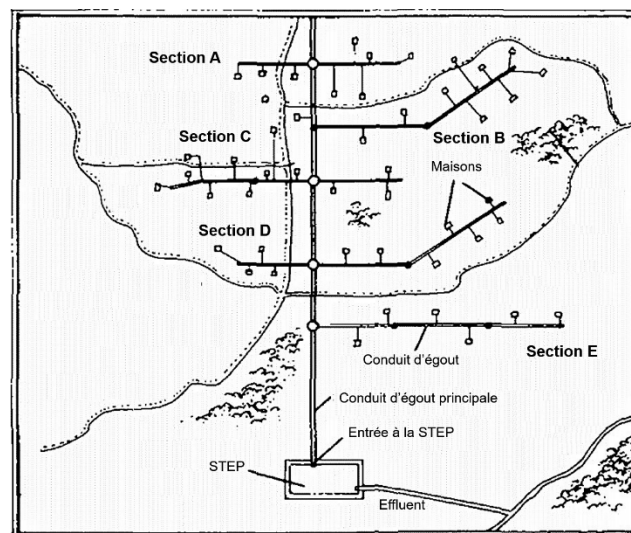


Figure 5. Schéma de système d'égout. Modifié de USAID (1982).

Objectifs de recherche

Le projet de recherche présent a pour but principal de développer un outil de normalisation populationnel pour l'analyse du SRAS-CoV-2 dans les eaux usées. Cet outil sera conçu à partir de la modélisation des concentrations du SRAS-CoV-2 dans les eaux usées avec d'autres facteurs tels la concentration de marqueurs populationnels et le débit. Ceci permettra d'augmenter la corrélation entre les concentrations du virus de la COVID-19 et le nombre de cas dans une ville choisie. Une forte corrélation entre ces données aiderait les autorités de santé publique dans la gestion de pandémie, présente et future, en ayant un outil efficace d'alerte précoce du virus. Un tel outil donnerait plus de temps pour la prise de décision des mesures sanitaires et la gestion du matériel et du personnel hospitalier par les gestionnaires appropriés. Le processus pour développer l'outil de normalisation populationnel se divise en trois objectifs spécifiques.

Le premier objectif spécifique est d'identifier un ou plusieurs marqueurs humains, de nature biologique, chimique ou autre, qui possèdent un potentiel d'agir en tant que marqueur populationnel pour normaliser les concentrations du SRAS-CoV-2 dans les eaux usées. Plusieurs facteurs peuvent influencer la performance d'un marqueur, lorsque comparés à d'autres

Le deuxième objectif spécifique est la quantification du SRAS-CoV-2 et des marqueurs populationnels dans les eaux usées de Rimouski. Le SRAS-CoV-2 et les marqueurs ont déjà été quantifiés dans les eaux usées avec succès à travers divers contextes et régions. Toutefois, pas tous n'ont été quantifiés dans les eaux usées de Rimouski spécifiquement; plusieurs facteurs externes tels que la démographie, le climat, le type d'échantillon et le système d'eau usée peuvent faire varier les mesures d'ÉBEU. Il est donc nécessaire d'assurer une quantification précise et exacte du SRAS-CoV-2 et des marqueurs populationnels dans les eaux usées de Rimouski spécifiquement.

Le troisième objectif spécifique est la création du modèle de normalisation populationnel à l'aide de R/RStudio et OpenSees. Le modèle a pour but de servir comme

outil de gestion aux autorités décisionnelles de la santé publique, autant au Québec qu'ailleurs, dans un contexte de pandémie en étant capable d'estimer rapidement, avec précision et précocité, le nombre de cas actif de COVID-19 dans une ville.

Méthodologie générale

Pour le premier objectif, le processus de sélection a été fait à partir de la littérature scientifique publiée non seulement dans un contexte de pandémie, mais dans un contexte d'analyse environnementale également. La sélection des critères sur lesquels l'évaluation des potentiels marqueurs populationnels est faite, les marqueurs populationnels considérés et la sélection des marqueurs populationnels sont tous décrits à la section 1.2, 1.3 et 1.4, respectivement, du CHAPITRE 1.

Pour le deuxième objectif, la quantification du SRAS-CoV-2 et des marqueurs populationnels est effectuée avec des méthodes d'analyses génétiques et chimiques selon le marqueur. Les échantillons d'eaux usées sont d'abord filtrés, puis la matière particulaire (sur le filtre) et les composés dissouts (filtrat) sont séparés et traités en parallèle. Le matériel génétique des marqueurs génétique et du SRAS-CoV-2 est extrait, purifié et concentré avant les analyses par qPCR. Les marqueurs chimiques sont extraits par phase solide, dérivés par silylation au besoin, puis analysés par GC-MS. Les méthodes détaillées sont décrites dans la section 2.5.2 et 2.5.3 du CHAPITRE 2.

Pour le troisième objectif, ayant plusieurs facteurs à inclure dans la modélisation des cas actifs, plusieurs modèles distincts sont créés. Pour chaque facteur, soit le débit et les marqueurs populationnels, il y aura un modèle simple. Il y aura aussi des modèles multiples, c'est-à-dire des modèles qui incluront plusieurs facteurs simultanément. Tous ces modèles seront comparés pour déterminer le plus performant. La performance des modèles est mesurée par la corrélation entre l'estimation des cas dans la population par le modèle et les données épidémiologiques de l'hôpital régional de Rimouski fournis par le CISSS et l'INSPQ. La méthode détaillée est décrite dans la section 2.5.4 du CHAPITRE 2.

Le plan expérimental et la méthodologie sont décrits dans la section 2.5 «Materials and Methods» au Chapitre 2. Elle se retrouve également sous forme de diagramme en annexe (ANNEXE III) pour une meilleure compréhension des différentes étapes et analyses qui sont faites en parallèle sur les mêmes échantillons d'eau usée.

CHAPITRE 1

IDENTIFICATION DES MARQUEURS POPULATIONNELS

1.1 BUT ET FONCTIONNEMENT DES MARQUEURS POPULATIONNELS

Premièrement, il est important de bien définir qu'est-ce qu'un marqueur populationnel. Un marqueur populationnel, dans un contexte global d'ÉBEU, est un composé, un produit, une molécule ou une donnée pouvant être mesuré qui représente soit un certain aspect d'une population ciblée ou qui peut être utilisé pour estimer le nombre de personnes dans la population ciblée (Gracia-Lor *et al.*, 2017). Donc, ceci inclut une grande quantité et diversité de substances pouvant être mesurées qui peuvent agir en tant que marqueur populationnel ; cela sera abordé dans la section 1.3.

Les marqueurs populationnels peuvent avoir plusieurs rôles, parfois simultanément. Un premier rôle possible est l'estimation de facteurs de style de vie ou d'utilisation d'une substance (Gracia-Lor *et al.*, 2017). Par exemple, il est possible d'analyser certaines drogues illicites, telles que la cocaïne, pour estimer leur patron et leur fréquence de consommation dans une certaine ville (Palardy *et al.*, 2016). Un deuxième rôle possible est l'estimation de l'exposition à des toxines ou des polluants provenant de l'environnement ou de la nourriture (Gracia-Lor *et al.*, 2017). Ceci pourrait être l'exposition à des pesticides ou à des filtres UV (Rousis *et al.*, 2016, 2017 ; Ramos *et al.*, 2016). Une troisième utilité des marqueurs est la santé publique (Gracia-Lor *et al.*, 2017). Le meilleur exemple récent est l'analyse du SRAS-CoV-2 dans le but de suivre le déplacement global du virus, de déterminer la prévalence et l'incidence puis d'utiliser cette information dans la prise de décision de gestion de la pandémie de la COVID-19. Le quatrième rôle possible est l'estimation de la taille d'une population (Gracia-Lor *et al.*, 2017). Les marqueurs populationnels, dans le contexte de la pandémie de la COVID-19, ont pour but de normaliser les concentrations et les données

d'ÉBEU pour ensuite en faire la comparaison avec d'autres régions ou pour faciliter la bonne prise de décision en ayant des données *per capita*. Pour ce projet-ci, les marqueurs populationnels ont été sélectionnés basés sur les critères leur permettant d'estimer la taille d'une population. Ces critères sont plus stricts et discriminants, assurant une meilleure sélection. Cependant, ils seront uniquement utilisés dans la normalisation des concentrations du SRAS-CoV-2. La sélection elle-même sera abordée plus en détail dans la prochaine section, soit la section 1.2.

D'abord, sur le plan des calculs, la normalisation est simple. Elle consiste à diviser les concentrations journalières du SRAS-CoV-2 avec les concentrations d'un marqueur populationnel pour avoir une nouvelle donnée : la concentration du SRAS-CoV-2 normalisée. Cette opération permet de balancer ou réduire les variations des concentrations du SRAS-CoV-2 et du marqueur populationnel qui sont dues à la STEP, ce qui laisse la concentration du SRAS-CoV-2 normalisée représentant davantage les variations dues aux changements dans la prévalence des cas de COVID-19. Comme illustré dans la Figure 6, la normalisation, lorsque bien faite, a comme effet d'augmenter la corrélation des données d'ÉBEU avec les cas cliniques. Toutefois, quoique cela n'est pas toujours nécessaire, la normalisation vient souvent améliorer la précision des données. Si la normalisation n'est pas effectuée proprement, elle vient simplement ajouter de la variation aux données, ce qui réduit leur qualité et rendent leur interprétation plus difficile. Dans une STEP ouverte, comme celle à Rimouski, c'est-à-dire lorsque le système inclut l'eau pluviale et sanitaire, la normalisation devient une étape nécessaire. Cependant, dans le cas spécifique d'une STEP fermé, qui inclurait uniquement l'eau utilisée par les divers bâtiments et usagers et qui ne possède pas de fuites ou de failles, la normalisation n'est pas strictement nécessaire pour avoir une bonne corrélation, quoiqu'elle puisse tout de même être bénéfique selon certains (Graham Construction and Engineering, 2023).

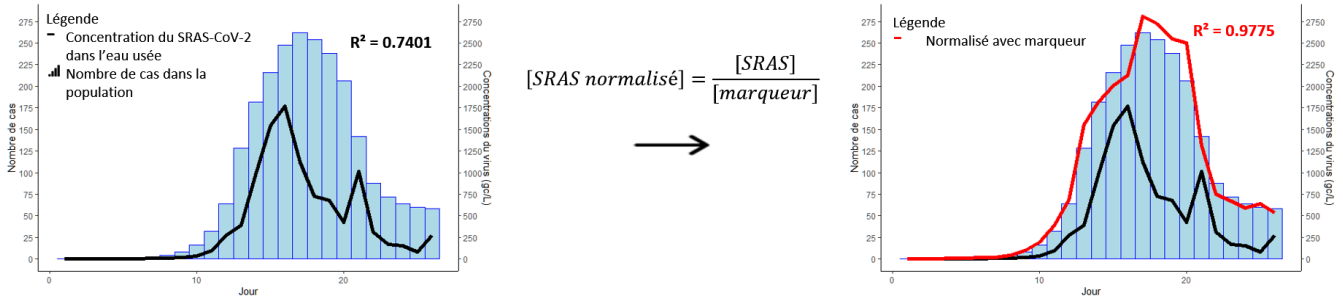


Figure 6. **Principe de la normalisation basée sur un marqueur populationnel.** [SRAS] : Concentration du SRAS-CoV-2 dans les eaux usées. [SRAS normalisé] : Concentration du SRAS-CoV-2 normalisé dans les eaux usées. [marqueur] : Concentration d'un marqueur populationnel dans les eaux usées.

Les données pharmacocinétiques sont une partie intégrante de l'ÉBEU. La pharmacocinétique décrit le cheminement métabolique des substances ingérées, et ainsi permet de connaître, pour une substance donnée, la proportion qui est excrétée en métabolites et celle excrétée directement ainsi de connaître le ou les métabolites qui sont excrétés. En connaissant les données de pharmacocinétique d'un certain marqueur populationnel, il est possible de rétrocalculer la quantité de personnes contribuant ce marqueur à une STEP à un moment précis, soit une journée particulière. Par exemple, pour calculer le nombre de personnes qui fument dans une ville, il faut connaître le métabolite principal de la nicotine, soit la cotinine, le taux d'excrétion moyen des personnes fumeuses, la concentration dans les eaux usées, et le débit journalier ou le volume d'eau traité en une journée par la STEP, le nombre de personnes fumeuses contribuant à la STEP est calculable facilement avec la formule générale suivante :

$$Nb_{pc} = \frac{[MP_{STEP}] \times V_j}{T_{ex}} \quad (1)$$

Où Nb_{pc} est le nombre de personnes contribuant à la STEP, $[MP_{STEP}]$ est la concentration du marqueur populationnel dans la STEP (g/L ou mg/mL), V_j est le volume journalier d'eau traitée par la STEP, parfois indiqué comme étant le débit (L/jour) et T_{ex} est le taux d'excrétion du marqueur populationnel (g/personne/jour). Ceci est un exemple relativement simplifié, mais la théorie est la même pour le fonctionnement des marqueurs populationnels de manière générale.

Que ce soit une STEP ouverte ou fermée, il y a plusieurs étapes, événements et transformations dans une STEP entre l'entrée d'une substance et la récolte des échantillons d'eau usée. La dilution, la dégradation bactérienne, la dégradation physico-chimique, le transport à travers la STEP, la sédimentation et la sorption sur certaines particules en sont tous des exemples qui doivent être pris en compte (Bayati *et al.*, 2022 ; Ye *et al.*, 2016 ; Hamouda *et al.*, 2021 ; Bivins *et al.*, 2020 ; Ahmed *et al.*, 2020a ; Wurtzer *et al.*, 2021). Toutefois, ces processus ne sont pas toujours facilement mesurables et nécessitent beaucoup d'effort et de temps consacré à la recherche de ces processus. Cela fait en sorte que les cibles primaires, comme les cibles d'ARN du SRAS-CoV-2, soient les gènes N1 et N2, auront une concentration variable dans le temps et dans l'espace à l'intérieur de la STEP (Bayati *et al.*, 2022 ; Ye *et al.*, 2016 ; Hamouda *et al.*, 2021 ; Bivins *et al.*, 2020 ; Ahmed *et al.*, 2020a ; Wurtzer *et al.*, 2021). Toutefois, cela est vrai aussi pour les marqueurs populationnels (Bayati *et al.*, 2022 ; Ye *et al.*, 2016 ; Ahmed *et al.*, 2020a). Lorsqu'un marqueur est bien choisi et bien mesuré, en théorie ce marqueur aura le même cheminement et les mêmes conséquences que la cible primaire. Donc si la variation à l'intérieur de la STEP est la même pour les deux, soit le marqueur populationnel (*e.g.* PMMoV) et la cible primaire (*e. g.* SRAS-CoV-2), toutes les pertes possibles et tous les changements dans la concentration de la cible primaire vont être pris en compte par le marqueur. Cela est ce qui permet de normaliser les concentrations avec un marqueur populationnel.

Assumant qu'un marqueur est parfaitement stable dans le temps, *i.e.* qui est excrété de manière constante dans une STEP pour une population donnée, les variations du SRAS-CoV-2 deviennent plus facilement interprétable. Si les concentrations du marqueur dans les eaux usées restent stables, mais que les concentrations du SRAS-CoV-2 dans les eaux usées varient, alors la variation est due à un changement dans le taux d'excrétion de particules virale du SRAS-CoV-2 dans la STEP, ce qui reflète typiquement le nombre de cas dans la population contribuant à la STEP (Figure 7-A). Si les concentrations du SRAS-CoV-2 varient, mais ne varient pas relativement aux concentrations du marqueur populationnel, c'est-à-dire que les deux varient dans la même direction et de la même amplitude, alors la variation de la concentration du SRAS-CoV-2 n'est pas dû à un changement dans le taux

d'excrétion de particules virale dans la STEP ou à un changement de la prévalence des cas dans la population contribuant (Figure 7-C). Cette variation est donc sûrement due à un changement soit dans un des processus inclus dans la STEP (dilution, dégradation, transport, sédimentation, sorption), dans le nombre total de personnes qui contribue à la STEP (événement culturel/sportif, jour férié, etc.) ou dans une autre forme de variation affectant les deux concentrations simultanément. La Figure 7-B et 7-D illustre aussi des situations où le marqueur populationnel n'est pas stable dans le temps dû à des taux d'excrétion variable.

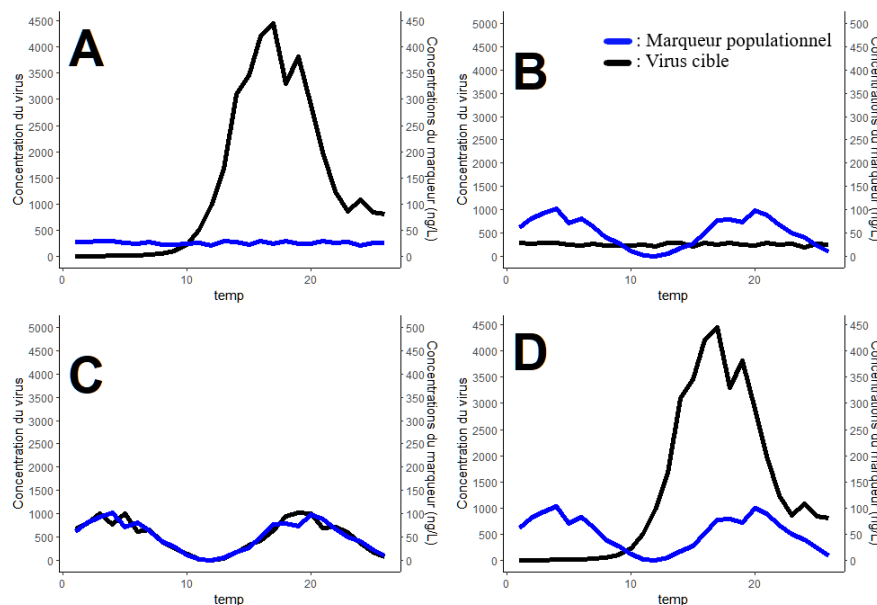


Figure 7. Exemples de relations et interprétations entre les concentrations d'un virus cible (noir) et d'un marqueur populationnel (bleu). A. Une situation idéale, le marqueur est stable dans le temps et permet de bien voir les variations dans la concentration du virus tout en permettant la normalisation adéquate des données. B. Une situation problématique, le marqueur varie grandement dans le temps alors que le virus est stable, ce qui augmenterait grandement la variabilité de la normalisation et pourrait produire des données qui sont non-représentative de la situation épidémiologique de cette région à ce moment donné. C. Une situation particulière, où le marqueur et le virus ont les mêmes variations de concentrations, démontrant soit un mauvais marqueur populationnel (qui varient trop dans le temps) ou des facteurs externes affectant les deux concentrations simultanément *e.g.* des variations dans le nombre de personnes dans la population contribuant. D. Une situation médiocre, le marqueur est plus ou moins stable dans le temps et permet de voir les variations dans la concentration du virus, mais pourrait négativement affecter les données lors de la normalisation.

Il existe d'autres manières de conceptualiser la normalisation ainsi que d'autres termes utilisés pour décrire les données ou les composés utilisés pour la normalisation, mais la majorité sont basés sur le même principe et le même fonctionnement. C'est le cas des indicateurs fécaux. Les indicateurs fécaux sont uniques des cibles d'ÉBEU qui proviennent

de la matière fécale, ce qui est le cas du SRAS-CoV-2 (Patey *et al.*, 2021). Les indicateurs fécaux permettent aussi de normaliser les concentrations du SRAS-CoV-2 dans les eaux usées, mais en étant une mesure de la quantité de matière fécale dans l'échantillon lui-même. En comparaison, les marqueurs populationnels sont une mesure de la contribution des personnes dans la STEP. Considérant l'aspect hautement aléatoire de l'échantillonnage, la quantité de matière fécale entre deux échantillons de la même STEP et de la même période peuvent avoir une variation non négligeable (Hundermark et Takashi, 2020). Cela fait en sorte qu'il est possible de normaliser les concentrations de SRAS-CoV-2 en se basant sur la concentration d'un indicateur fécal. Une même cible ou substance peut avoir les deux rôles simultanément, mais cela n'est pas toujours le cas. Par exemple, puisque le PMMoV est associé à la matière fécale, il peut être considéré et utilisé comme un indicateur fécal, mais aussi comme un marqueur populationnel dû à ses autres caractéristiques (Gracia-Lor, *et al.*, 2017). Cependant, la caféine peut être utilisée comme marqueur populationnel, mais puisqu'elle provient de l'urine, elle ne peut pas être utilisée en tant qu'indicateur fécal. Pour le projet présent, le terme marqueur populationnel est priorisé puisque certains marqueurs utilisés ne peuvent pas être considérés comme des indicateurs fécaux, même si deux parmi eux le peuvent. Les critères de sélections à la section 1.2 ont aussi été basés sur le concept de marqueurs populationnels et non sur le concept d'indicateur fécal.

La normalisation des concentrations par un marqueur populationnel est une étape importante de l'ÉBEU. Cependant, ceci ne remplace pas les étapes et les précautions nécessaires pour assurer et contrôler la qualité des données produites. La normalisation s'ajoute à la méthode sans remplacer ou enlever une autre étape. De plus, elle ne remplace pas les données de dépistage clinique, des données plus souvent utilisées par les autorités de santé publique, dans un contexte de pandémie. L'ÉBEU et la normalisation sont un excellent outil qui performe mieux, lorsqu'utilisé en tandem avec d'autres méthodes tel le dépistage clinique. Pour permettre aux autorités décisionnelles d'intervenir de manière éclairée et juste, il est idéal d'utiliser ces deux méthodes simultanément puisque les failles de l'une sont les forces de l'autre.

1.2 CRITÈRES DE SÉLECTION

La précision et la fiabilité de la normalisation par un marqueur populationnel dépendent grandement du choix du marqueur. Chaque STEP possède une population contributive unique. C'est-à-dire que chaque groupe de personnes formant la population qui contribue à une STEP est différent. Cela entraîne plusieurs défis liés à la mise en œuvre et l'utilisation adéquate des marqueurs populationnels. Malgré l'efficacité potentielle d'un certain marqueur, il est possible que ce marqueur soit efficace dans certaines régions, mais que ce ne soit pas le cas dans d'autres (Mitrancescu *et al.*, 2022). Il est alors très important, mais souvent peu facile, de faire un bon choix de marqueur populationnel pour une STEP en particulier.

Les critères de sélection suivants ont pour but premier de déterminer quelles substances, molécules ou données pourraient servir de marqueur populationnel dans le cadre d'ÉBEU pour le SRAS-CoV-2. Cela étant dit, les critères peuvent avoir une application plus large. Ils peuvent être réutilisés dans la sélection d'autres marqueurs populationnels pour des utilisations futures *e.g.* d'autres pandémies.

Malgré plusieurs critères de sélection, ceci ne constitue pas l'entièreté du processus de sélection et d'application des marqueurs populationnels. Ceci n'est qu'une sélection préliminaire ; une manière de réduire les centaines de marqueurs populationnels potentiels pour ensuite évaluer, en laboratoire, la performance de quelques-uns d'entre eux qui semblent les plus prometteurs. Il est impossible d'effectuer un choix singulier uniquement à partir de la littérature. Comme mentionné plus haut, il y a une variation considérable entre chaque STEP, même à l'intérieur d'une petite région. Il est impératif de valider l'efficacité et la performance des marqueurs populationnels choisis pour confirmer ou infirmer leur habileté à normaliser les concentrations de la cible primaire, dans ce cas-ci le SRAS-CoV-2. Même après une sélection rigoureuse, il est possible qu'aucun marqueur potentiel identifié ne fonctionne bien dans la situation désirée. Cela ne veut pas dire qu'il n'en existe pas, cela

démontre d’abord la variabilité dans le domaine de l’ÉBEU et ensuite le besoin d’expériences exhaustives pour arriver à des finalités acceptables.

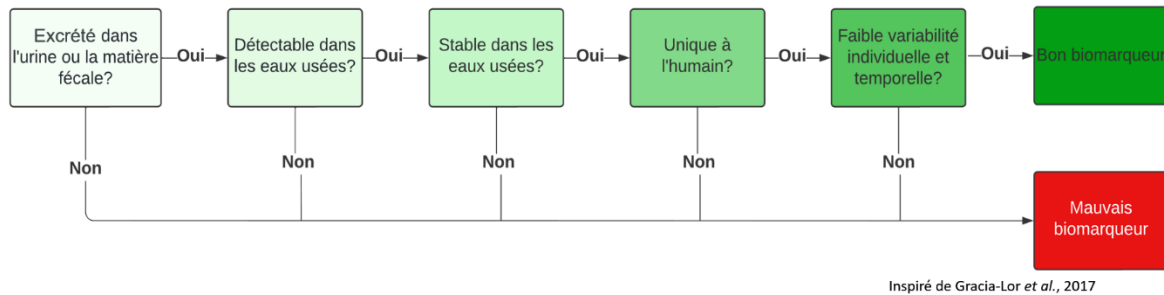


Figure 8. Diagramme des principaux critères de sélection pour les marqueurs populationnels (biomarqueurs)

1.2.1 L’excrétion par l’urine ou les fèces

Le premier critère est à la base même du principe d’ÉBEU : l’excrétion. Pour qu’un marqueur puisse représenter un aspect de la santé d’une population, ou pour être capable d’estimer la taille de la population, le marqueur doit d’abord être excrété et être acheminé dans la STEP par les égouts. Le mode d’excrétion est aussi un facteur important à considérer puisqu’il peut avoir un impact sur la méthode d’analyse utilisée lors de la quantification. Les marqueurs populationnels qui sont excrétés par l’urine sont échantillonnés et analysés à partir de la phase liquide de l’eau usée alors que les marqueurs populationnels qui sont excrétés par les fèces sont typiquement échantillonnés et analysés à partir de la phase solide *i.e.* les particules de matières fécales en suspension. Si le composé n’est jamais excrété dans les eaux usées, l’ÉBEU n’est pas la bonne approche et il est impossible d’utiliser ces méthodes pour en tirer des conclusions valides.

1.2.2 La détectabilité dans les eaux usées

Le deuxième critère est la détection du marqueur dans les eaux usées. Ce critère comprend plusieurs aspects, mais est résumé par la fiabilité de la détectabilité. Malgré les avancées technologiques en matière de détection et quantification, certaines choses restent inatteignables pour l’instant. Il est important de considérer les limites de détection et de

quantification des méthodes envisagées puisque cela va influencer le choix des marqueurs populationnels. Si les quantités ou les concentrations du marqueur potentiel sont trop basses, soit la détection ou la quantification ne sera pas être fiable à travers le temps dû à l'aspect aléatoire de l'échantillonnage, soit la détection ou la quantification sera simplement impossible. Même si certains composés semblent être de bons candidats, si leurs concentrations dans la STEP ciblé sont régulièrement sous la limite de quantification des méthodes utilisées, alors il vaut mieux ne pas considérer ce composé.

1.2.3 La stabilité dans les eaux usées

Le troisième critère est la stabilité physico-chimique et biologique du marqueur populationnel dans les eaux usées. Comme mentionné dans la section 1.1, il existe plusieurs processus qui affectent les concentrations des marqueurs populationnels dans les eaux usées ; ceux-ci doivent être pris en compte. Plus le marqueur est stable, moins ces processus auront un effet négatif sur la fiabilité et la reproductibilité des analyses. Une bonne stabilité aide à la production de données utilisables dans la prise de décision. Toutefois, il faut aussi considérer la stabilité du SRAS-CoV-2. Les marqueurs possédant une stabilité similaire au SRAS-CoV-2 peuvent bien représenter les dynamiques et les interactions dans les eaux usées, même s'ils possèdent une stabilité plus faible que d'autres marqueurs.

1.2.4 L'unicité de la provenance humaine

Le quatrième critère est l'unicité de la provenance humaine. C'est-à-dire qu'un bon marqueur populationnel provient uniquement de l'excrétion humaine et ne possède pas d'autres sources qui pourraient s'acheminer dans les eaux usées. Cela peut inclure des sources telles que les animaux de compagnies, les produits ménagers, des produits industriels, les fertilisants, pesticides et autres produits d'agriculture, les déchets, les produits pharmaceutiques, la nourriture et d'autres produits consommables, ainsi que le milieu naturel. Il est important de ne pas avoir d'autres sources pour un marqueur populationnel pour assurer que les données produites représentent la population humaine contribuant à la

STEP et non d'autres aspects de la société ou de la région comme la pollution ou la production de certain produit par les industries contribuant à la STEP. La façon la plus efficace pour faire cela est de considérer les métabolites humains, les produits consommés uniquement par les humains ou le matériel génétique spécifiquement associé aux humains ; ceci sera abordé plus en détail dans la section 1.3.

1.2.5 La variabilité temporelle, spatiale et individuelle

La variabilité est le cinquième et dernier critère incontournable dans la sélection d'un marqueur populationnel adéquat. La variabilité temporelle, spatiale et individuelle réfère à la variabilité du taux d'excrétion d'un marqueur populationnel dans le temps (jours, semaines, saisons, etc.), à travers différentes villes ou régions et entre les individus contribuant à la STEP, respectivement. Quoique la variabilité puisse être la cible ou la finalité d'une étude et peut expliquer certains aspects reliés à l'épidémiologie d'une population ciblée, dans le cadre de la normalisation telle qu'ici, elle doit être tenue à un minimum. Il est important d'avoir une excrétion stable, peu variable, du marqueur pour être en mesure d'identifier efficacement les causes de variations dans la cible primaire, le SRAS-CoV-2. Si le marqueur a plusieurs sources de variations, alors il est plus difficile d'expliquer et d'interpréter la variation dans les concentrations de SRAS-CoV-2 normalisées. Cela fait en sorte que si les concentrations normalisées varient, et que le marqueur est stable *i.e.* a une faible variation, cette variation des données normalisées est plus facilement attribuable au changement de l'incidence de la COVID-19 dans la population.

Il existe plusieurs types de variations et chacun a sa particularité. D'abord, il y a la variation temporelle. Celle-ci réfère spécifiquement à la variation dans le temps, mais elle peut être à plusieurs échelles différentes, soit séparément ou simultanément. Par exemple, un marqueur populationnel pourrait avoir un taux d'excrétion qui varie selon les saisons ou les mois de l'année. Il pourrait aussi avoir une variation hebdomadaire, avec une plus forte excrétion la fin de semaine par rapport à la semaine, ou vice versa. L'alcool est un bon

exemple de composé avec une forte variation hebdomadaire, ayant une plus haute excrétion la fin de semaine et une excrétion plus faible la semaine (Reid *et al.*, 2011).

Ensuite, il y a la variation spatiale. Celle-ci réfère à la différence entre certaines régions ou parties du monde. C'est-à-dire qu'une région pourrait avoir un taux d'excrétion élevé pour un certain marqueur populationnel alors qu'une autre pourrait avoir un taux d'excrétion plus faible pour ce même marqueur. Ceci n'est pas intrinsèquement négatif, c'est possible d'utiliser un marqueur populationnel avec une forte variation spatiale, mais cette variation ajoute un niveau de complexité supplémentaire si ce marqueur est pour être utilisé à travers plusieurs régions pour les mêmes utilisations. Cette variation réduit l'universalité d'un marqueur, forçant certaines régions à utiliser un marqueur différent, ce qui peut nuire aux comparaisons d'incidence de COVID-19 entre des régions différentes. La nicotine et ses métabolites comme la cotinine sont des bons exemples de marqueurs qui possèdent une forte variation spatiale due à la différence dans la taille de la population fumeuse à travers certains pays (Chen *et al.*, 2014). Toutefois, cela n'empêche pas son utilisation dans un contexte d'ÉBEU en Italie et en Australie par exemple (Chen *et al.*, 2014 ; Senta *et al.*, 2015).

La dernière forme de variation est la variation individuelle. Chaque individu est différent, alors chaque individu excrète divers marqueurs à des taux différents ; ceci est inévitable. Toutefois, il est important de minimiser cette variation ou de sélectionner des marqueurs qui sont moins prône à ce type de variation. Une grande variabilité individuelle pourrait augmenter l'aspect aléatoire de l'échantillonnage de l'eau usée, ce qui pourrait mener à des résultats moins fiables et moins constants, réduisant l'exactitude des données normalisées produites. Les produits pharmaceutiques ou les métabolites associés à des conditions médicales sont des exemples de composés ayant typiquement une haute variabilité individuelle (Castiglioni *et al.*, 2006).

1.2.6 La facilité et rapidité d'analyse

La technique d'analyse nécessaire pour quantifier un certain marqueur n'est pas un critère qui devrait être priorisé, mais il peut être pertinent de le considérer. Avoir les

ressources nécessaires pour la production de données exactes, précises et fiables est une partie importante de l'ÉBEU. Ceci est encore plus vrai pour la quantification des marqueurs populationnels. Plus la quantification d'un marqueur est accessible, plus des régions différentes, possédant possiblement moins de ressources, pourront utiliser ce marqueur populationnel avec la même méthode. Lorsque le but de la sélection est d'identifier des marqueurs populationnels avec un potentiel d'applicabilité universel, ceci devient important à considérer. Toutefois, il n'est pas nécessaire de se limiter à une méthode ou une autre si cela engendre une diminution significative de la qualité des données produites par un marqueur ou si cela enlève la possibilité d'utiliser un bon marqueur populationnel.

D'autre part, il est important de considérer la rapidité de l'analyse, lorsque comparée avec la rapidité de l'analyse de la cible primaire. Si la méthode pour la quantification du marqueur populationnel ne peut pas être effectuée simultanément ou la même journée que la cible primaire, cela peut causer certains problèmes. Ce n'est pas systématiquement problématique, mais ça pourrait ralentir la publication et l'utilisation des données de normalisation et affecter négativement le temps nécessaire aux autorités de santé publique pour prendre des décisions importantes. La vitesse à laquelle les données d'ÉBEU sont produites est son plus grand avantage, alors ralentir ce processus pourrait être désavantageux.

1.2.7 La pharmacocinétique

La pharmacocinétique est une branche de la pharmacologie qui englobe les quatre grandes étapes liées à la prise de médication ou de substance : l'absorption, la distribution, la métabolisation et l'excrétion. Donc, c'est un domaine qui traite des processus utilisés par le corps humain lors de la consommation de certaines substances ou molécules. Les données liées à la pharmacocinétique des marqueurs populationnels peuvent aider à faire un choix judicieux. Comme mentionné plus haut dans la section 1.2.4 et élaboré dans la section 1.3, les métabolites et leurs composés parents peuvent être de bons marqueurs populationnels. Cependant, pour en connaître leur existence et leur taux d'excrétion, des études sur leur pharmacocinétique doivent avoir été faites. C'est-à-dire que ces études sont la base des

connaissances pour plusieurs marqueurs populationnels potentiels. Cela indique aussi l'importance et le potentiel de ce type de connaissances et d'études.

La pharmacocinétique sert d'abord à identifier les métabolites qui peuvent se retrouver dans les eaux usées, mais permettent aussi de connaître le taux d'excrétion des différents métabolites. Pour la normalisation du SRAS-CoV-2, connaître la valeur exacte d'excrétion n'est pas nécessaire, mais cela permet d'évaluer la variabilité individuelle de l'excrétion, qui est un critère important dans la sélection d'un marqueur populationnel. Cependant, pour des applications d'ÉBEU plus larges ou différentes des applications dans le cadre du projet présent, les données de pharmacocinétique deviennent cruciales puisqu'elles permettent, entre autres, de rétrocalculer la consommation ou l'exposition d'un composé et de confirmer la provenance de certaines substances (Gracia-Lor *et al.*, 2017).

1.3 IDENTIFICATION DES NOUVEAUX MARQUEURS POPULATIONNELS

Il existe plusieurs groupes très distincts de marqueurs populationnels. Chacun de ces groupes possède des avantages et des inconvénients ainsi que des possibilités d'applications et d'utilisations différentes. Ces catégories seules ne sont pas assez pour déterminer si un marqueur populationnel est efficace, il est important de regarder chaque marqueur individuellement pour déterminer leur potentiel pour normaliser les concentrations du SRAS-CoV-2 dans les eaux usées. La catégorisation est effectuée uniquement à guise de structure et ne reflète pas nécessairement l'efficacité de ceux-ci, même si la majorité des marqueurs appartenant à la même catégorie ont souvent des caractéristiques similaires.

1.3.1 Matériel génétique

Le matériel génétique possède plusieurs voies pour s'acheminer dans une STEP. D'abord, il est retrouvé dans l'urine et dans la matière fécale de manière systématique. De plus, la perte de cellules épithéliales, telles que des cellules de peau morte lors d'une douche,

est assez grande pour en retrouver dans les eaux usées également. Grâce aux avancées dans le domaine de la PCR et du séquençage génétique (l'identification des séquences génétiques qui ont été amplifiées par PCR), il est possible de détecter de manière fiable de très petites quantités de matériel génétique. L'ADN est l'un des composés avec la plus grande stabilité dans les eaux usées, alors sa quantification est fiable et reproductible dans le temps. L'ARN est assez stable pour en faire des analyses d'assez bonne qualité, mais sa stabilité est plus faible que celle de l'ADN. Cette différence de stabilité est dû à deux raisons. D'abord, l'ADN est généralement sous la forme de double brin, alors que l'ARN est généralement sous la forme d'un simple brin. Quoiqu'il y ait des exceptions, la structure de double brin de l'ADN est moins susceptible à différents processus de dégradation. De plus, le désoxyribose, retrouvé dans l'ADN, comparativement au ribose dans l'ARN, est moins réactif, ce qui augmente la stabilité de l'ADN. Généralement, les marqueurs génétiques ont une excellente traçabilité, c'est-à-dire qu'il est relativement facile d'identifier l'origine et la provenance du marqueur pour confirmer, par exemple, que l'humain en est la seule source. Enfin, il peut y avoir une assez forte variabilité spatiale ou individuelle pour les marqueurs génétiques, mais cela dépend du marqueur spécifique.

Le matériel génétique inclut plusieurs sous-catégories : l'ADN humain, l'ADN viral, l'ARN viral, l'ADN bactérien et l'ARN bactérien. L'ADN et l'ARN viral sont de loin le type de marqueur génétique le plus utilisé. Par exemple, le crAssphage, le PMMoV, les polyomavirus et les adénovirus ont tous déjà été utilisés dans un contexte d'ÉBEU, plus spécifiquement pour la normalisation du SRAS-CoV-2 (Ahmed *et al.*, 2020b ; D'Aoust *et al.*, 2021 ; CDC, 2023 ; Bofill-Mas *et al.*, 2006). Ce qui distingue ces marqueurs des autres est leur abondance dans les eaux usées, leur ubiquité à travers plusieurs régions dans le monde, leur forte stabilité dans les eaux usées et leur détection fiable. Toutefois, ils peuvent avoir une variation du taux d'excrétion modérée et dans le cas du PMMoV avoir d'autres sources telles que la nourriture non consommée et l'agriculture. Malgré cela, le PMMoV a été le premier marqueur recommandé par le CDC et le crAssphage fait maintenant partie de cette liste depuis la fin de l'année 2022 (CDC, 2023). Cependant, le PMMoV possède une variation spatiale non négligeable forçant certaines régions, comme le Québec, d'identifier

un autre marqueur populationnel plus efficace (Rioux *et al.*, 2023). Le plus grand avantage d'utiliser des marqueurs d'ARN viral, comme le PMMoV, est qu'ils peuvent tous être analysés simultanément avec le SRAS-CoV-2 avec peu ou aucun changement dans la méthode. Ceci permet une production et publication des données très rapide en plus d'éviter l'achat de nouvel équipement ou la formation d'une nouvelle main-d'œuvre.

L'ARN bactérien a aussi déjà été utilisé dans le contexte de la pandémie de la COVID-19 avec un succès moyen. Les deux cibles les plus utilisées sont l'ARN ribosomal 16s HF183 d'un bactéroïde humain et l'ARN ribosomal 18s (Greenwald *et al.*, 2021). Ceux-ci possèdent également une bonne abondance dans les eaux usées et une bonne traçabilité, mais leur variabilité laisse à désirer au niveau spatial et individuel. De plus, l'ARN ribosomal 16s HF183 peut aussi provenir de certains animaux, influençant la fiabilité des données produites.

Le matériel génétique est une des catégories qui se marie le mieux dans le contexte présent avec le SRAS-CoV-2. Cela est dû en partie à leurs similarités structurales et épidémiologique, ce qui explique leur grande utilisation à travers le monde depuis le début de la pandémie de la COVID-19. Il est important de noter que la similarité structurale, menant à des comportements similaires dans l'eau usée, n'est pas le seul critère qui fait un bon marqueur populationnel. Un marqueur de la même nature que la cible peut aider à avoir une meilleure performance et est très avantageux pour l'analyse en laboratoire. Cependant la nature du marqueur ne sera jamais le seul élément menant à une bonne performance. Ils ne sont pas tous adéquats et peuvent posséder une variation spatiale et individuelle, ou tout autre défaut, nuisible au potentiel universel de ces marqueurs populationnels.

1.3.2 Bactéries et parasites

Les bactéries et les parasites ont d'abord été utilisés comme indicateurs de pollution humaine dans l'environnement (Tolouei *et al.*, 2019). Ce n'est que récemment que leur potentiel comme marqueur populationnel dans un contexte d'ÉBEU a été exploré (Tolouei *et al.*, 2019). Ils sont significativement moins utilisés que d'autres catégories et cela est dû à certains inconvénients (Tolouei *et al.*, 2019). Les bactéries et les parasites sont uniquement

excrétés à travers la matière fécale et leur stabilité dans les eaux usées dépend uniquement de leur capacité à survivre aux conditions dans la STEP. Ils sont détectables avec facilité, habituellement via dénombrement suivant une filtration par membrane, mais d'autres méthodes peuvent aussi être utilisées (Tolouei *et al.*, 2019). Leurs plus grands inconvénients sont leur variabilité et leurs autres sources. Leur variabilité temporelle et individuelle peut être assez élevée, ce qui diminue grandement leur fiabilité en tant que marqueur populationnel. De plus, la majorité des bactéries et des parasites qui peuvent être utilisés comme marqueur populationnel se retrouvent aussi naturellement dans l'environnement *i.e.* les sédiments et les cours d'eau (Tolouei *et al.*, 2019). Cela diminue grandement la fiabilité des bactéries et des parasites comme marqueurs populationnels dans le contexte présent.

Les bactéries dont le potentiel comme marqueur a été exploré incluent *Escherichia coli*, *Clostridium perfringens* tandis que les parasites incluent les oocystes de *Cryptosporidium* et les kystes de *Giardia*. Parmi ces micro-organismes, les kystes de *Giardia* et les colonies de *E. coli* n'ont pas pu être directement associés avec la matière fécale, ce qui rend leur utilisation comme marqueur populationnel invalide (Tolouei *et al.*, 2019). *C. perfringens* peut s'accumuler dans les sédiments à l'intérieur d'une STEP qui peuvent ensuite être remis en suspension lors d'un événement de haut débit tel que la fonte de la neige ou des averses importantes ; cela peut fortement influencer les analyses de concentrations dans l'eau usée et biaiser l'interprétation des données de normalisation (Tolouei *et al.*, 2019). Les oocystes de *Cryptosporidium* sont très difficilement détectables et sont rarement au-dessus du seuil de quantification. Donc les bactéries et les parasites ont une bonne utilité comme marqueur de pollution dans l'environnement, mais manquent plusieurs qualités et caractéristiques importantes pour leur utilisation en tant que marqueur populationnel dans le contexte présent.

1.3.3 Édulcorants

Les édulcorants sont omniprésents dans la nourriture et les boissons de la majorité des régions à travers le monde (Kokotou *et al.*, 2013). Leur grande présence, au niveau de leur popularité à travers le globe et de quantité consommée, permet de les utiliser adéquatement

dans un contexte d'ÉBEU (Kokotou *et al.*, 2013). Ils sont tous excrétés à travers l'urine et possèdent une stabilité grandement variable, soit très haute ou très basse (Kokotou *et al.*, 2013). Ceci est de même pour la détection, pour certain elle est assez élevée et donc très fiable, alors que pour d'autres elle est très basse et alors peu fiable. Ce qui est constant pour tous les édulcorants est la variabilité et les sources de contamination (Kokotou *et al.*, 2013). La variabilité spatiale est typiquement faible ou modérée, mais la variation individuelle et temporelle peut être importante. La consommation, donc la concentration dans les eaux usées, peut varier à l'intérieur d'une même semaine avec des différences entre la fin de semaine (samedi et dimanche) et les jours de semaine (lundi à vendredi) (Kokotou *et al.*, 2013 ; Gracia-Lor *et al.*, 2017). Aussi, la consommation individuelle peut grandement varier ; certains individus peuvent activement éviter de consommer des édulcorants, alors que d'autres en consomment régulièrement à travers leur diète (Kokotou *et al.*, 2013). Les différentes sources de contamination peuvent inclure, mais ne sont pas limitées à, l'agriculture, la production et la consommation de la nourriture animale et la nourriture et les boissons non consommées qui sont directement jetées dans la STEP par les usagers (Kokotou *et al.*, 2013). Le tout fait des édulcorants un groupe polaire de potentiels marqueurs populationnels, certains ayant un grand potentiel, et d'autres ayant un pauvre potentiel.

Parmi les édulcorants qui pourraient servir de marqueur populationnel, il y a l'acésulfame, le cyclamate, la saccharine et le sucralose. Ces quatre édulcorants ont une haute stabilité et une bonne détection (Kokotou *et al.*, 2013). En plus, ces quatre édulcorants ne sont pas métabolisés par le corps humain (Kokotou *et al.*, 2013). Donc tous les édulcorants qui sont consommés sont directement retrouvés dans les eaux usées. Il n'y a pas de perte engendrée par le métabolisme humain. Le besoin de données de pharmacocinétique est alors réduit. Toutefois, cela fait en sorte que lors d'une contamination quelconque, il est plus difficile de distinguer l'origine exacte de ces édulcorants, puisque si leur provenance est d'un déversement quelconque qui s'achemine dans la STEP, alors il est indistinguable de ceux qui proviennent de la consommation personnelle. Cela est relativement rare, mais ce type d'évènement n'est pas impossible.

Les édulcorants qui ne pourraient pas servir de marqueur populationnel incluent l'alitame, l'aspartame, le néotame et la néohespéridine dihydrochalcone. Ces édulcorants possèdent une stabilité très faible et une détection non fiable (Kokotou *et al.*, 2013). Cependant, à l'opposé des quatre édulcorants mentionnés plus haut, ils sont tous métabolisés à différents degrés par le corps humain (Kokotou *et al.*, 2013). Il est donc nécessaire d'analyser les métabolites respectifs de ces édulcorants. Cela nécessite des données de pharmacocinétique, mais permettrait de valider la provenance humaine des composés, ce qui améliorerait la fiabilité des données par rapport à la contamination ou aux autres sources possibles de ces édulcorants. Toutefois, ces quatre édulcorants sont tout de même inadéquats en tant que marqueur populationnel.

Les édulcorants, de manière générale, ont donc un certain potentiel en tant que marqueur populationnel (Kokotou *et al.*, 2013). Cependant, le choix exact de l'édulcorant est très important, puisque cela peut grandement influencer la qualité des résultats obtenus. Ceci ne retire pas leur utilité à servir comme marqueur de santé ou même de pollution dans certaines situations (Gracia-Lor *et al.*, 2017). Il est important aussi de considérer les méthodes d'analyse possible pour ces composés. Les méthodes les plus populaires sont la chromatographie en tandem avec un spectromètre de masse, plus spécifiquement la chromatographie en phase liquide haute performance (HPLC).

1.3.4 Drogues

Les drogues illicites sont parmi les premiers composés à être quantifiés dans les eaux usées dans un contexte d'ÉBEU (Palardy *et al.*, 2016). Le but était d'identifier les patrons de consommation ainsi que l'intensité et la quantité de substance illicite consommée par une certaine population (Palardy *et al.*, 2016). Puisque les méthodes étaient déjà développées pour la quantification de plusieurs de ces substances, leur potentiel en tant que marqueur populationnel a rapidement été mis à l'épreuve.

Les drogues incluent non seulement les substances illicites telles que la cocaïne, les méthamphétamines et anciennement le THC et le CBD, mais incluent aussi des drogues

consommées régulièrement à travers le monde telles que l'alcool, la nicotine et la caféine. Cette catégorie comporte de grandes différences à l'intérieur même de la catégorie malgré le nombre limité de composés potentiels.

Les drogues illicites ont été considérées que très brièvement. Leur consommation et leur présence dans les eaux usées sont très faibles et extrêmement variables au niveau temporel, spatial et individuel. La proportion de consommateurs de ces substances dans la population en générale est beaucoup trop faible pour en faire de bons marqueurs populationnels.

Les autres drogues communément consommées, comme l'alcool, la nicotine et la caféine, ont été utilisées dans un contexte d'ÉBEU dans le passé (Chen *et al.*, 2014 ; Senta *et al.*, 2015). Ils sont majoritairement excrétés à travers l'urine et possèdent une bonne stabilité dans les eaux usées avec un certain succès au niveau de la nicotine et de la caféine (Chen *et al.*, 2014 ; Senta *et al.*, 2015 ; Oloye *et al.*, 2023). Il est possible de quantifier les métabolites de certaines de ces drogues, comme la cotinine pour la nicotine et la paraxanthine pour la caféine (Chen *et al.*, 2014 ; Senta *et al.*, 2015). Ceci nécessite des données de pharmacocinétique, mais permet d'éliminer plusieurs sources de contaminations telles que des boissons non consommées ou des mégots de cigarettes qui se retrouvent dans la STEP. L'analyse et la quantification des métabolites dans ce cas-ci requièrent également une limite de quantification plus basse. C'est-à-dire que la quantité de métabolites est typiquement plus faible que la quantité du composé parent, donc ils sont plus difficiles à quantifier (Chen *et al.*, 2014 ; Senta *et al.*, 2015).

1.3.5 Métabolites naturels

Les métabolites naturels humains possèdent déjà plusieurs utilités, notamment dans le domaine médical (Shi *et al.*, 2017). Les métabolites sont définis comme étant des substances fabriquées ou utilisées lors de la décomposition ou la formation des divers composés biochimiques dans le corps. Elles sont les produits du métabolisme. Cette catégorie inclut seulement les métabolites qui sont naturellement produits par le métabolisme humain et

n'inclut pas les métabolites provenant de la consommation de produits transformés ou manufacturés. Des exemples de métabolites non naturels incluent les métabolites de la caféine (paraxanthine), de la nicotine (cotinine), et des édulcorants. Cette classification est unique à ce document et a pour but de regrouper les composés parents avec leur métabolite potentiel et d'éviter à les séparer dans des sections différentes. Ceci permet aussi d'explorer les métabolites naturels dans une section distincte et en regroupant toutes les informations générales sur les métabolites dans la même section.

Il existe plusieurs métabolites qui sont excrétés à travers l'urine de manière quotidienne par les humains et qui peuvent être détectables de manière efficace dans les eaux usées par chromatographie en tandem avec un spectromètre de masse (Shi *et al.*, 2017 ; Senta *et al.*, 2015 ; Chen *et al.*, 2014 ; Brewer *et al.*, 2012 ; Hutchison *et al.*, 2022). La stabilité des métabolites peut grandement varier et il peut y avoir quelques sources de contamination, c'est-à-dire que pas tous les métabolites sont uniques à l'humain. Certains animaux par exemple peuvent excréter les mêmes métabolites à des taux relativement similaires (Hutchison *et al.*, 2022). Cependant, la variabilité est faible de manière générale.

Les deux métabolites les plus populaires dans un contexte d'ÉBEU, excluant les métabolites des produits consommés mentionnés plus haut (drogues, édulcorants), sont les métabolites de la créatine et de la sérotonine, soit la créatinine et l'acide 5-hydroxyindoleacétique (5-HIAA) respectivement (Hutchison *et al.*, 2022 ; Hsu *et al.*, 2022 ; Chen *et al.*, 2014). La créatinine possède une stabilité relativement faible dans les eaux usées, mais la quantité qui est excrétée compense au niveau de la détection pour qu'elle soit fiable malgré la stabilité (Hutchison *et al.*, 2022). Elle a une variabilité modérée au niveau individuel et spatial, mais une faible variabilité temporelle (Hutchison *et al.*, 2022). Toutefois, le plus gros désavantage est la présence de plusieurs sources de contamination. Tous les animaux excrètent la créatinine, ce qui inclut les animaux domestiques et les animaux de fermes et d'élevage. La présence de l'industrie animale à proximité d'une STEP rend l'utilisation de la créatinine comme marqueur populationnel très peu efficace (Hutchison *et al.*, 2022). Le 5-HIAA possède une stabilité élevée, une détectabilité fiable,

une très faible variation globale et aucune source de contamination significative (Hsu *et al.*, 2022 ; Chen *et al.*, 2013). Ceci fait de ce composé un marqueur populationnel avec un potentiel très élevé.

Il existe plusieurs autres métabolites qui sont détectables dans les eaux usées : le cholestérol, le cortisol, le coprostérol, l'aminocétone et l'androstènedione en sont des exemples. Toutefois, leur stabilité et leur variabilité font en sorte qu'ils ne sont pas de bons candidats pour être des marqueurs populationnels dans le contexte de la normalisation des concentrations de SRAS-CoV-2 (Gracia-Lor *et al.*, 2017). De plus, ils sont souvent retrouvés à des niveaux relativement bas, faisant en sorte que leur détection et leur quantification dans les eaux usées sont peu fiables. Cela n'empêche pas de les considérer pour une autre fonctionnalité en ÉBEU (Gracia-Lor *et al.*, 2017).

1.3.6 Produits pharmaceutiques et produits de soins personnels

Les produits pharmaceutiques ainsi que les produits de soins personnels (PCP, « Personal Care Products ») sont consommés de manière quotidienne par une grande partie de la population dans les pays développés et en développement (Gracia-Lor *et al.*, 2017). Ces produits peuvent inclure les médicaments de prescription (opioïdes, drogue antiinflammatoire non stéroïdienne (NSAID), pilules contraceptives, benzodiazépines, antibiotiques, bêtabloquants, etc.), les médicaments en vente libre (ibuprofène, acétaminophène, etc.), les contrastes pour l'imagerie médicale, les produits cosmétiques (maquillage, crèmes, etc.) et les produits d'hygiène (savon, détergent, désinfectant, dentifrice, shampooing, etc.). Il y a donc une grande variété de produits et de composés qui sont inclus dans cette famille de substances ; pour la liste complète, voir l'annexe I. Tous sont excrétés à des niveaux et des fréquences différentes, rendant leur applicabilité parfois difficile dans le contexte de marqueur populationnel (Gracia-Lor *et al.*, 2017).

La majorité des produits pharmaceutiques et PCP sont excrétés en trop petite quantité par trop peu de personnes sur une base trop peu régulière pour être considérés à des fins de normalisation. De plus, leur consommation n'est pas assez répandue, c'est-à-dire qu'il n'y a

pas assez de personnes qui consomment et qui excrètent un certain produit à un temps donné pour que ce produit agisse efficacement en tant que marqueur populationnel. Aussi, il faut considérer le lien entre la consommation de ces produits, notamment les médicaments tels l'ibuprofène et l'acétaminophène, et la hausse du nombre de cas de COVID-19 (ou de maladies semblable) dans la population (Lux *et al.*, 2022). Par exemple, s'il y a une augmentation du nombre de cas de COVID-19 dans une ville, la concentration du SRAS-CoV-2 dans les eaux usées sera en hausse. Alors, il est attendu que la consommation d'ibuprofène augmente grâce à la hausse de cas de COVID-19, augmentant ainsi les concentrations dans les eaux usées. Donc si les concentrations de SRAS-CoV-2 étaient normalisées par l'ibuprofène, il y aurait une hausse dans chacune des concentrations respectives, ce qui en quelque sorte diminuerait, ou cacherait, la véritable augmentation des cas dans la population basée sur les données d'ÉBEU contrairement à si la concentration d'ibuprofène était stable temporellement, ce qui serait l'objectif d'un marqueur populationnel.

Certains produits pharmaceutiques ont déjà été utilisés en tant que marqueur populationnel avec un succès limité (O'Brien *et al.*, 2014). En utilisant la combinaison de plusieurs produits pharmaceutiques, jusqu'à 14 produits différents, dans une STEP couvrant une grande population (*i.e.* $\geq 300\,000$ personnes) c'était possible de bien estimer la population contribuant à une STEP en utilisant un modèle bayésien (O'Brien *et al.*, 2014). Toutefois, ce type de modélisation n'est pas aussi efficace et rapide lorsque comparé à l'utilisation d'autres marqueurs de populations.

Donc, les produits pharmaceutiques et les PCP ne feraient pas des marqueurs populationnels idéaux pour la normalisation des concentrations de SRAS-CoV-2 dans les eaux usées. Leur taux d'excrétion et leur détectabilité sont trop faibles, leur stabilité est hautement variable et leur variabilité temporelle et individuelle est trop élevée, cela malgré la très faible présence d'autres sources de contamination.

1.3.7 Polluants et substances toxiques

Les polluants et les substances toxiques sont des vastes classes de substances qui sont retrouvées à des niveaux variables dans les eaux usées. Ils sont souvent issus de l'activité humaine et non de l'excrétion des individus directement, mais certains polluants persistants et plusieurs substances toxiques peuvent être consommés et ensuite excrétés. Les pesticides, les parabènes, les mycotoxines, les filtres-UV, les plastifiants et les ignifuges font partie de la catégorie de polluants et substances toxiques.

Considérant que la source première de ces substances dans les eaux usées est généralement leur présence dans l'environnement et non l'excrétion par les individus contribuant à la STEP, la majorité de ces composés n'ont pas d'applicabilité en tant que marqueur populationnel en ÉBEU (Gracia-Lor *et al.*, 2017). Toutefois, ils peuvent tout de même agir comme indicateur de santé ou d'exposition puisque leurs effets négatifs sur l'environnement et la santé humaine sont hautement étudiés (Gracia-Lor *et al.*, 2017).

Certains polluants et substances toxiques peuvent être métabolisés par le corps humain (Gracia-Lor *et al.*, 2017). Ceci permettrait de les utiliser comme marqueur populationnel potentiel. Cependant, il manque encore beaucoup d'information sur le sujet et les connaissances de pharmacocinétiques pour ce type de substance ne sont pas assez développées pour en faire une utilisation efficace pour l'instant.

Les mycotoxines, contrairement aux autres polluants, sont retrouvées en majorité dans la nourriture, notamment le riz, le maïs et le blé (Marroquín-Cardona *et al.*, 2014 ; Gracia-Lor *et al.*, 2017). Ces substances sont des métabolites secondaires des champignons qui peuvent avoir un effet néfaste sur la santé humaine, si elles sont présentes en trop grande quantité (Marroquín-Cardona *et al.*, 2014 ; Gracia-Lor *et al.*, 2017). Les mycotoxines et leurs métabolites sont excrétés à travers l'urine en faible quantité, ce qui peut rendre leur quantification difficile, mais possible. Il existe trop peu d'information sur leur stabilité et leur variabilité pour être considéré comme un marqueur populationnel à ce jour. Cependant, avec les changements climatiques et l'augmentation des demandes alimentaires en hausse, cela

pourrait justifier leur utilisation de plus en plus dans un contexte d'ÉBEU, ce qui pourrait mener à leur utilisation comme marqueur populationnel dans le futur (Marroquín-Cardona *et al.*, 2014).

1.3.8 Autres marqueurs

Il existe d'autres substances et d'autres données qui sont liées aux eaux usées qui possèdent une possible applicabilité en ÉBEU. Certaines sont déjà utilisées pour normaliser des données de concentrations d'ÉBEU alors que d'autres ont été considérées ou appliquées à des fins différentes.

Les composés azotés tels que l'acide ammoniacal ou l'ammonium et les composés phosphorés tels que l'acide phosphorique ou le phosphate sont des exemples de marqueurs populationnels qui ont déjà été utilisés (Gracia-Lor *et al.*, 2017, Hutchison *et al.*, 2022). Cependant, malgré leur bonne stabilité, détectabilité et variabilité, ces composés ne constituent pas de bons marqueurs populationnels puisqu'ils dépendent hautement de la caractérisation et la composition de l'eau usée (industriel, domestique ou mixte) en plus d'avoir plusieurs sources de contamination (Gracia-Lor *et al.*, 2017 ; Hutchison *et al.*, 2022). Ils peuvent davantage être utilisés comme variable latente au débit de la STEP, mais ils perdent leur utilité rapidement puisque le débit est une donnée facilement acquise de manière quotidienne.

Les données hydrochimiques telles que la demande chimique en oxygène (COD) et la demande biologique en oxygène (BOD) ont aussi été considérées. Cependant, comme les composés azotés et phosphorés, ceux-ci ne sont pas de bons marqueurs populationnels puisqu'il existe trop de sources différentes, en dehors de l'excrétion, pouvant influencer ces données de manière significative (Gracia-Lor, *et al.*, 2017 ; Hutchison *et al.*, 2022). De plus, ces données peuvent varier grandement dans le temps et l'espace avec des événements de contaminations ou météorologiques quelconques (Gracia-Lor *et al.*, 2017 ; Hutchison *et al.*, 2022).

Les données physiques, comme le débit journalier de la STEP et la matière particulaire en suspension (MPS), ne sont pas des marqueurs populationnels en tant que tels, mais peuvent tout de même jouer le rôle de normaliser les concentrations du SRAS-CoV-2 ou d'autres cibles primaires dans les eaux usées (Hrudey *et al.*, 2022). Le débit journalier de la STEP est une des données capables de mieux normaliser les concentrations de SRAS-CoV-2 (Fang *et al.*, 2022 ; Hrudey *et al.*, 2022 ; Maal-Bared *et al.*, 2023 ; Mitranescu *et al.*, 2022). Comme expliqué dans la section 1.1, il existe plusieurs processus modifiant les concentrations de SRAS-CoV-2 dans les eaux usées. La majorité de ces processus et les plus importants *i.e.* ceux ayant le plus grand impact, varie en fonction du débit (Hrudey *et al.*, 2022). Cela inclut la dilution, le transport, la sédimentation et la sorption (Mitranescu *et al.*, 2022). Alors il est logique de comprendre pourquoi le débit est capable de normaliser les données d'ÉBEU de manière efficace.

Les MPS ne servent pas directement à normaliser les données d'ÉBEU, mais peuvent être des facteurs ou des variables pouvant augmenter la précision et la prédictibilité de la modélisation (Hrudey *et al.*, 2022). Comme le débit, les MPS sont intrinsèquement liées à certains processus présents dans une STEP, notamment la sorption, la sédimentation et le transport, mais il est aussi lié au débit directement (Mitranescu *et al.*, 2022). Un haut débit a tendance à remettre en suspension les particules qui se sont sédimentées dans une STEP, ce qui pourrait augmenter les données de concentrations, sans nécessairement avoir plus d'excrétion dans la STEP à la base (Tolouei *et al.*, 2019). De plus, le débit n'est pas le seul phénomène qui peut causer ceci (Tolouei *et al.*, 2019). Cela fait des MPS une donnée potentiellement pertinente à prendre en compte lors de la normalisation et de la modélisation des concentrations de SRAS-CoV-2 dans les eaux usées.

1.4 SÉLECTION

Considérant les critères de sélections abordés dans la section 1.2, quatre marqueurs populationnels ont été sélectionnés. En ajout à ces critères, d'autres considérations ont aussi

été prises en compte, notamment de ne pas sélectionner quatre marqueurs populationnels faisant partie de la même catégorie ni de sélectionner quatre marqueurs populationnels qui sont excrétés suivant le même mécanisme. Pour cette raison, les marqueurs populationnels choisis font partie de 3 catégories différentes : matériel génétique, drogue et métabolite naturel.

Le premier marqueur populationnel qui a été sélectionné est le virus bactériophage *Carjivirus communis*, aussi nommé crAssphage. Ce bactériophage fait maintenant partie de la liste des marqueurs populationnels recommandés par le CDC, mais ne l'était pas au moment de la sélection (CDC, 2023). Ce bactériophage a été découvert et séquencé en 2014 par Dutilh *et al.* grâce au programme « Cross Assembly » qui a identifié ce virus à partir d'échantillons de matière fécale. Ce bactériophage est six fois plus abondant dans la matière fécale humaine que tout autre phage connu combiné (Dutilh *et al.*, 2014). Son extrême abondance dans la matière fécale fait de ce marqueur populationnel un candidat idéal. Ce virus est également excrété à travers la matière fécale, tout comme le virus du SRAS-CoV-2, alors son analyse requiert peu de modifications aux méthodes développées pour le SRAS-CoV-2 en plus de bien représenter la dynamique à l'intérieur d'une STEP dû à l'association avec la matière fécale. De plus, malgré le fait que certains gènes tels le gène KP06_gp31 puissent être retrouvés chez certains animaux, notamment le bétail, le porc et la volaille, le gène CPQ_056 est unique à l'humain et ne possède aucune autre source connue pouvant contaminer ou affecter l'analyse de ce gène dans les eaux usées (Garcia-Aljaro *et al.*, 2017). Aussi, ce bactériophage est retrouvé en quantité élevée chez tous les humains habitant dans une région avec un certain niveau d'industrialisation (Honap *et al.*, 2020). Plus précisément, ce virus est présent chez tous les humains, mais à une plus faible prévalence chez les populations traditionnelles de chasseurs-cueilleurs tels que les Hadza en Tanzanie et les Matsés au Pérou comparé au monde industriel moderne (Honap *et al.*, 2020). Cette différence significative nous importe peu puisque ces groupes n'utilisent pas de systèmes d'égouts ou des STEPs. Alors, elles vont rarement se retrouver à faire partie d'une population contribuant à une STEP, limitant leur impact sur des études d'ÉBEU. Il n'y a pas de patron ou des différences statistiques dans la prévalence du crAssphage selon la géographie ou des

variables démographiques telles que l'âge, le sexe, l'indice de masse corporelle et la santé des individus (Honap *et al.*, 2020). Cela fait en sorte que le virus est excrété par toutes et tous de manière universelle avec très peu de variabilité dans le temps, l'espace et au niveau individuel. Pour ce qui en est de sa stabilité dans les eaux usées, le crAssphage est un virus à double brin d'ADN (Dutilh *et al.*, 2014). Cela signifie qu'il est très stable dans les eaux usées et sa dégradation dans les eaux usées ne devrait pas influencer l'analyse et la quantification du bactériophage de manière significative (Gracia-Lor *et al.*, 2017).

Le deuxième marqueur populationnel qui a été sélectionné est l'ADN mitochondrial humain de la sous-unité 5 du NADH déshydrogénase. Considérant que le premier choix de marqueur populationnel est un virus à ADN, le choix de l'ADN mitochondrial est venu facilement, car les deux pourront être analysés simultanément en utilisant les mêmes méthodes de filtrations, extractions et de quantification. D'un point de vue de marqueur populationnel, l'ADN mitochondrial est très similaire au crAssphage : il est excrété en majorité par la matière fécale, il est excrété de manière universelle par les humains avec peu ou aucune variabilité significative, il est stable dans les eaux usées et est unique à l'humain. De plus, ce même gène a déjà été utilisé pour normaliser les concentrations de SRAS-CoV-2 dans les eaux usées avec succès au Kansas (É-U-A), démontrant qu'il pourrait avoir un succès similaire au Québec (Hutchison *et al.*, 2022). Toutefois, le plus gros défi de son utilisation est les sources de contamination lors de l'analyse et des travaux de laboratoire. Ce marqueur populationnel nécessite des mesures uniques et exhaustives pour éviter la contamination des analyses et assurer que sa quantification est non seulement précise, mais également exacte.

Le troisième marqueur populationnel qui a été choisi est l'acide 5-hydroxyindole-3-acétique (5-HIAA). Cette molécule est le métabolite principal de la sérotonine, aussi connu sous le nom de 5-hydroxytryptamine (5-HT). Provenant d'un neurotransmetteur essentiel et universel chez l'humain, le 5-HIAA est excrété de manière universelle chez tous les humains, sans grande variabilité spatiale et temporelle. Toutefois, il existe une variabilité individuelle non négligeable, pouvant être causée par certaines conditions médicales telles que la

dépression, des migraines sévères répétées, la présence de tumeurs carcinoïdes et non carcinoïdes, la fibrose kystique et des troubles de malabsorption (University of Rochester Medical Center, 2023). Cependant, l'effet potentiel causé par ces conditions et la prévalence de ces conditions ne sont pas assez élevés pour avoir un effet important à l'échelle populationnel en ce qui concerne les taux d'excrétion du 5-HIAA. Le composé parent du 5-HIAA, soit la sérotonine, est aussi présent chez d'autres animaux faisant partie du taxon *bilateria*, mais le métabolite humain, le 5-HIAA, n'a pas d'autres sources connues (Shi *et al.*, 2017). Ce métabolite est aussi connu pour sa stabilité dans les eaux usées, facilitant sa détection (Thai *et al.*, 2019). Le 5-HIAA est excrété à travers l'urine, ce qui diffère de l'excrétion à travers la matière fécale des deux premiers marqueurs populationnels, mais cela est intentionnel ; son mode d'excrétion est une des raisons pour lesquelles le 5-HIAA a été choisi. Ceci a été fait pour garder une certaine diversité parmi les marqueurs populationnels et simultanément apporter un point de vue préliminaire sur les différences de performance entre les marqueurs populationnels excrétés à travers l'urine et ceux excrétés à travers la matière fécale sans en faire l'objectif du projet.

Le dernier marqueur populationnel qui a été choisi est la caféine. Ce marqueur est le seul choisi qui n'est pas un marqueur qui est retrouvé naturellement chez l'humain et qui dépend de la consommation de certains produits par les individus. Malgré cela, la caféine est parmi les drogues les plus consommées et les plus répandues sur la planète, offrant tout de même une opportunité à plusieurs régions de l'utiliser comme marqueur populationnel si celui-ci est efficace. Le choix de la caféine est en partie dû à la facilité de l'analyse en tandem avec le 5-HIAA, les deux étant excrétés à travers l'urine et pouvant être analysés avec la même méthode, mais aussi dû à ces caractéristiques acceptables comme marqueur populationnel selon les critères utilisés. Cette molécule est stable dans l'eau usée et est facilement détectable (Chen *et al.*, 2014). De plus, mis à part les boissons caféinées non consommées qui sont déversées dans la STEP, il n'y a pas d'autres sources de contaminations connues (Gracia-Lor *et al.*, 2017). Son seul inconvénient est sa variabilité spatiale et individuelle. La consommation des boissons et d'autres produits caféinés varie de personne en personne et varie aussi au niveau régional (Chen *et al.*, 2014). Toutefois, considérant que la variabilité

temporelle est minimale et que le but de son utilisation n'est pas d'estimer la taille de la population, mais bien de normaliser les concentrations du SRAS-CoV-2, il a été jugé que cette molécule était un candidat acceptable. Aussi, la caféine est une molécule très facile à analyser, ce qui fait en sorte qu'elle pourrait constituer une manière simple, facile, rapide et surtout accessible de normaliser les concentrations du SRAS-CoV2 pour des municipalités qui possèdent moins de ressources matérielles, financières ou humaines si son efficacité est adéquate. Le métabolite principal de la caféine, la paraxanthine, a été considéré. Cependant, celui-ci se retrouve généralement en plus petites quantités dans les eaux usées et risque de parfois ne pas être détecté par la méthode d'analyse utilisée puisque sa concentration peut se retrouver sous la limite de détection (Hsu *et al.*, 2022).

CHAPITRE 2

NORMALISATION DU SRAS-COV-2 DANS LES EAUX USÉES

L'article « Biomarkers as a normalisation tool for Wastewater-Based Epidemiology to correlate with COVID-19 case data » a été rédigé avec l'aide des professeurs Marc-Denis Rioux et Richard St-Louis de l'UQAR. Cet article sera soumis dans *Water Research*, un journal revu par les pairs, pour sa publication. Ma contribution dans la production de cet article au-delà de la rédaction de l'article inclut aussi la conception des expériences et les manipulations pour les réaliser ainsi que le traitement des données, leur interprétation et la création des figures. Le professeur Marc-Denis Rioux a contribué à la conception des expériences, l'échantillonnage et la révision de l'article. Le professeur Richard St-Louis a contribué à la conception des expériences et la révision de l'article également.

Les travaux décrits dans cet article ont également été présentés en partie à l'oral et par affiche dans le passé dans plusieurs congrès et rencontres scientifiques. Cela inclut le *Canadian Ecotoxicity Workshop* (CEW) à Winnipeg en 2022, la rencontre annuelle du *Society of Environmental Toxicology And Chemistry* à Pittsburgh en 2022, le colloque de Chapitre Saint-Laurent à Québec en 2023 et à la journée annuelle du Collectif de Recherche Appliquée aux Bioprocédés et à la chimie de l'Environnement en 2024.

2.1 RÉSUMÉ EN FRANÇAIS

L'épidémiologie basée sur les eaux usées (ÉBEU) est utilisée pour estimer les cas cliniques de COVID-19 dans plusieurs régions à travers le monde. Pour ce faire, les concentrations du SRAS-CoV-2 dans les eaux usées sont normalisées avec un biomarqueur pour améliorer les corrélations entre les données d'eaux usées et les données cliniques. Le virus de la marbrure légère du poivron (PMMoV) fut un biomarqueur populaire pour la

normalisation, mais produit des résultats mitigés. Notre objectif est donc d'identifier et quantifier différents biomarqueurs pouvant avoir une meilleure performance dans la normalisation du SRAS-CoV-2 dans les eaux usées. Six critères ont été utilisés pour la sélection des biomarqueurs : la voie d'excrétion, la détectabilité, la stabilité, l'unicité, la variation spatio-temporelle et la méthode d'analyse. Les biomarqueurs sélectionnés sont le gène de la sous unité 5 du NADH déshydrogénase de l'ADN mitochondrial humain (mth-ND5), le gène CPQ_056 du virus bactériophage *Carjivirus communis* (CrAss), la caféine et l'acide hydroxyindoleacétique (5-HIAA). Les biomarqueurs ainsi que le SRAS-CoV-2 ont été analysés par réaction en polymérase en chaîne quantitative en temps réel (qPCR) et par chromatographie gazeuse couplée avec un spectromètre de masse (GC-MS). En utilisant les concentrations de SRAS-CoV-2 dans les eaux usées, les biomarqueurs, le débit journalier de la station d'épuration d'eau de la ville de Rimouski (pop. $\pm 50\ 000$) et une estimation du taux d'excrétion virale, plusieurs modèles d'estimation de nouveau cas actifs ont été créés par des régressions linéaires. La performance des modèles est calculée et comparée en utilisant des corrélations de Spearman. Ces modèles démontrent que le CrAss, le mth-ND5, le 5-HIAA et le débit journalier sont tous capables de mieux normaliser les concentrations du SRAS-CoV-2 que le PMMoV. L'utilisation de ces nouveaux biomarqueurs en ÉBEU mènerait à une estimation plus exacte des cas actifs de COVID-19 dans la population. Ceci aiderait les autorités de santé publique dans la prise de décision rapide et la gestion de la pandémie actuelle et dans le futur.

2.2 BIOMARKERS AS A NORMALISATION TOOL FOR WASTEWATER-BASED EPIDEMIOLOGY TO CORRELATE WITH COVID-19 CASE DATA

Authors: Elliott O. H. Chartrand^{1*}, Marc-Denis Rioux², Richard St-Louis¹

¹ Département de biologie, chimie et géographie, Université du Québec à Rimouski, 300 allée des Ursulines, Rimouski (Québec), G5L 2Z9, Canada

² Département de mathématiques, informatique et génie, Université du Québec à Rimouski, 300 allée des Ursulines, Rimouski (Québec), G5L 2Z9, Canada

* Corresponding author :

E. O. H. Chartrand

Elliottchartrand@hotmail.ca

2.3 ABSTRACT

Wastewater-based epidemiology (WBE) is used, partly, to estimate the number of new active cases of COVID-19 in multiple regions across the globe. To do so, the concentrations of the virus, SARS-CoV-2, in wastewater are normalized using a biomarker to increase the correlation between the wastewater and clinical data. The pepper mild mottle virus (PMMoV) is a popular biomarker used in this normalisation but has mitigated results. The objective of this study is to identify and quantify different biomarkers that could outperform PMMoV in the normalisation of the SARS-CoV-2 concentrations in wastewater of a small city (pop. ~ 50 000). Six criteria were used in the selection process: excretion pathway, detectability, stability, uniqueness, temporal and spatial variation, and analysis method. The four chosen biomarkers are caffeine (CAF), 5-hydroxyindolacetic acid (5-HIAA), *Carjivirus communis* gene CPQ_056 (CrAss), and human mitochondrial DNA gene for the sub-unit 5 NADH dehydrogenase (mth-ND5). The biomarkers and SARS-CoV-2 concentrations were obtained via Real-Time quantitative PCR and GC-MS. Using the biomarker and SARS-CoV-2 concentrations along with the daily flowrate of the wastewater treatment sampled and a viral shedding rate model, multiple models were constructed using linear regressions to estimate new active cases based on the wastewater data. The performance of these models was

evaluated using Spearman correlations and compared to models using PMMoV. The results show CrAss, mth-ND5, 5-HIAA, and daily flowrate all increase the correlations between the SARS-CoV-2 concentrations and the clinical cases more than PMMoV. Using these biomarkers in WBE leads to a more exact estimation of new COVID-19 cases in the catchment population. This will help health authorities in their decision making and overall management of the current and all future pandemics.

2.4 INTRODUCTION

Wastewater-based epidemiology (WBE) has gained significant traction as a vital surveillance and monitoring tool since the onset of the SARS-CoV-2 pandemic (Hrudey *et al.*, 2020; Rioux *et al.*, 2023). WBE offers numerous advantages over the classic clinical testing; minimal resource, no-need for specific screening of individuals, passive surveillance of cases, scalable early warning system, and capacity to detect various pathogens (Hutchison *et al.*, 2022; Daughton, 2020; D'Aoust *et al.*, 2021; Kumar *et al.*, 2021; Patey *et al.*, 2021). These aspects of WBE are particularly relevant during a massive outbreak like the COVID-19 pandemic, where timely information is essential for decision-making. However, despite recent advancements in WBE, there remains a pressing need for more accurate, usable, and accessible data modeling.

For WBE to function as a reliable predictor and estimator prevalent and incident COVID-19 cases within a city, normalisation of virion load is imperative. The primary challenge to do this lies in identifying an effective population-based biomarkers to normalise the SARS-CoV-2 concentrations in wastewater (Hutchison *et al.*, 2022; Greenwald *et al.*, 2021; Ahmed *et al.*, 2021; Hsu *et al.*, 2022). The Centers for Disease Control and Prevention recommend the use of pepper mild mottle virus (PMMoV) and viruses from the Crassvirales order, previously referred to as CrAssphage (Dutilh *et al.*, 2014). The effectiveness of PMMoV for normalisation remains a topic of ongoing debate (Maal-Bared *et al.*, 2023). While it has been shown to enhance the correlation between WBE data and clinical case numbers in some instances (D'Aoust *et al.*, 2021; Maal-Bared *et al.*, 2023), others show no

statistically significant effects or a significant reduction in the correlation with case data (Hsu *et al.*, 2022; Li *et al.*, 2024; Maal-Bared *et al.*, 2023; Rioux *et al.*, 2023).

The objective of the present study is to identify potential population-based biomarkers suitable for application across diverse regions capable of wastewater-based epidemiology and to evaluate their performance in a smaller municipality (pop. ~ 50 000) against the current benchmark, PMMoV. By conducting this comparison in a smaller city rather than a large metropolitan area, we can minimise sources of variability, reduce the residence time of biomarkers in wastewater, and better highlight the population's impact on the wastewater treatment plant (WWTP). The selected biomarkers were chosen based on six critical criteria: excretion pathway, detectability in wastewater, stability in wastewater, human specificity, variability, and analytical feasibility. Each criterion serves as a filter to narrow down the pool of viable biomarkers. The excretion pathway influences quantification methods and ensures the biomarker's presence in wastewater. Good detectability in wastewater is needed for accurate quantification. Stability in wastewater is essential for reliable measurement, minimising degradation-related variability. Human specificity further reduces variability by exogenous sources. Variability in all forms—temporal, spatial, and individual, including fluctuations in individual excretion rates—should be minimised to ensure that the variation in quantification data is attributable to the wastewater system and population-related factors. Additionally, the method of quantification must be scrutinised, as some are faster, more accessible, or more reliable than others.

The four biomarkers selected are caffeine (CAF), 5-hydroxyindolacetic acid (5-HIAA), *Carjivirus communis* gene CPQ_056 (CrAss), and human mitochondrial DNA gene for the sub-unit 5 NADH dehydrogenase (mth-ND5). With the use of these biomarkers, the SARS-CoV-2 Nucleocapsid 2 gene (N2) and other factors such as WWTP daily flowrate (Flow), a normalisation tool will be developed to help increase the correlation between WBE data and clinical case data. This will further help public health authorities by supplying a cost-effective and accurate early warning surveillance system to aid in their timely decision making in the context of a pandemic.

Additionally, the dynamics of viral shedding through excretion is a critical element that warrants attention (Puhach *et al.*, 2023; Rioux *et al.*, 2023). An infected individual will continue to excrete viral particles throughout their infectious period, not solely on the day their positive test result contributes to the case data (Puhach *et al.*, 2023; Rioux *et al.*, 2023). Consequently, viral shedding must be incorporated into models to improve their accuracy (Rioux *et al.*, 2023). Moreover, the population case data employed plays a significant role. Ideally, this data should represent the sampled sewershed and be gathered in real-time through clinical testing. However, this data has become increasingly scarce as public health entities have scaled back extensive clinical testing. In the absence of high-quality data, epidemiological estimates and models are often employed as substitutes but may yield misleading conclusions if not adequately modelled, particularly if the dynamics of viral shedding rates are overlooked (Rioux *et al.*, 2023). Ultimately, these considerations underscore the necessity for a standardised method and biomarker for the normalisation of WBE data in the effective monitoring and surveillance of pathogens.

2.5 MATERIALS AND METHODS

2.5.1 Biomarker selection process

From the literature available in 2022, a total of 43 peer-reviewed articles were sifted through and 48 potential biomarkers were identified from those articles. These articles were found through Web of Science using the key words “biomarker” and “wastewater-based epidemiology” and therefore not exclusively linked to the COVID-19 pandemic. The six criteria mentioned above were used to make the final selection. The four chosen biomarkers all exhibited outstanding qualities for each criterion. The principal considered biomarkers, and their respect of the criteria used, is summed in Table 1 along with the most popular method used for their quantification and one reference as an example associated with each biomarker.

Table 1 List of potential biomarkers considered, and their respect of the criteria used

Name	Type	Excretion Pathway	Stability	Exogenous sources	Variability	Detectability	Method	Example Reference
PMMoV	RNA	Feces	High	Food, Agriculture	Moderate	High	RT-qPCR	D'Aoust <i>et al.</i> , 2021
Crassvirales	DNA	Feces	High	None	Low	High	qPCR	Ahmed <i>et al.</i> , 2020b
Human Polyomavirus	DNA	Feces	High	None	Low	High	qPCR	Bofill-Mas <i>et al.</i> , 2006
Adenovirus	DNA	Feces	High	None	Moderate	High	qPCR	Bofill-Mas <i>et al.</i> , 2006
Eukaryotic RNA	RNA	Feces	Low	None	Moderate	Moderate	RT-qPCR	Greenwald <i>et al.</i> , 2021
Mitochondrial DNA	DNA	Feces	High	None	Low	High	qPCR	Hutchison <i>et al.</i> , 2022
Human Bacteroides	DNA	Feces	Low	Animals	Moderate	High	qPCR	Greenwald <i>et al.</i> , 2021
<i>E. coli</i>	Bacteria	Feces	N/A	Environnement	Moderate	High	CFU count	Tolouei <i>et al.</i> , 2019
<i>C. perfringens</i>	Bacteria	Feces	N/A	Environnement	Moderate	High	CFU count	Tolouei <i>et al.</i> , 2019
Acesulfame	Sweetener	Urine	High	Agriculture, Animal feed	Moderate	High	LC-MS	Kokotou <i>et al.</i> , 2013
Cyclamate	Sweetener	Urine	High	Agriculture, Animal feed	Moderate	High	LC-MS	Kokotou <i>et al.</i> , 2013
Saccharine	Sweetener	Urine	High	Agriculture, Animal feed	Moderate	High	LC-MS	Kokotou <i>et al.</i> , 2013
Sucralose	Sweetener	Urine	High	Agriculture, Animal feed	Moderate	High	LC-MS	Kokotou <i>et al.</i> , 2013
Alitame	Sweetener	Urine	N/A	Agriculture, Animal feed	Moderate	Low	LC-MS	Kokotou <i>et al.</i> , 2013
Aspartame	Sweetener	Urine	Low	Agriculture, Animal feed	Moderate	Low	LC-MS	Kokotou <i>et al.</i> , 2013
Neotame	Sweetener	Urine	N/A	Agriculture, Animal feed	Moderate	Low	LC-MS	Kokotou <i>et al.</i> , 2013
Neohesperidin dihydrochalcone	Sweetener	Urine	Low	Agriculture, Animal feed	Moderate	Low	LC-MS	Kokotou <i>et al.</i> , 2013
Caffeine	Drug	Urine	High	Food, Beverages	Moderate	High	LC-MS	Senta <i>et al.</i> , 2014
Nicotine	Drug	Urine	High	Cigarette butts, pharmaceuticals	Moderate	High	LC-MS	Senta <i>et al.</i> , 2014
Creatinine	Metabolite	Urine	Low	Animals, Animal products, Food	Moderate	Moderate	LC-MS	Chen <i>et al.</i> , 2014
Serotonin	Metabolite	Urine	High	None	Low	Moderate	LC-MS	Chen <i>et al.</i> , 2014
Ammonium	Other	Urine	High	Agriculture	Low	High	Ammonia kit	Hutchison <i>et al.</i> , 2022

2.5.2 Sampling and treatment

Sampling was conducted in Rimouski, Québec, Canada, from January 2022 to May 2022, covering the winter and spring period, which matches with the end of the 5th wave and covers the 6th wave of the pandemic in the Bas-Saint-Laurent region of Québec. The Rimouski WWTP services a population of approximately 50 000, with an estimated average daily flow around 30,000 cubic metres, varying between 17 000 and 76 000. Composite samples were collected over 24-hour periods from the influent of the aeration basin using one-litre opaque high-density polyethylene bottles. This sampling was performed by the city of Rimouski on average three times per week—on Mondays, Wednesdays, and Fridays—yielding a total of 54 one-litre samples. The filled bottles were maintained at 4°C during transport to the research facility, where they were subsequently frozen at -80°C until processing. Storage in these conditions does not incur sample degradation for the targets (Hokajärvi *et al.*, 2021). Prior to treatment, samples were thawed at 20°C for 8 to 9 hours in a dark environment. Once thawed and thoroughly agitated, 625 µL of 2M MgCl₂ was added to each 50 mL aliquot of wastewater. This increased the retention of viral particle on the filters via salt-bridging (Wallis *et al.*, 1972; Lukasik *et al.*, 2000). The resulting mixture was then filtered through plain 47mm mixed cellulose ester filters with a porosity of 0.45 µm using a vacuum pump (Fisher, Ottawa, Canada). This filtration process was repeated six times for each 1L sample. Two filters were designated for the analysis of DNA biomarkers (CrAss and mth-ND5), two for RNA biomarkers (PMMoV and N2), one filter was archived and stored at -80°C, and the final filter was treated with distilled water as a blank control. The filtrate was collected in a conical flask and stored at 4°C until processing for the analysis of chemical biomarkers (5-HIAA and CAF).

The extraction of genetic material was conducted simultaneously for RNA and DNA using the RNeasy Power Microbiome Kit from QIAGEN. Filters were excised and placed into 1.5 mL centrifuge tubes, into which 20 µL of bovine respiratory syncytial virus (BRSV) and 5 µL of an internal positive control (IPC) (Kavlick, 2018) were introduced to the RNA and DNA filters, respectively, serving as method controls. Both controls were also applied

to the blank filter. The samples underwent DNA and RNA extraction according to a modified protocol based on the manufacturer's guidelines, aimed at minimising waste byproducts and ensuring optimal viral recovery from wastewater. This extraction, with a recovery rate of 75%, yielded 100 μ L of genetic material per filter which was then frozen at -80°C in 1.5 mL tubes until quantification.

Chemical biomarkers were extracted using solid-phase extraction (SPE) based on previous works (Shi *et al.*, 2017; Hsu *et al.*, 2022). As surrogate, 0.5 μ g of caffeine-trimethyl-d₉ and 0.2 μ g of 5-HIAA-d₆ (Sigma-Aldrich, Toronto, Canada) were added to the filtrate prior to the SPE process. Oasis HLB 3cc cartridges from Waters were employed for the extraction of 5-HIAA, while Supelclean LC-18 500 mg cartridges from Supelco were utilised for caffeine extraction (Sigma-Aldrich, Toronto, Canada). The cartridges were conditioned with 6 mL of methanol followed by 3 mL of nanopure water. Subsequently, 100 mL of the sample filtrate was loaded onto the cartridges at a flow rate of 10 mL/min. The biomarkers were eluted with 4 mL of methanol and the eluate stored at -20°C until quantification. The recovery rates were approximately $92.5\% \pm 2\%$ and $19.5\% \pm 7\%$ for CAF and HIAA respectively.

2.5.3 Quantification

The quantification of the genetic biomarkers, controls, and SRAS-CoV-2 was done using an Opus FX 96 well thermocycler from Bio-Rad. RNA and DNA were amplified separately via RT-qPCR and qPCR respectively. For each filter, the amplification was done in triplicates, resulting in six wells per sample for each RT-qPCR and qPCR reaction. Reaction mixtures for RT-qPCR contained 5 μ L of sample RNA, 10 μ L of $2 \times$ Luna[®] Universal Probe One-Step RT-qPCR reaction mix (BioLabs Inc., New England), 0.25 μ M of each primer, 0.125 μ M of each probe, 1 μ L of RT enzyme mix and nanopore RNase-free water for a total well volume of 20 μ L. The thermal cycling conditions were an initial RT denaturation at 55°C for 10 minutes, 5 minutes at 95°C for denaturation, then 40 cycles comprised of 10 seconds at 95°C and 30 seconds at 60°C . The amplification was done

partially in multiplex, with BRSV and PMMoV targets done together while N2 done separately from the former two. Calibration curves were done with each plate. Reactions mixtures and cycling conditions for the DNA amplification via qPCR were similar with minor differences. The mixture did not include RT enzyme, it being replaced with water, and used 10µL of 2 × Luna® Universal Probe One-Step qPCR reaction mix (BioLabs Inc., New England) for the reaction mix. As for the cycling conditions, the initial RT denaturation process was excluded, with the rest of the conditions being identical. All target sequences and references are listed in Table 2.

Table 2 Genetic biomarker target sequences

Target	Types	Sequences	Reference
BRSV	Forward primer	GCA ATG CTG CAG GAC TAG GTA TAA T	Boxus <i>et al.</i> , 2005
BRSV	Reverse primer	ACA CTG TAA TTG ATG ACC CCA TTC T	
BRSV	Probe Cy5	Cy5 ACC AAG ACT /ZEN/ TGT ATG ATG CTG CCA AAG CA IABkRQ	
PMMoV	Forward primer	TAC TTC GGC GTT AGG CAA TCA G	Kitajima <i>et al.</i> , 2018
PMMoV	Reverse primer	TGA AAC CAG TAG CAG GAA ATC TAA	
PMMoV	Probe HEX	HEX CAG CAG TTC /ZEN/ TCT GAT GTG TGG IABkFQ	
nCoV N2	Forward primer	TTA CAA ACA TTG GCC GCA AA	CDC, 2020
nCoV N2	Reverse primer	GCG CGA CAT TCC GAA GAA	
nCoV N2	Probe TEX615	TEX615-ACA ATT TGC /ZEN/ CCC CAG CGC TTC AG-3IABkRQ	
CrAss	Forward primer	CAG AAG TAC AAA CTC CTA AAA AAC GTA GAG	Stachler <i>et al.</i> 2017
CrAss	Reverse primer	GAT GAC CAA TAA ACA AGC CAT TAG C	
CrAss	Probe FAM	FAM AAT AAC GAT TTA CGT GAT GTA AC MGB	
IPC	Forward primer	CGCGAGATAC ACTGCCAGAA	Kavlick <i>et al.</i> 2018
IPC	Reverse primer	GACCACAGCC AGATTAAATT TACCA	
IPC	Probe HEX	HEX-TCCGCGTGAT TACG-BHQ1	
ADNmth	Forward primer	CAG CAG CCA TTC AAG CAA TGC	Hutchison <i>et al.</i> , 2022
ADNmth	Reverse primer	GGT GGA GAC CTA ATT GGG CTG ATT AG	
ADNmth	Probe FAM	FAM-TAT CGG CGA TAT CGG TTT CAT CCT CG-BHQ-1	

The quantification of the chemical biomarkers was done by GC-MS with an Agilent Technologies 8890 gas chromatogram coupled with a 5977B mass spectrometer system.

Electron ionization with positive ion detection was used to analyse all molecules between 50 and 550 m/z. The column used was a silica capillary column by Restek (Rxi-5ms, 30m, 0,25mmID, 0,25um). For the thermal profile, the GC oven was held at 70°C for 7 minutes, followed by a ramp up to 325°C at a rate of 20°C/min at which that temperature was held until the end of the run at 28 minutes. Prior to the splitless injection, derivatization with trimethylsilyl of 5-HIAA into 5-HIAA-3-TMS is necessary to assure the volatility and thermal stability of the analyte, making its identification and quantification possible via GC-MS. Thus, the MeOH extracts were evaporated under a nitrogen stream at room temperature and reconstituted with 80µL of a toluene and hexane mixture 1:1. Then, 80µL of BFTSA + TMCS 1% (Sigma-Aldrich, Toronto, Canada) was added directly to the samples which were then heated to 90°C for 60 minutes. After derivatization, samples were cooled to room temperature before injection.

2.5.4 Statistical analysis and modelling

Statistical analysis and data manipulation were performed using R version 4.3.3 (R Core Team, 2024) in conjunction with RStudio. The shedding rate modelling, based on the work of Rioux *et al.* (2023), was executed with OpenSees (McKenna *et al.*, 2010) and subsequently imported into R for further analysis. Linear regression models were employed, and Spearman's rank correlations (Spearman, 1904) were calculated to compare each model's performance. All models used the concentrations of the SARS-CoV-2 N2 gene target in wastewater as the primary variable against which case data were correlated. The anonymous case data were sourced from the Rimouski Regional Hospital, representing the incidence of cases among all patients under its care. Time-lag was not included in the models due to the nature of the sampling; a lack of sample resolution makes defining an accurate time-lag difficult. Therefore, it was omitted.

2.6 RESULTS

2.6.1 Marker concentrations

Each biomarker, including N2, was quantified in wastewater sampled from January 21st to May 24th, 2022 (Figure 9). As anticipated, variations over time were observed for each biomarker, although the degree of fluctuation differed significantly among them. CrAss (B1-B2) exhibited the least variation among the five biomarkers measured, whereas PMMoV (D1-D2) and CAF (E1-E2) demonstrated substantial temporal fluctuations. In contrast, 5-HIAA (F1-F2) and mth-ND5 (C1-C2) displayed moderate levels of variation. With the daily flow rate of the WWTP increasing during the spring, the observed downward trend in concentrations is expected (1). This trend is evident across all measured concentrations of biomarkers but is particularly pronounced in CrAss (B1) and mth-ND5 (C1).

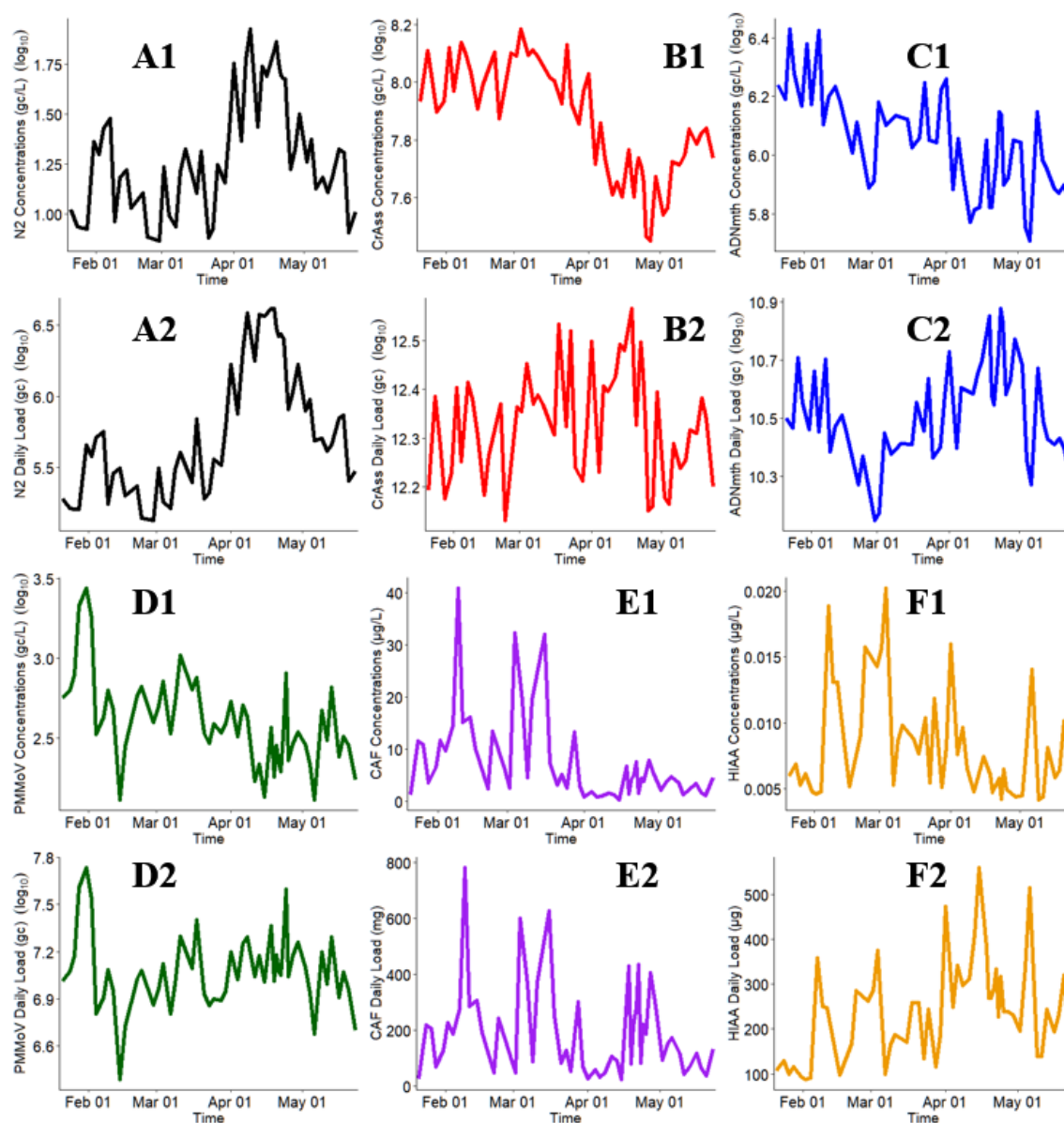


Figure 9. SARS-CoV-2 N2 gene and biomarker concentrations (A1-F1) and daily loads over time (A2-F2). A1-2: SARS-CoV-2 N2 gene concentrations and daily load. B1-2: CrAss concentrations and daily load. C1-2: mth-ND5 concentrations and daily load. D1-2: PMMoV concentrations and daily load. E1-2: Caffeine concentrations and daily load. F1-2: 5-HIAA concentrations and daily load.

2.6.2 Case Data

The anonymous case data used in this study was provided by the Rimouski Regional Hospital as population-wide case data from mass clinical testing was unavailable. This represents the case data of patients that were admitted into care during the study's period (HR). The virus shedding dynamics developed by Rioux *et al.* (2023) were used to generate a new set of case data, depicting a more realistic state of estimated active cases (HROS). Both data sets vary similarly, with some differences in intensity and amplitude (Figure 10). The end of January 2022 shows the tail end of the 5th wave during the pandemic. In the month of April, both data sets show the 6th wave of the COVID-19 pandemic with an increase in active cases during that time.

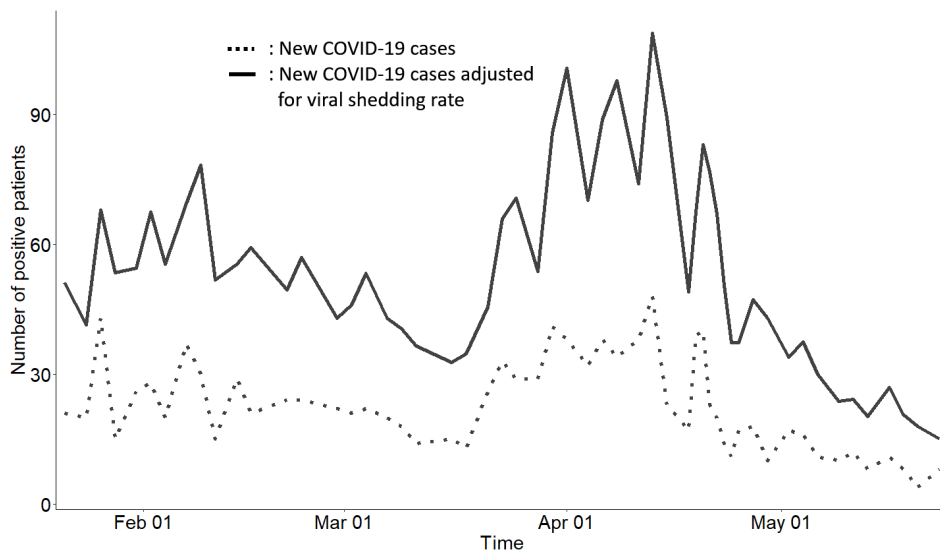


Figure 10. Case data from the study period. Dotted line: HR, number of patients who tested positive for COVID-19 in the Rimouski Regional Hospital per day. Line: HROS, shedding rate modelling based on the number of patients who tested positive per day.

2.6.3 Modelling

The correlation between all variables can be observed within Figure 11. Certain variables are expected to have little to no correlations, as other are expected to be highly correlated. Flow is expected to be negatively correlated with every WW target, such as N2

and every biomarker measured (Polo *et al.*, 2020). This is what is observed for every biomarker analysed here except for N2. Another expected correlation is the correlation between both case data sets, HR and HROS. As HROS is based on HR, they are highly correlated, as they should. All three genetic biomarkers (PMMoV, CrAss, and mth-ND5) are positively correlated with each other, as expected due to their common excretion pathway. Moreover, none of the biomarkers are correlated on their own with case data. This is favourable, as biomarkers that are independent of the case data is preferred to normalise viral loads in WW.

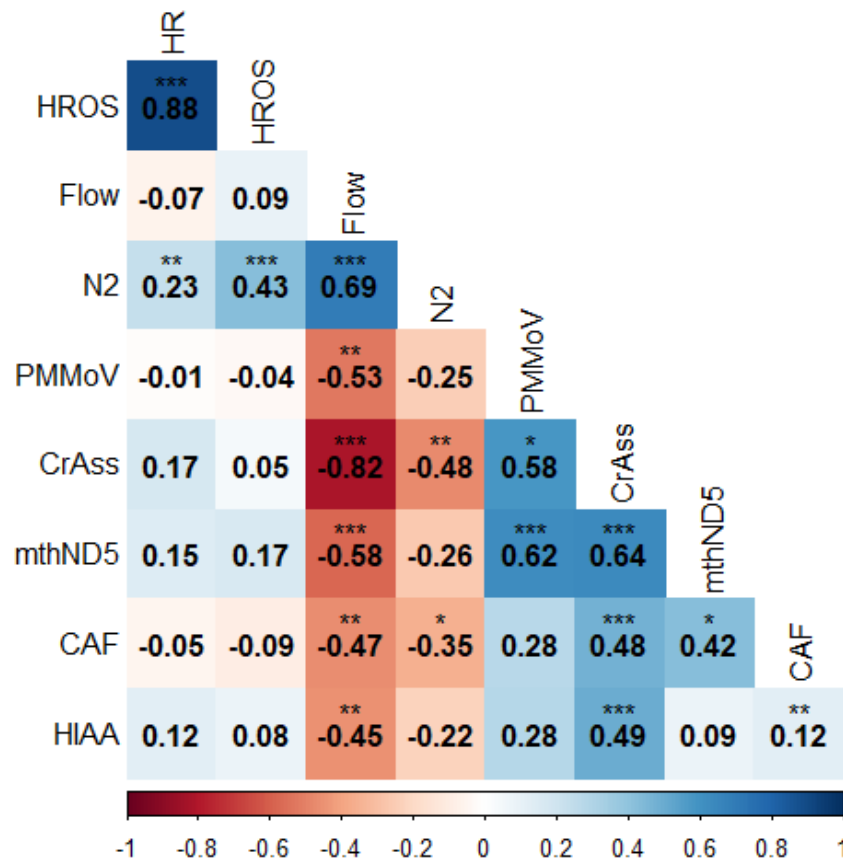


Figure 11. Spearman Correlation Matrix for all variables used for modelling, including case data. HR (Hospital Reported cases), HROS (Hospital Reported cases with Shedding dynamics), Flow (daily flowrate from the WWTP), N2 (SARS-CoV-2 N2 gene target). Significant correlations are noted with * < 0.05 , ** < 0.01 , and *** < 0.001 .

In total, eleven linear models were systematically developed, incorporating N2, Flow, and the five biomarkers. Each model combined N2 with either one of the biomarkers or Flow as a secondary variable, or a conjunction of Flow and an additional biomarker as the second and third variables. These models, along with both case datasets, were represented in a correlation matrix plot to elucidate which variables influenced correlations (Figure 12). Two significant trends emerged from the analysis. Firstly, HROS consistently augmented correlations with wastewater data, regardless of the biomarkers employed. Secondly, the observed horizontal patterns indicated that the choice of biomarker or biomarker combinations was the principal determinant of a model's correlation strength. Models incorporating PMMoV or CAF demonstrated non-significant correlations, except for the N2_PMMoV model with HROS. In contrast, the highest correlations were obtained from models utilising a single variable apart from N2, specifically Flow, CrAss, or mth-ND5. These models exhibited similar correlation coefficients, ranging from 0.40 to 0.47 for HR and between 0.53 and 0.54 for HROS. Models using those biomarkers and using 5-HIAA surpassed the correlation observed with N2 alone across both case datasets. Furthermore, the introduction of additional variables in three-variable models did not yield any further enhancement in correlation.

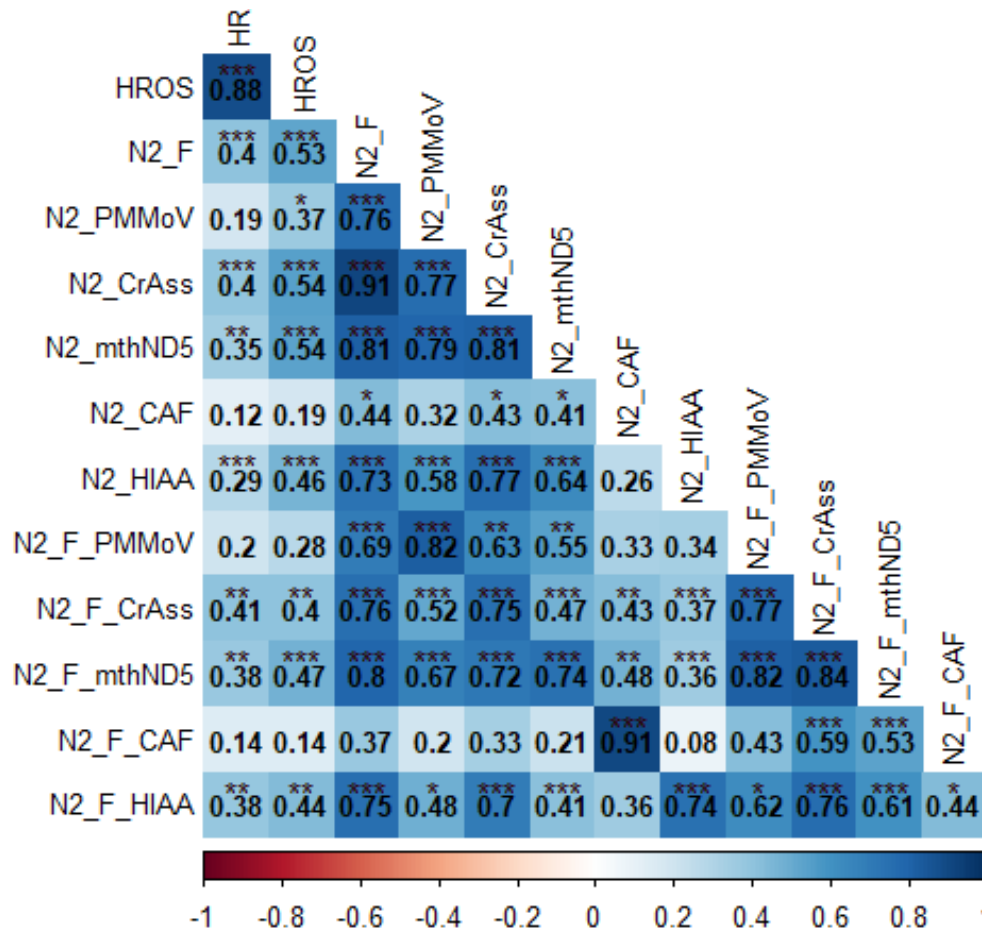


Figure 12. Spearman Correlation Matrix for all models and case data sets. Significant correlations are noted with * (<0.05), ** (<0.01), and *** (<0.001).

Each model was plotted over HROS case data to observe the relationship between the trends of each model in relation to the case data (Figure 13). All models behave similarly; their output increases as the number of cases increases and go back down once the cases drop. Some models increase and decrease more rapidly, creating a high and narrow peak (H: N2-Flow-CAF) while others do so more slowly, having more rounded peaks that last longer in time (B: N2-Flow, F: N2-Flow-CrAss, J: N2-Flow-ND5mth).

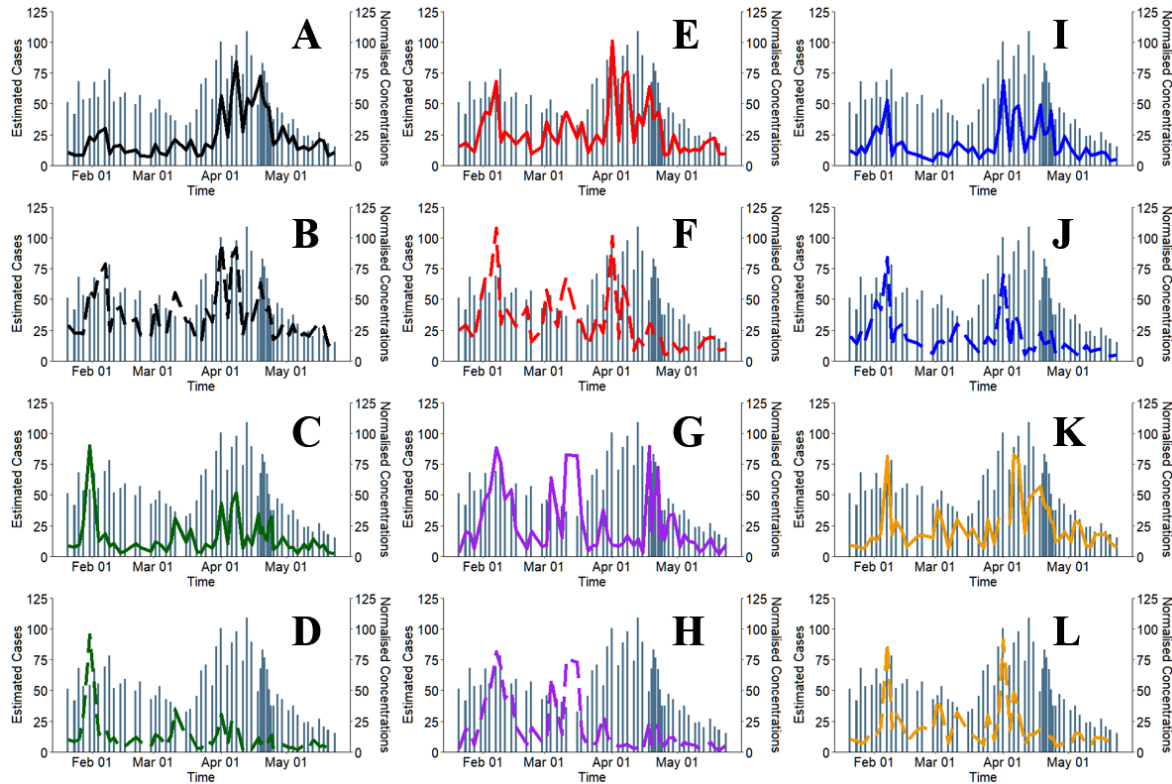


Figure 13. WBE Models with Case Data (HROS) over time. Full lines are models excluding Flow; dotted lines are models including Flow. Colours are based on the biomarkers used. A: N₂, B: N₂_F, C: N₂_PMMoV, D: N₂_F_PMMoV, E: N₂_CrAss, F: N₂_F_CrAss, G: N₂_CAF, H: N₂_F_CAF, I: N₂_mthND5, J: N₂_F_mthND5, K: N₂_HIAA, L: N₂_F_HIAA.

2.7 DISCUSSION

The use of populational biomarkers to normalise SARS-CoV-2 viral loads in wastewater is a crucial step in the use of WBE which can either improve or weaken correlation with case data. Previous studies can help narrow down the choices of appropriate biomarkers but testing them in real time to verify their efficacy is crucial. Not all biomarkers perform properly in every region or throughout the year. This is evident with PMMoV. Here, PMMoV decreases the correlation between WBE data and case data, with or without considering flowrate (Figure 12). PMMoV not only reduces correlation but also turns a significant relationship into a non-significant relationship most of the time it is included in a model (Figure 12). This differs from previous studies which conclude PMMoV increase such correlations and helps with WBE modelling in the context of the COVID-19 pandemic

(D'Aoust *et al.*, 2021; Mitranescu *et al.*, 2022). However, other studies have found the performance of PMMoV as a normalization tool can vary with sampling site based on WWTP flowrate (Wartell *et al.*, 2022) or other factors such as population diet, the precision and accuracy of different laboratories, or size of sewersheds (Rabe *et al.*, 2023; Maal-Bared *et al.*, 2023). Some studies concluded that PMMoV did not perform well (Li *et al.*, 2024 ; Rioux *et al.*, 2023). This study adds to the growing number of findings reevaluating the performance of PMMoV and concluding it is inconsistent and a liability in case modeling.

One of the main criteria to evaluate the potential performance of a biomarker is variation, mainly temporal variation of the concentration in-sewer. This criterion is harder to determine or assess via literature alone, compared to excretion pathway or detectability in wastewater, which in most cases is already well documented. Over the study period, the concentration of all biomarkers varied through time as seen in Figure 9. This is expected due to the variation in the WWTP flowrate in the spring. Flowrate is partly dependent on rainfall and snowmelt as the WW system in Rimouski is a combined system, where stormwater and wastewater are transported together. In the spring, when all the snow which accumulated during the winter melts, the flowrate increases. This increase translates into a decrease in the concentrations of the biomarkers, as seen in Figure 8 B1-C1. Even with this seasonal variation, the biomarker which vary the least (CrAss and mth-ND5) are the ones which increases the correlation between WBE data and case data the most (Figure 12). This can also be seen with the variation of the biomarkers' daily load (Figure 9 A2-F2). This is not a coincidence, but rather the result of the theory behind normalization. A biomarker which has fewer external forces acting upon it, such as excretion variations amongst individuals and contamination from non-human sources, varies less over time. Therefore, it is a better indicator of the specific variation within the WWTP system, which affects all biomarkers and targets present, including SARS-CoV-2. A biomarker able to accurately portray that type of variation leads to a better normalization of the target pathogen, correcting the variation caused by WWTP (Gracia-Lor *et al.*, 2017). This leads to a better view of the other types of variations affecting the pathogen concentrations, such as populational case data. Flow works similarly, as flow is responsible and linked to a great number of possible variation factors in

the WWTP such as particle transport and dilution (Hrudey *et al.*, 2022; Mitranescu *et al.*, 2022).

The relationship between the biomarkers and the WWTP flowrate is demonstrated well by the spearman correlation (Figure 11). The strong negative correlation observed is expected. For N2 however, a positive correlation was measured. This may seem like an anomaly at first, but it is simply due to the timing of the 6th wave with the 2022 spring season. N2 concentrations are expected to rise and fall with case data, which is a relationship that was observed in this study. The highest peak of flowrate can be observed during the spring as the snow melts, which, here, happens around early and mid April for the study period (Annex II). This coincides with the highest peak of case data measured, which is in early and mid April, during the 6th wave of the COVID-19 pandemic (Figure 10). So, we see an independent increase in N2 concentrations due to case incidence and in flowrate due to seasonal variations. The positive correlation between those two variables, N2 and flowrate, is a coincidence. The relative increase in N2 concentration was much higher than the relative increase in flowrate, which contributes to the positive correlation observed between those two variables. The same phenomenon explains the negative correlations between the five biomarkers and N2 (Figure 11). As all biomarkers get diluted due to the flowrate, N2 increased drastically due to the rise in case data at the same time, leading to a significant negative correlation when no correlation should be observed. This highlights the importance of measuring flowrate and considering such seasonal variations when normalising WBE data to correlate to case data. The same correlation matrix shows the significant relationship between all three genetic biomarkers (PMMoV, CrAss, mth-ND5). This is an interesting result as it shows how similarly they perform based on their excretion pathway *i.e.* biomarkers of similar origins (*e.g.* genetic material) are more likely to behave and react the same way in wastewater compared to other biomarkers from a different origin (*e.g.* metabolites) (Dutilh *et al.*, 2014; Hutchison *et al.*, 2022; Kitajima *et al.*, 2018). Despite good correlations, there are small differences in their origin, variation, and excretion rate in wastewater (Hutchison *et al.*, 2022; Honap *et al.*, 2020; Kitajima *et al.*, 2018). CAF and 5-HIAA, regardless of having the same excretion pathway, are much more different from each

other in terms of provenance, excretion rate and variation in wastewater, while also having some exogenous sources for CAF, such as wasted or unfinished caffeinated drinks poured down the drain. These differences contribute to the weak correlation between these two biomarkers (Chen *et al.*, 2014; Hsu *et al.*, 2022).

From the Spearman correlation matrix between all models and case data sets (Figure 12), two trends can be noticed. Firstly, all the models correlated with HROS case data have an equal or higher correlation than when compared to HR case data. This increase in the correlation caused by the HROS case data varies, sometimes being small (N2_F_HIAA), sometimes being much higher (N2_mthND5). The only two models for which the HROS case data does not increase the correlation is N2_F_CAF and N2_F_CrAss, two models using a biomarker and flowrate. This trend shows the importance of considering the shedding rate of SARS-CoV-2 viral particles. As demonstrated in Rioux *et al.* (2023), this is a crucial step in building a more accurate model for WBE. Our study furthers that principle. When only using new COVID-19 cases for the case data, *i.e.* case incidence, there are many active cases not accounted for. Any person who was already tested and counted in the incidence but is still positive with COVID-19 will continue to excrete viral particles that can and will be detected by WBE (Patey *et al.*, 2021). There are still uncertainties regarding viral particle shedding in COVID-19 patients, which makes constructing a precise shedding rate model difficult (Wu *et al.*, 2020; Lowry *et al.*, 2023). A simple, theoretical model such as the one in Rioux *et al.* (2023) is sufficient to roughly represent the shedding rate of viral particles caused by the continuous infection of the virus in positive patients. This transforms the case incidence into an estimation of case prevalence. Case prevalence would be ideal for such modelling; however, case prevalence is extremely difficult to obtain for COVID-19 patients (Doernberg *et al.*, 2022). Not only are most healthcare systems not equipped to accommodate this form of data acquisition on large scale nor on a long-term basis, but most COVID-19 patients are not hospitalised where there is continuous testing (INSPQ, 2023). Therefore, most patients would have to go back to a testing point to obtain a negative test for the health authorities to have the necessary information to calculate case prevalence. This is why considering the viral shedding from new active patients is crucial. It is the quickest, most

efficient, and most accessible way to account for the continuous infection of COVID-19 patients, the continuous shedding of viral particles, and how it affects WBE data. This not only greatly improves the correlation of wastewater data and case data as seen in Figure 12, but also furthers our understanding of the many factors at play in WBE.

The second very noticeable trend in Figure 12 is the effect of the chosen biomarker used in the model on the correlations with case data. Visually there are two groups based mainly on their correlations and significance with case data. The first comprised of Flow, CrAss, mth-ND5, and 5-HIAA; the second comprised of CAF and PMMoV. These groups are also based on concentration variation through time, which is a good indicator of a biomarker's potential performance. This separation holds true for both case data sets used and regardless of the number of variables in the chosen model. When CAF or PMMoV are included in the model, it has poor performance compared to the other models and N2 itself, and it is no longer a significant relationship. This is consistent with many studies reevaluating the performance and pertinence of using PMMoV as a normalisation biomarker in WBE for the COVID-19 pandemic (Li *et al.*, 2024 ; Rioux *et al.*, 2023). The poor performance of CAF is unsurprising for the same reasons as PMMoV. Both are diet-based, having some variations in food consumption throughout different individuals, regions, and seasons (Gracia-Lor *et al.*, 2017). What made CAF such an attractive biomarker was its ease of analysis by most analytical laboratories. However, according to our results, its poor performance cannot be compensated by this technical advantage, and as such, should not be considered or used as a normalisation biomarker for WBE. However, much like PMMoV, some studies have found it could perform well in certain cases (Oloye *et al.*, 2023), but its inconsistency is a major drawback. CrAss, mth-ND5, 5-HIAA, and flowrate all tend to increase the correlation of WBE data with case data. The highest correlations observed are from the models using CrAss, mth-ND5, and flowrate, all very similarly (Figure 12). This is true between the models using only one biomarker at once, and between the models using flowrate and a biomarker simultaneously. Those biomarkers and flowrate have a similar effect on the increase in the correlation, suggesting they all have similar performances in this setting. Moreover, when using flowrate in conjunction with either CrAss, mth-ND5, or 5-HIAA, the effect on the

correlation is once again very similar between those biomarkers, suggesting similar normalisation performances. These findings are consistent with previous works on the use of CrAss, mth-ND5, and 5-HIAA in the context of WBE (Ahmed *et al.*, 2020b; Hutchison *et al.*, 2022; Hsu *et al.*, 2022; Polo *et al.* 2020). These biomarkers seem to be the best candidates at helping accurately normalise SARS-CoV-2 concentrations in wastewater to better correlate with case data.

Another important result is the differences in performance between the models using flowrate with a biomarker versus the models using a single biomarker at a time. The highest correlations were obtained with the simplest models, using a single variable (excluding N2) (Figure 12). The more complex models still outperformed N2 on its own. However, adding information and complexity to the simple models did not improve the correlation with case data further, and rather often lowered the correlations. It's the case for the models using CrAss and mth-ND5, both having higher correlations on their own then when paired with flowrate. This effect maybe be caused simply by the doubled effect of flowrate. The effect flowrate has on biomarkers and targets in wastewater is intrinsic within the biomarker concentrations measured. The concentrations fluctuate in great part due to the variations in flowrate, diluting or concentrating the biomarkers such as CrAss. Therefore, flowrate variations are included within the models using biomarker concentrations. Adding flowrate as a separate variable simply adds redundant information, adding unnecessary variation and noise in the model, reducing the correlation. One could use the flowrate data to calculate the daily load of each biomarker and use the load as a variable when modelling, but this would simply take out flowrate information of the model, which is counter productive. Using daily loads and adding back in flowrate as a variable into modelling amounts to the same as simply including biomarker concentrations alone.

It is important to remember the context of the study. The anonymous case data used was provided by the Rimouski Regional Hospital and as such, represents case incidence solely within the hospital of all and any patient in care. The case data do not represent case incidence across the sewershed sampled. In theory, the incidence of COVID-19 cases would

be proportionally similar between the hospital and the whole city of Rimouski. However, as with any pandemic or pathogen, there are complex variations and dynamics leading to differences between the case data of the hospital and the hypothetical case data of the sewershed. Comparing a sample which represents a large scale to a data set representing a portion of that scale is less ideal than comparing case data and wastewater data of the same scale. However, city-wide clinical testing was unavailable for the period of the study. Modeling for shedding rate does help increase the correlation between the case data and the wastewater data. Nevertheless, clinical data would be beneficial to thoroughly test if a certain biomarker, method or process is performing as expected from a study such as ours.

2.8 CONCLUSION

This study's foremost aim was two-fold. The first step was to measure the concentrations of four biomarkers with a good potential as a normalisation tool for the analysis of SRAS-CoV-2 in wastewater in a small city. The second step consisted of measuring their performance based on the correlation with COVID-19 case incidence to compare them with the most used biomarker, PMMoV. This study demonstrates the capabilities of CrAss, mth-ND5, 5-HIAA and flowrate as good normalisation tools in small communities and suggests their performance will be as good in most other settings. The results also demonstrate the poor performance of PMMoV and caffeine as a normalisation tool as they reduce the correlation with case data. As such, future studies aiming at increasing this correlation through WBE should avoid the use of diet related biomarkers, such as PMMoV and caffeine, and focus on biomarkers more specific to human physiology, such as CrAss, 5-HIAA, and mth-ND5. The link between temporal variations in the concentration of biomarkers and their performance suggests temporal variation could be used as a reliable criterion to test certain biomarkers from literature without having to rely on pandemic related data. This study also shows that while flow normalisation is a good way to better correlate wastewater data to case incidence, when using a biomarker, flowrate can become redundant and diminish the correlation, making it a good tool in situations where access to accurate biomarker data cannot be obtained.

Despite the limitations related to the use of case incidence provided from the municipality's regional hospital, due to the lack of mass clinical testing, this study's findings align with general trends found in other studies. Further testing would be required to confirm the high performance of CrAss, mth-ND5, 5-HIAA, and flowrate, but these are nevertheless promising results, as those biomarkers outperform PMMoV. Access to real-time, sewershed-wide case data is crucial for the proper testing of modelling and normalisation methods as it helps illustrate a more accurate picture of the dynamics between wastewater and the population. Although case incidence from a portion of the population or sewershed can represent the case incidence of its city, this type of data should only be used when city-wide case incidence is unavailable. It is also paramount to remember the crucial role of viral shedding dynamics in any model estimating or predicting case incidence based on wastewater data. Considering this leads to more accurate and reliable models.

CONCLUSION GÉNÉRALE

Le présent projet de recherche concerne un des plus gros manquements dans le domaine de l'ÉBEU, soit l'absence d'un marqueur populationnel capable de bien normaliser les concentrations du SRAS-CoV-2 dans les eaux usées. Ce mémoire avait comme premier but de réviser la littérature existante concernant les marqueurs populationnels et de déterminer les caractéristiques les plus importantes pour leur utilisation dans un contexte d'ÉBEU, soit normaliser les concentrations du SRAS-CoV-2 dans les eaux usées. De tous les critères de sélections proposés, quoiqu'ils aient tous une importance, ceux qui ressortent davantage sont la variation temporelle et l'unicité. La faible variation de la concentration dans l'eau usée est cruciale puisqu'elle permet de réduire les sources de variations externes dans l'eau usée tout en augmentant l'effet des nouveaux cas actifs dans la population. L'unicité d'un marqueur est tout aussi importante. Ayant une seule origine permet non seulement à réduire les sources de variations dans les eaux usées, mais aussi assure que les variations observées sont associées soit avec un changement au niveau populationnel ou un changement dans l'eau usée et non à une source externe pouvant réduire l'exactitude de la modélisation. L'association entre la performance d'un marqueur et ces deux critères peut être observée au CHAPITRE 2 dans la section 2.6. Nous avons aussi remarqué que la provenance exacte et la nature d'un marqueur doivent être considérées. Parmi les marqueurs étudiés ici, les seuls marqueurs ayant des performances inadéquates sont ceux associés à la diète et à la consommation de certains produits (la caféine et le PMMoV). L'origine du marqueur n'était pas un critère utilisé à lui seul, il était indirectement inclus dans la variation individuelle élaborée dans la section 1.2.5. Toutefois, il serait pertinent de le considérer comme son propre critère dans de futures études puisque l'origine d'un marqueur exerce une grande influence sur sa performance dans le contexte de la normalisation populationnelle en ÉBEU.

Le deuxième objectif de ce projet était de quantifier le SRAS-CoV-2 et les marqueurs choisis dans l'eau usée de Rimouski. Cet objectif a été atteint avec quelques défis. Les méthodes choisies ont été adéquates pour une quantification exacte et précise de toutes les cibles dans l'étude. Les rendements d'extraction pour les cibles génétiques (75%) et la caféine (92%) sont excellents et s'alignent avec la littérature. Toutefois, le 5-HIAA a un faible rendement (<20%), mais considérant la méthode d'extraction choisie, cela s'aligne avec la littérature (Jang *et al.*, 2019). L'objectif n'était pas d'optimiser les méthodes d'extractions, mais bien simplement de quantifier les composés dans l'eau usée, alors l'amélioration de ces méthodes d'extraction est laissée à de futurs projets.

Le dernier objectif de ce projet, soit comparer la performance des nouveaux marqueurs au PMMoV, comporte des résultats peu surprenants considérant la sélection des marqueurs choisis. Comme démontré dans le CHAPITRE 2, le débit journalier, le CrAss, le mth-ND5 et le 5-HIAA possèdent tous des corrélations avec les nouveaux cas cliniques plus élevés que le PMMoV. Ces marqueurs sont tous plus performants que le PMMoV pour normaliser les concentrations du SRAS-CoV-2 dans les eaux usées. Cela est dû principalement à leur faible variation dans les eaux usées. C'est une caractéristique importante que le PMMoV ne possède pas puisque son origine dans la diète mène à de trop grandes sources de variations dans l'eau usée. Il est important aussi de noter l'influence de l'estimation du taux d'excrétion virale. Considérer ce phénomène dans la modélisation des données d'ÉBEU apporte une plus grande exactitude dans les modèles. Cela ne fait qu'augmenter la corrélation entre les données d'ÉBEU et les cas cliniques et nous permet de développer notre compréhension de la relation entre la transmission du virus dans une ville et l'excrétion des particules virales dans l'eau usée. Cette meilleure corrélation démontre notre connaissance grandissante des dynamiques entre l'eau usée, les composés qui s'y retrouvent et la transmission du virus de la COVID-19 à travers une ville.

Selon les critères de sélections, quelques autres marqueurs populationnels pourraient être intéressants d'explorer comme alternative. C'est le cas de la paraxanthine. Ce marqueur, présenté dans le CHAPITRE 1, pourrait surpasser le PMMoV et la caféine sur le plan de la

performance. Son seul inconvénient est une faible concentration, donc une faible détectabilité, dans les eaux usées. Toutefois, cela est un problème facilement résolu par l'utilisation d'une méthode davantage sensible, telle la chromatographie liquide avec un spectromètre de masse (LC-MS). Le polyomavirus humain JCPyV pourrait aussi être un candidat acceptable pour la normalisation des concentrations du SRAS-CoV-2 dans les eaux usées, tel que mentionné dans le CHAPITRE 1. La performance de ces deux marqueurs n'a pas été évaluée dans ce projet-ci par question de ressource et de méthode.

A priori, les résultats et leur étendue sont restreints. Les données de cas actifs de COVID-19 utilisées proviennent d'un seul hôpital d'une seule ville. Quoique les données d'eaux usées proviennent de cette même ville, l'échelle de ces deux jeux de données reste différente ; la première est d'un seul bâtiment, alors que la deuxième échelle couvre l'entièreté des usagers couverts par le réseau d'eau usée de Rimouski, soit la majorité de la ville. Comparer des données de deux échelles de grandeurs différentes implique une interprétation des résultats limitée. Même si les nouveaux cas de COVID-19 dans l'hôpital régional de Rimouski reflètent en théorie les nouveaux cas dans le reste de la ville, il y aura toujours de différence et des interactions qui ne seront jamais mesurées et laissé au hasard. Ceci souligne l'importance d'avoir accès à des données de cas de COVID-19 en temps réel qui couvre une plus grande proportion de la population qu'uniquement les patients d'un hôpital. La prise de ces données de cas clinique est cruciale et nécessaire à la recherche et à l'avancement des modèles prédictifs d'ÉBEU.

Malgré les limites, les retombées de cette recherche sont grandes. La pandémie de la COVID-19 n'est pas la première et sera loin d'être la dernière pandémie où l'eau usée joue un rôle crucial dans la détection et le suivi d'un virus à travers la communauté ainsi que dans la gestion générale au sein des autorités de santé publique. Être en mesure de mieux corréler les concentrations d'un virus dans l'eau usée, dans ce cas-ci le SRAS-CoV-2, avec les nouveaux cas dans la population est applicable à tout autre pathogène détectable dans les eaux usées. Pour les pandémies futures, nous aurons une longueur d'avance en gestion si les résultats du présent projet sont mis en fonction adéquatement. Avoir une meilleure

approximation des nouveaux cas à venir ne fait qu'aider les autorités de santé publique à travers le monde. Cela aide à mieux préparer les lits d'hôpitaux pour les personnes avec des symptômes graves, mieux planifier les horaires du personnel de plus ou de moins pour assurer une main-d'œuvre suffisante dans les centres de soins, assurer un bon approvisionnement du matériel et l'équipement pour le bon soin des patients et surtout aider dans les décisions et les annonces concernant les mesures sanitaires à prendre pour réduire la transmission d'un pathogène dans la communauté tout en assurant un impact minime sur la vie quotidienne de la population. Non seulement cela, mais les résultats de ce projet sont applicables au-delà de la ville de Rimouski. Les marqueurs choisis sont universels, c'est-à-dire qu'ils peuvent être utilisés à travers le monde. Leur performance peut varier selon la région, mais comme démontré au CHAPITRE 2, ils améliorent la corrélation entre les données d'eaux usées et de cas dans la population. Connaissant bien la théorie derrière ces marqueurs, ils seront sans doute d'une grande aide dans la gestion des pandémies futures.

ANNEXES

ANNEXE I

Marqueurs populationnels considérés

Produits génétiques	PMMoV
	ADN mitochondrial humain
	CrAssphage
	Polyomavirus humain
	Adénovirus
	Bactéroïde HF138
Bactéries et parasites	<i>Clostridium perfringens</i>
	<i>Escherichia coli</i>
	Cystes <i>Gardia</i>
	Oocystes de <i>Cryptosporidium</i>
Édulcorants	Acésulfame
	Cyclamate
	Saccharine
	Sucralose
	Alitame
	Aspartame
	Néotame
	Néohepéridine dihydrochalcone
Drogues	Cocaïne
	Caféine
	Nicotine
	Méthamphétamine
	Alcool
Métabolites	Cholestérol
	Créatine/Créatinine
	Sérotonine/5-HIAA
	Androstènedione
	Cortisol
	Coprostérol

	1-aminopropan-2-one
Produits pharmaceutiques et de soins personnels	Aténolol
	Gabapentine
	Hydrochlorothiazide
	Venlafaxine
	Carbamazépine
	Codéine
	Ibuprofène
	Paracétamol
	Furosémide
	Iopromide
	Naproxène
	Acide salicylique
	Benzodiazépines
Polluants	Parabènes
	Phtalates
	Mycotoxines
	Ignifuges
	Pesticides
	Filtres UV
Autres	BOD
	COD
	Composés phosphorés
	Composés azotés
	Débit journalier
	Matière particulaire en suspension

ANNEXE II

Valeurs de concentration des marqueurs chimiques et du débit journalier

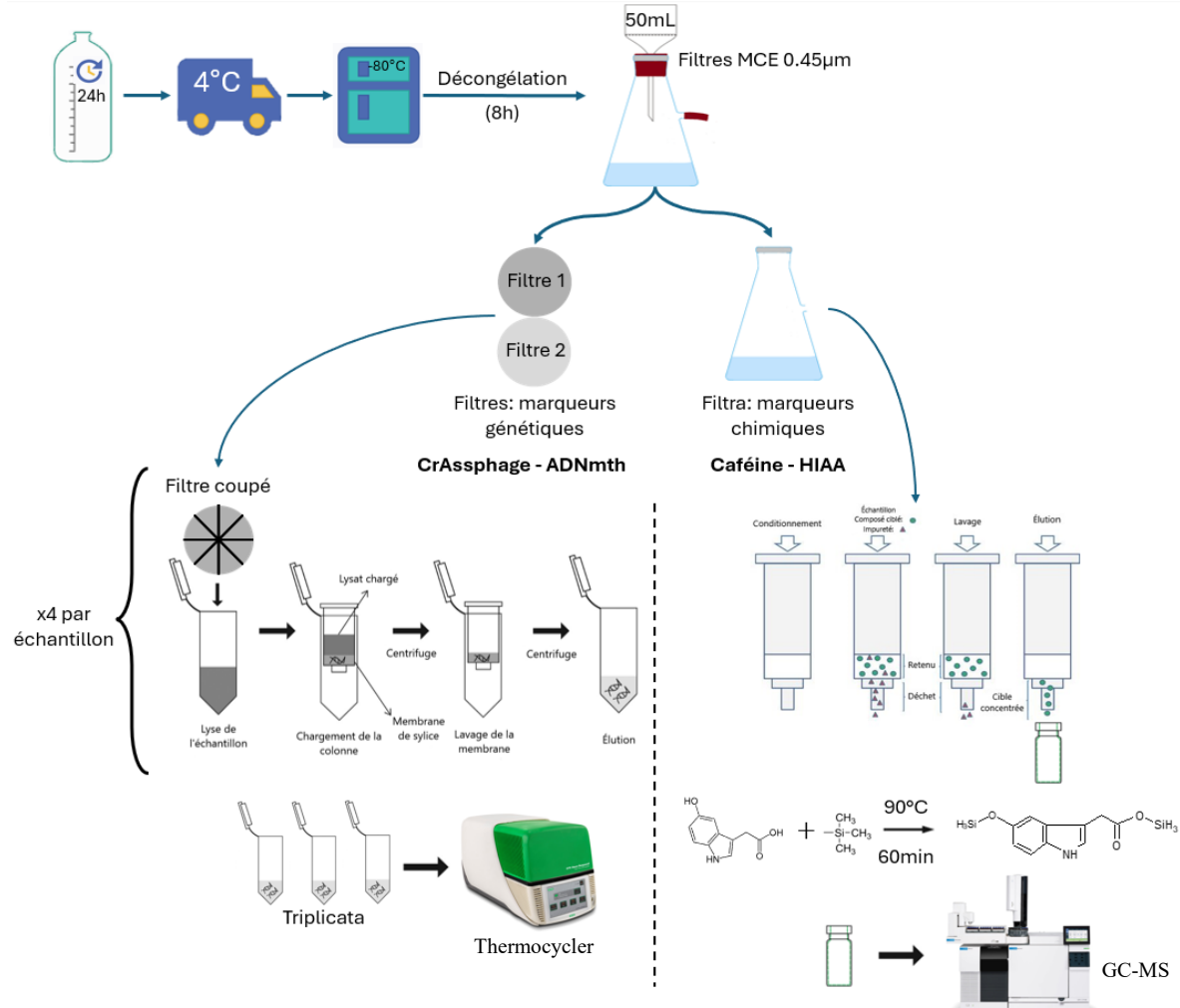
Date	Débit (m ³)	Caféine (µg/L)	5-HIAA (ng/L)
2022-01-21	18239	1.34	5.93
2022-01-24	18870	11.53	6.88
2022-01-26	18967	10.85	5.25
2022-01-28	19002	3.60	6.13
2022-01-31	19770	6.41	4.79
2022-02-02	19238	11.75	4.57
2022-02-04	19137	9.77	4.75
2022-02-07	18973	14.62	18.89
2022-02-09	19130	40.86	13.08
2022-02-11	18818	15.09	13.11
2022-02-14	18871	16.16	8.45
2022-02-16	18792	10.14	5.18
2022-02-21	18448	2.50	9.09
2022-02-23	18093	13.43	15.76
2022-02-28	18323	5.80	14.30
2022-03-02	18381	2.58	15.63
2022-03-04	18566	32.28	20.23
2022-03-07	18960	19.24	5.28
2022-03-09	18907	4.64	8.65
2022-03-11	19068	19.43	9.81
2022-03-16	19555	31.99	8.76
2022-03-18	33733	7.46	7.67
2022-03-21	25038	3.21	10.36
2022-03-23	24579	5.20	5.43
2022-03-25	20595	2.58	11.84
2022-03-28	22799	13.23	5.11
2022-03-30	22909	3.06	8.12
2022-04-01	29597	0.84	15.99
2022-04-04	32693	1.80	7.61
2022-04-06	35410	0.87	9.65
2022-04-08	45825	1.01	6.50
2022-04-11	65293	1.61	4.77
2022-04-13	68765	1.31	5.95

Date	Débit (m³)	Caféine (µg/L)	5-HIAA (ng/L)
2022-04-15	75543	0.31	7.41
2022-04-18	63130	6.79	6.07
2022-04-19	57125	1.40	4.70
2022-04-20	53089	4.07	5.07*
2022-04-21	53690	5.19	5.44*
2022-04-22	57583	7.54	5.80
2022-04-23	53790	1.53	4.22
2022-04-24	49223	4.50	6.45
2022-04-25	48390	3.82	4.93
2022-04-27	51235	7.90	4.65*
2022-04-29	52855	5.37	4.37
2022-05-02	43555	2.10	4.51
2022-05-04	40001	3.75	8.41
2022-05-06	36686	4.68	14.04
2022-05-09	33527	3.54	4.15
2022-05-11	32102	1.31	4.36
2022-05-13	30053	2.17	8.13
2022-05-16	33450	3.44	5.82
2022-05-18	36381	1.72	6.40
2022-05-20	31584	1.18	10.16
2022-05-23	29016	4.60	5.07

*Valeurs calculées basées sur une moyenne mobile utilisée pour le calcul et la création des matrices de corrélation. Originellement non-détecté.

ANNEXE III

Diagramme résumé de la méthodologie



RÉFÉRENCES BIBLIOGRAPHIQUES

Abd-Elshafy, D. N., Nadeem, R., & Bahgat, M. M. (2022). Using Molecular Tools to Compare Various Concentration Protocols Prior to Detecting SARS-Cov-2 RNA in Wastewater. *Egyptian Journal of Chemistry*, 65(131), 1241–1248.
<https://doi.org/10.21608/EJCHEM.2022.165190.7030>

Ahmed, W., Bertsch, P. M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic, A., Riddell, S., Sherchan, S. P., Simpson, S. L., Sirikanchana, K., Symonds, E. M., Verhagen, R., Vasan, S. S., Kitajima, M., & Bivins, A. (2020a). Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research*, 191, 110092.
<https://doi.org/10.1016/J.ENVRES.2020.110092>

Ahmed, W., Bertsch, P. M., Bivins, A., Bibby, K., Farkas, K., Gathercole, A., Haramoto, E., Gyawali, P., Korajkic, A., McMinn, B. R., Mueller, J. F., Simpson, S. L., Smith, W. J. M., Symonds, E. M., Thomas, K. v., Verhagen, R., & Kitajima, M. (2020b). Comparison of virus concentration methods for the RT-qPCR-based recovery of murine hepatitis virus, a surrogate for SARS-CoV-2 from untreated wastewater. *Science of the Total Environment*, 739. <https://doi.org/10.1016/J.SCITOTENV.2020.139960>

Bayati, M., Hsieh, H. Y., Hsu, S. Y., Li, C., Rogers, E., Belenchia, A., Zemmer, S. A., Blanc, T., LePage, C., Klutts, J., Reynolds, M., Semkiw, E., Johnson, H. Y., Foley, T., Wieberg, C. G., Wenzel, J., Lyddon, T., LePique, M., Rushford, C., Lin, C. H. (2022). Identification and quantification of bioactive compounds suppressing SARS-CoV-2 signals in wastewater-based epidemiology surveillance. *Water Research*, 221, 118824.
<https://doi.org/10.1016/J.WATRES.2022.118824>

Bivins, A., Greaves, J., Fischer, R., Yinda, K. C., Ahmed, W., Kitajima, M., Munster, V. J., & Bibby, K. (2020). Persistence of SARS-CoV-2 in Water and Wastewater. *Environmental Science and Technology Letters*, 7(12), 937–942.
https://doi.org/10.1021/ACS.ESTLETT.0C00730/ASSET/IMAGES/LARGE/EZ0C00730_0002.JPEG

Boehm, A. B., Hughes, B., Duong, D., Chan-Herur, V., Buchman, A., Wolfe, M. K., & White, B. J. (2023). Wastewater concentrations of human influenza, metapneumovirus, parainfluenza, respiratory syncytial virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: a surveillance study. *The Lancet Microbe*, 4(5), e340–e348. [https://doi.org/10.1016/S2666-5247\(22\)00386-X](https://doi.org/10.1016/S2666-5247(22)00386-X)

Bofill-Mas, S., Albinana-Gimenez, N., Clemente-Casares, P., Hundesa, A., Rodriguez-Manzano, J., Allard, A., Calvo, M., & Girones, R. (2006). Quantification and Stability of Human Adenoviruses and Polyomavirus JCPyV in Wastewater Matrices. *Applied and Environmental Microbiology*, 72(12), 7894–7896. <https://doi.org/10.1128/AEM.00965-06>

Boxus, M., Letellier, C., & Kerkhofs, P. (2005). Real Time RT-PCR for the detection and quantitation of bovine respiratory syncytial virus. *Journal of Virological Methods*, 125, 125–130. <https://doi.org/10.1016/j.jviromet.2005.01.008>

Brewer, A. J., Ort, C., Banta-Green, C. J., Berset, J.-D., & Field, J. A. (2012). *Normalized Diurnal and Between-Day Trends in Illicit and Legal Drug Loads that Account for Changes in Population*. <https://doi.org/10.1021/es202447r>

Castiglioni, S., Bagnati, R., Fanelli, R., Pomati, F., Calamari, D., & Zuccato, E. (2006). Removal of pharmaceuticals in sewage treatment plants in Italy. *Environmental Science and Technology*, 40(1), 357–363. https://doi.org/10.1021/ES050991M/SUPPL_FILE/ES050991MSI20050930_013329.PDF

Center for Disease Control and Prevention (CDC). (2023). National Wastewater Surveillance System (NWSS). *Wastewater Surveillance Testing Methods*. Last Reviewed: March 10, 2023. https://archive.cdc.gov/www_cdc_gov/nwss/testing.html

Centers for Disease Control and Prevention (CDC). (2020). National Wastewater Surveillance System (NWSS). *Developing a Wastewater Surveillance Sampling Strategy*. Last Reviewed: June 26, 2024. https://archive.cdc.gov/www_cdc_gov/nwss/sampling.html

Cevik, M., Tate, M., Lloyd, O., Maraolo, A. E., Schafers, J., & Ho, A. (2020). SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *The Lancet. Microbe*, 2(1), e13. [https://doi.org/10.1016/S2666-5247\(20\)30172-5](https://doi.org/10.1016/S2666-5247(20)30172-5)

Chen, C., Kostakis, C., Gerber, J. P., Tschärke, B. J., Irvine, R. J., & White, J. M. (2014). *Towards finding a population biomarker for wastewater epidemiology studies*. <https://doi.org/10.1016/j.scitotenv.2013.11.075>

COVID - CentraEau. (2021). Retrieved September 27, 2023, from <https://centreau.org/covid/>

Cutrupi, F., Rossi, M., Cadonna, M., Poznanski, E., Manara, S., Postinghel, M., Palumbi, G., Bellisomi, M., Nicosia, E., Allaria, G., Dondero, L., Veneri, C., Mancini, P., Giusy, ·, Ferraro, B., Rosa, G., Suffredini, E., Foladori, P., & Grasselli, E. (2023). *Evaluation of concentration procedures, sample pre-treatment, and storage condition for the detection of SARS-CoV-2 in wastewater*. 30, 106660–106670. <https://doi.org/10.1007/s11356-023-29696-y>

D'Aoust, P. M., Mercier, E., Montpetit, D., Jia, J. J., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Langlois, M. A., Servos, M. R., MacKenzie, M., Figeys, D., MacKenzie, A. E., Graber, T. E., & Delatolla, R. (2021). Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence. *Water Research*, 188. <https://doi.org/10.1016/j.watres.2020.116560>

Daughton, C. G. (2020). Wastewater surveillance for population-wide Covid-19: The present and future. <https://doi.org/10.1016/j.scitotenv.2020.139631>

Dehghan Banadaki, M., Torabi, S., Rockward, A., Strike, W. D., Noble, A., Keck, J. W., & Berry, S. M. (2024). Simple SARS-CoV-2 concentration methods for wastewater surveillance in low resource settings. *Science of The Total Environment*, 912, 168782. <https://doi.org/10.1016/J.SCITOTENV.2023.168782>

Doernberg, S. B., Holubar, M., Jain, V., Weng, Y., Lu, D., Bollyky, J. B., Sample, H., Huang, B., Craik, C. S., Desai, M., Rutherford, G. W., & Maldonado, Y. (2022). Incidence and Prevalence of Coronavirus Disease 2019 Within a Healthcare Worker Cohort During the First Year of the Severe Acute Respiratory Syndrome Coronavirus 2 Pandemic. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, 75(9), 1573–1584. <https://doi.org/10.1093/CID/CIAC210>

Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., Barr, J. J., Speth, D. R., Seguritan, V., Aziz, R. K., Felts, B., Dinsdale, E. A., Mokili, J. L., & Edwards, R. A. (2014). A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. *Nature Communications* 2014 5:1, 5(1), 1–11. <https://doi.org/10.1038/ncomms5498>

Environment and Climate Change Canada (2023) Canadian Environmental Sustainability Indicators: Municipal wastewater treatment. Available at: www.canada.ca/en/environment-climate-change/services/environmental-indicators/municipal-wastewater-treatment.html.

Environmental Protection Agency (EPA). (1998). How Wastewater Treatment Works...The Basics. In *Office of Water*. <https://www3.epa.gov/npdes/pubs/bastre.pdf>

Environmental Protection Agency (EPA). (2003). Wastewater Technology Fact Sheet - Disinfection for Small Systems EPA 832-F-03-024. In *Office of Water*.

Fang, Z., Roberts, A. M. I., Mayer, C. D., Frantsuzova, A., Potts, J. M., Cameron, G. J., Singleton, P. T. R., & Currie, I. (2022). Wastewater monitoring of COVID-19: a perspective from Scotland. *Journal of Water and Health*, 20(12), 1688–1700. <https://doi.org/10.2166/WH.2022.082>

Gao, J., Burgard, D. A., Tschärke, B. J., Lai, F. Y., O'Brien, J. W., Nguyen, H. D., Zheng, Q., Li, J., Du, P., Li, X., Wang, D., Castiglioni, S., Cruz-Cruz, C., Baz-Lomba, J. A.,

Yargeau, V., Emke, E., Thomas, K. v., Mueller, J. F., & Thai, P. K. (2022). Refining the estimation of amphetamine consumption by wastewater-based epidemiology. *Water Research*, 225, 119182. <https://doi.org/10.1016/J.WATRES.2022.119182>

García-Aljaro, C., Ballesté, E., Muniesa, M., & Jofre, J. (2017). Determination of crAssphage in water samples and applicability for tracking human faecal pollution. *Microbial Biotechnology*, 10(6), 1775–1780. <https://doi.org/10.1111/1751-7915.12841>

Gracia-Lor, E., Castiglioni, S., Bade, R., Been, F., Castrignanò, E., Covaci, A., González-Mariño, I., Hapeshi, E., Kasprzyk-Hordern, B., Kinyua, J., Lai, F. Y., Letzel, T., Lopardo, L., Meyer, M. R., O'Brien, J., Ramin, P., Rousis, N. I., Rydevik, A., Ryu, Y., Santos, M. M., Senta, I., Thomaidis, N. S., Veloutsou, S., Yang, Z., Zuccato, E., Bijlsma, L. (2017). Measuring biomarkers in wastewater as a new source of epidemiological information: Current state and future perspectives. In *Environment International* (Vol. 99, pp. 131–150). Elsevier Ltd. <https://doi.org/10.1016/j.envint.2016.12.016>

Graham Construction and Engineering, *Waterloo Wastewater Treatment Plant Upgrade – Contract 3 - Graham Construction & Engineering Inc.* (n.d.). Retrieved February 22, 2023, from <https://grahambuilds.com/waterloo-wastewater-treatment-plant-upgrade-contract-3/>

Green, H. C., Haugland, R. A., Varma, M., Millen, H. T., Borchardt, M. A., Field, K. G., Walters, W. A., Knight, R., Sivaganesan, M., Kelty, C. A., & Shanks, O. C. (2014). Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. *Applied and Environmental Microbiology*, 80(10), 3086–3094. <https://doi.org/10.1128/AEM.04137-13/ASSET/571759A2-B8EE-463E-9A33-85DC323649E9/ASSETS/GRAPHIC/ZAM9991053420003.JPEG>

Greenwald, H. D., Kennedy, L. C., Hinkle, A., Whitney, O. N., Fan, V. B., Crits-Christoph, A., Harris-Lovett, S., Flamholz, A. I., Al-Shayeb, B., Liao, L. D., Beyers, M., Brown, D., Chakrabarti, A. R., Dow, J., Frost, D., Koekemoer, M., Lynch, C., Sarkar, P., White, E., ... Nelson, K. L. (2021). Tools for interpretation of wastewater SARS-CoV-2 temporal and spatial trends demonstrated with data collected in the San Francisco Bay Area. *Water Research X*, 12. <https://doi.org/10.1016/j.wroa.2021.100111>

Government of Canada. (2012). *Wastewater Systems Effluent Regulations*, (testimony of Government of Canada). <https://laws-lois.justice.gc.ca/eng/regulations/sor-2012-139/fulltext.html>

Hamming, I., Timens, W., Bulthuis, M., Lely, A. T., Navis, G. J., & van Goor, H. (2004). Rapid Communication Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *Journal of Pathology J Pathol*, 203, 631–637. <https://doi.org/10.1002/path.1570>

Hamouda, M., Mustafa, F., Maraqa, M., Rizvi, T., & Aly, A. (2021). Wastewater surveillance for SARS-CoV-2 : Lessons learnt from recent studies to define future

applications. *Science of the Total Environment*, 759, 143493.
<https://doi.org/10.1016/j.scitotenv.2020.143493>

Hokajärvi, A.-M., Rytönen, A., Tiwari, A., Kauppinen, A., Oikarinen, S., Lehto, K.-M., Kankaanpää, A., Gunnar, T., Al-Hello, H., Blomqvist, S., Miettinen, I. T., Savolainen-Kopra, C., & Pitkänen, T. (2021). *The detection and stability of the SARS-CoV-2 RNA biomarkers in wastewater influent in Helsinki, Finland*.
<https://doi.org/10.1016/j.scitotenv.2021.145274>

Honap ID, T. P., Sankaranarayanan, K., Schnorr, S. L., Ozga, A. T., Warinner, C., & Lewis Jr, C. M. (2020). *Biogeographic study of human gut-associated crAssphage suggests impacts from industrialization and recent expansion*.
<https://doi.org/10.1371/journal.pone.0226930>

Hrudey, S. E., Ashbolt, N. J., Isaac-Renton, J. L., McKay, R. M., & Servos, M. R. (2020). ÉPIDÉMIOLOGIE BASÉE SUR LES EAUX USÉES POUR LE SRAS-COV-2.

Hrudey, S. E., Bischel, H. N., Charrois, J., Chik, A. H. S., Conant, B., Delatolla, R., Dorner, S., Graber, T. E., Hubert, C., Isaac-Renton, J., Pons, W., Safford, H., Servos, M., & Sikora, C. (2022). Wastewater Surveillance for SARS-CoV-2 RNA in Canada.
<https://doi.org/10.1139/Facets-2022-0148>, 7, 1493–1597.
<https://doi.org/10.1139/FACETS-2022-0148>

Hsu, S. Y., Bayati, M., Li, C., Hsieh, H. Y., Belenchia, A., Klutts, J., Zemmer, S. A., Reynolds, M., Semkiw, E., Johnson, H. Y., Foley, T., Wieberg, C. G., Wenzel, J., Johnson, M. C., & Lin, C. H. (2022). Biomarkers selection for population normalization in SARS-CoV-2 wastewater-based epidemiology. *Water Research*, 223, 118985.
<https://doi.org/10.1016/J.WATRES.2022.118985>

Hundermark, E. L., & Takahashi, M. K. (2020). Improving the yield of environmental DNA from filtered aquatic samples. *Conservation Genetics Resources*, 12(1), 49–51.
<https://doi.org/10.1007/S12686-018-1067-3>

Hutchison, J. M., Li, Z., Chang, C.-N., Hiripitiyage, Y., Wittman, M., & Sturm, B. S. M. (2022). *Improving correlation of wastewater SARS-CoV-2 gene copy numbers with COVID-19 public health cases using readily available biomarkers*.
<https://doi.org/10.1093/femsmc/xtac010/6562076>

INSPQ. *Archives des données COVID-19 | Institut national de santé publique du Québec*. (2023, January 2). <https://www.inspq.qc.ca/covid-19/donnees/archives>

Jackson, C. B., Farzan, M., Chen, B., & Choe, H. (2021). Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews Molecular Cell Biology* 2021 23:1, 23(1), 3–20.
<https://doi.org/10.1038/s41580-021-00418-x>

Jang, M., Chernyshov, V.D., Pirogov, A.V., Tataurova, O. G., Rozhanets, V. V., Shpigun, O. A. (2019). Determination of 5-Hydroxyindole-3-Acetic Acid in Wastewater by High Performance Liquid Chromatography Coupled with Tandem Mass Spectrometric Detection. *Inorg Mater* **55**, 1352–1358. <https://doi.org/10.1134/S002016851914005X>

Kavlick, M. F. (2018). Development of a universal internal positive control. *BioTechniques*, 65(5), 275–280. <https://doi.org/10.2144/BTN-2018-0034/ASSET/IMAGES/LARGE/FIGURE4.JPEG>

Kitajima, M., Sassi, H. P., & Torrey, J. R. (2018). *Pepper mild mottle virus as a water quality indicator*. 1, 19. <https://doi.org/10.1038/s41545-018-0019-5>

Khreefa, Z., Barbier, M. T., Koksai, A. R., Love, G., & del Valle, L. (2023). Pathogenesis and Mechanisms of SARS-CoV-2 Infection in the Intestine, Liver, and Pancreas. *Cells*, 12(2). <https://doi.org/10.3390/CELLS12020262>

Kokotou, M. G., & Thomaidis, N. S. (2013). Determination of eight artificial sweeteners in wastewater by hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Analytical Methods*, 5(16), 3825–3833. <https://doi.org/10.1039/c3ay40599k>

Kumar, M., Joshi, M., Kumar Patel, A., & Joshi, C. G. (2021). Unravelling the early warning capability of wastewater surveillance for COVID-19: A temporal study on SARS-CoV-2 RNA detection and need for the escalation. *Environmental Research*, 196, 110946. <https://doi.org/10.1016/j.envres.2021.110946>

Leo, B. F., Lin, C. Y., Markandan, K., Saw, L. H., Mohd Nadzir, M. S., Govindaraju, K., Shariffuddin, I. I., Sankara, R., Tiong, Y. W., Pakalapati, H., & Khalid, M. (2023). An overview of SARS-CoV-2 transmission and engineering strategies to mitigate risk. *Journal of Building Engineering*, 73, 106737. <https://doi.org/10.1016/J.JOBE.2023.106737>

Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology*, 3(Volume 3, 2016), 237–261. <https://doi.org/10.1146/ANNUREV-VIROLOGY-110615-042301/CITE/REFWORKS>

Li, M. Y., Li, L., Zhang, Y., & Wang, X. S. (2020). Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infectious Diseases of Poverty*, 9(1). <https://doi.org/10.1186/S40249-020-00662-X>

Li, Y., Ash, K., Alamilla, I., Joyner, D., Edward Williams, D., McKay, P. J., Green, B., DeBlander, S., North, C., Kara-Murdoch, F., Swift, C., Hazen, T. C., Jores, J., Mazziotta, C., & Moore, M. D. (2024). COVID-19 trends at the University of Tennessee: predictive insights from raw sewage SARS-CoV-2 detection and evaluation and PMMoV as an indicator for human waste. *Frontiers in Microbiology*, 15, 1379194. <https://doi.org/10.3389/FMICB.2024.1379194>

- Lowry, S. A., Wolfe, M. K., & Boehm, A. (2023). Respiratory virus concentrations in human excretions that contribute to wastewater: a systematic review and meta-analysis Uncorrected Proof. *Journal of Water and Health*, 00, 1. <https://doi.org/10.2166/wh.2023.057>
- Lukasik, J., Scott, T. M., Andryshak, D., & Farrah, S. R. (2000). Influence of Salts on Virus Adsorption to Microporous Filters. *Applied and Environmental Microbiology*, 66(7), 2914. <https://doi.org/10.1128/AEM.66.7.2914-2920.2000>
- Lux K, Jorns M, Vogler C. Creation of a State Drug Shortage Dashboard in the COVID-19 Pandemic. *Hospital Pharmacy*. 2022;57(4):474-481. doi:[10.1177/00185787211046861](https://doi.org/10.1177/00185787211046861)
- Maal-Bared, R., Qiu, Y., Li, Q., Gao, T., Hrudey, S. E., Bhavanam, S., Ruecker, N. J., Ellehoj, E., Lee, B. E., & Pang, X. (2023). Does normalization of SARS-CoV-2 concentrations by Pepper Mild Mottle Virus improve correlations and lead time between wastewater surveillance and clinical data in Alberta (Canada): comparing twelve SARS-CoV-2 normalization approaches. *Science of The Total Environment*, 856, 158964. <https://doi.org/10.1016/J.SCITOTENV.2022.158964>
- Marroquín-Cardona, A. G., Johnson, N. M., Phillips, T. D., & Hayes, A. W. (2014). Mycotoxins in a changing global environment – A review. *Food and Chemical Toxicology*, 69, 220–230. <https://doi.org/10.1016/J.FCT.2014.04.025>
- McKenna, F., Scott, M. H., and Fenves, G. L. (2010) Nonlinear finite-element analysis software architecture using object composition. *Journal of Computing in Civil Engineering*, 24(1):95-107. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000002](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000002)
- Mitrancescu, A., Uchaikina, A., Kau, A.-S., Stange, C., Ho, J., Tiehm, A., Wurzbacher, C., & Drewes, J. E. (2022). Wastewater-Based Epidemiology for SARS-CoV-2 Biomarkers: Evaluation of Normalization Methods in Small and Large Communities in Southern Germany. *ACS ES&T Water*. <https://doi.org/10.1021/ACSESTWATER.2C00306>
- North, D., & Bibby, K. (2023). Comparison of viral concentration techniques for native fecal indicators and pathogens from wastewater. *Science of The Total Environment*, 905, 167190. <https://doi.org/10.1016/J.SCITOTENV.2023.167190>
- O'Brien, J. W., Thai, P. K., Eaglesham, G., Ort, C., Scheidegger, A., Carter, S., Lai, F. Y., & Mueller, J. F. (2014). A model to estimate the population contributing to the wastewater using samples collected on census day. *Environmental Science and Technology*, 48(1), 517–525. https://doi.org/10.1021/ES403251G/SUPPL_FILE/ES403251G_SI_001.PDF
- Oloye, F. F., Xie, Y., Challis, J. K., Femi-Oloye, O. P., Brinkmann, M., McPhedran, K. N., Jones, P. D., Servos, M. R., & Giesy, J. P. (2023). Understanding common population markers for SARS-CoV-2 RNA normalization in wastewater – A review. *Chemosphere*, 333, 138682. <https://doi.org/10.1016/J.CHEMOSPHERE.2023.138682>

- Palardy, A., Gagné, J. P., & Tremblay, L. (2016). Presence of illicit drugs and pharmaceutical residues in the wastewaters of an eastern Canadian city. *Journal of Xenobiotics*, 5(2), 37–39. <https://doi.org/10.4081/xeno.2015.5773>
- Patey, G., Huot, C., & Levallois, P. (2021). Conditions de faisabilité et utilité de la surveillance de la COVID-19 à l'aide du monitoring du SRAS-CoV-2 dans les eaux usées - REVUE DE LITTÉRATURE. *Institut National de La Santé Publique Du Québec*.
- Paul, J. R., Trask, J. D., & Culotta, C. S. (1939). Poliomyelitic virus in sewage. *Science*, 90(2333), 258–259. <https://doi.org/10.1126/SCIENCE.90.2333.258/ASSET/115DB4B7-8023-46A9-ABD4-6A45ED13ADFF/ASSETS/SCIENCE.90.2333.258.FP.PNG>
- Polo, D., Quintela-baluja, M., Corbishley, A., Jones, D. L., Singer, A. C., Graham, D. W., & Romalde, J. L. (2020). Making waves: Wastewater-based epidemiology for COVID-19 – approaches and challenges for surveillance and prediction. *Water Research*, 186, 116404. <https://doi.org/10.1016/j.watres.2020.116404>
- Puhach, O., Meyer, B., & Eckerle, I. (2023). SARS-CoV-2 viral load and shedding kinetics. *Nature Reviews Microbiology*, 21(3), 147–161. <https://doi.org/10.1038/S41579-022-00822-W>
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rabe, A., Ravuri, S., Burnor, E., Steele, J. A., Kantor, R. S., Choi, S., Forman, S., Batjiaka, R., Jain, S., León, T. M., Vugia, D. J., & Yu, A. T. (2023). Correlation between wastewater and COVID-19 case incidence rates in major California sewersheds across three variant periods. *Journal of Water and Health*, 21(9), 1303–1317. <https://doi.org/10.2166/WH.2023.173>
- Rainey, A. L., Liang, S., Bisesi, J. H., Sabo-Attwood, T., & Maurelli, A. T. (2023). A multistate assessment of population normalization factors for wastewater-based epidemiology of COVID-19. *PLOS ONE*, 18(4), e0284370. <https://doi.org/10.1371/JOURNAL.PONE.0284370>
- Ramos, S., Homem, V., Alves, A., & Santos, L. (2016). A review of organic UV-filters in wastewater treatment plants. *Environment International*, 86, 24–44. <https://doi.org/10.1016/J.ENVINT.2015.10.004>
- Reid, M. J., Langford, K. H., Mørland, J., & Thomas, K. v. (2011). Analysis and Interpretation of Specific Ethanol Metabolites, Ethyl Sulfate, and Ethyl Glucuronide in Sewage Effluent for the Quantitative Measurement of Regional Alcohol Consumption. *Alcoholism: Clinical and Experimental Research*, 35(9), 1593–1599. <https://doi.org/10.1111/J.1530-0277.2011.01505.X>

Rioux, M. D., Guillemette, F., Lemarchand, K., Doiron, K., Lemay, J. F., Maere, T., Dolcé, P., Quessy, P., Abonnenc, N., Vanrolleghem, P. A., & Frigon, D. (2023). Wastewater-based epidemiology: the crucial role of viral shedding dynamics in small communities. *Frontiers in Public Health*, 11. <https://doi.org/10.3389/FPUBH.2023.1141837>

Rousis, N. I., Zuccato, E., & Castiglioni, S. (2016). Monitoring population exposure to pesticides based on liquid chromatography-tandem mass spectrometry measurement of their urinary metabolites in urban wastewater: A novel biomonitoring approach. *Science of The Total Environment*, 571, 1349–1357. <https://doi.org/10.1016/J.SCITOTENV.2016.07.036>

Rousis, N. I., Zuccato, E., & Castiglioni, S. (2017). Wastewater-based epidemiology to assess human exposure to pyrethroid pesticides. *Environment International*, 99, 213–220. <https://doi.org/10.1016/J.ENVINT.2016.11.020>

Senta, I., Gracia-Lor, E., Borsotti, A., Zuccato, E., & Castiglioni, S. (2015). Wastewater analysis to monitor use of caffeine and nicotine and evaluation of their metabolites as biomarkers for population size assessment. *Water Research*, 74, 23–33. <https://doi.org/10.1016/j.watres.2015.02.002>

Shi, H., Wang, B., Niu, L., Cao, M., Kang, W., Lian, K., & Zhang, P. (2017). Trace level determination of 5-hydroxytryptamine and its related indoles in amniotic fluid by gas chromatography-mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis*, 143, 176–182. <https://doi.org/10.1016/j.jpba.2017.05.044>

Singer, A. C., Thompson, J. R., Filho, C. R. M., Street, R., Li, X., Castiglioni, S., & Thomas, K. v. (2023). A world of wastewater-based epidemiology. *Nature Water* 2023 1:5, 1(5), 408–415. <https://doi.org/10.1038/s44221-023-00083-8>

Soller, J., Jennings, W., Schoen, M., Boehm, A., Wigginton, K., Gonzalez, R., Graham, K. E., McBride, G., Kirby, A., & Mattioli, M. (2022). Modeling infection from SARS-CoV-2 wastewater concentrations: promise, limitations, and future directions. *Journal of Water and Health*. <https://doi.org/10.2166/WH.2022.094>

Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1), 72. <https://doi.org/10.2307/1412159>

Stachler, E., Kelty, C., Sivaganesan, M., Li, X., Bibby, K., & Shanks, O. C. (2017). *Quantitative CrAssphage PCR Assays for Human Fecal Pollution Measurement*. <https://doi.org/10.1021/acs.est.7b02703>

Temmam, S., Vongphayloth, K., Baquero, E., Munier, S., Bonomi, M., Regnault, B., Douangboubpha, B., Karami, Y., Chrétien, D., Sanamxay, D., Xayaphet, V., Paphaphanh, P., Lacoste, V., Somlor, S., Lakeomany, K., Phommavanh, N., Pérot, P., Dehan, O., Amara, F., Donati, F., Bigot, T., Nilges, M., Rey, F.A., Van der Werf, S., Brey, P. & Eloit, M.

(2022). Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature*, 604(7905), 330–336. <https://doi.org/10.1038/s41586-022-04532-4>

Thai, P. K., O'Brien, J. W., Banks, A. P. W., Jiang, G., Gao, J., Choi, P. M., Yuan, Z., & Mueller, J. F. (2019). *Evaluating the in-sewer stability of three potential population biomarkers for application in wastewater-based epidemiology*. <https://doi.org/10.1016/j.scitotenv.2019.03.231>

Tolouei, S., Burnet, J. B., Autixier, L., Taghipour, M., Bonsteel, J., Duy, S. V., Sauvé, S., Prévost, M., & Dorner, S. (2019). Temporal variability of parasites, bacterial indicators, and wastewater micropollutants in a water resource recovery facility under various weather conditions. *Water Research*, 148, 446–458. <https://doi.org/10.1016/j.watres.2018.10.068>

Turner, A. J. (2015). ACE2 Cell Biology, Regulation, and Physiological Functions. *The Protective Arm of the Renin Angiotensin System (RAS)*, 185. <https://doi.org/10.1016/B978-0-12-801364-9.00025-0>

University of Rochester Medical Center, *5-Hydroxyindoleacetic Acid (Urine) – Health Encyclopedia*. (n.d.). Retrieved March 17, 2023, from https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=167&contentid=5_hiaa

Verenitch, S. S., Lowe, C. J., & Mazumder, A. (2006). Determination of acidic drugs and caffeine in municipal wastewaters and receiving waters by gas chromatography-ion trap tandem mass spectrometry. *Journal of Chromatography. A*, 1116 (1–2), 193–203. <https://doi.org/10.1016/J.CHROMA.2006.03.005>

Wallis, C., Henderson, M., & Melnick, J. L. (1972). Enterovirus Concentration on Cellulose Membranes. *Applied Microbiology*, 23(3), 476. <https://doi.org/10.1128/AM.23.3.476-480.1972>

Wartell, B. A., Ballare, S., Ghandehari, S. S., Arcellana, P. D., Proano, C., Kaya, D., Niemeier, D., & Kjellerup, B. v. (2022). Relationship between SARS-CoV-2 in wastewater and clinical data from five wastewater sheds. *Journal of Hazardous Materials Advances*, 8, 100159. <https://doi.org/10.1016/J.HAZADV.2022.100159>

West, N. W., Vasquez, A. A., Bahmani, A., Khan, M. F., Hartrick, J., Turner, C. L., Shuster, W., & Ram, J. L. (2022). Sensitive detection of SARS-CoV-2 molecular markers in urban community sewersheds using automated viral RNA purification and digital droplet PCR. *Science of The Total Environment*, 847, 157547. <https://doi.org/10.1016/J.SCITOTENV.2022.157547>

WHO. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483), 1260–1263.

https://doi.org/10.1126/SCIENCE.ABB2507/SUPPL_FILE/ABB2507_REPRODUCIBILITY_CHECKLIST.PDF

Wu, C., Liu, Y., Yang, Y., Zhang, P., Zhong, W., Wang, Y., Wang, Q., Xu, Y., Li, M., Li, X., Zheng, M., Chen, L., & Li, H. (2020). Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, 10(5), 766–788. <https://doi.org/10.1016/J.APSB.2020.02.008>

Wurtzer, S., Waldman, P., Ferrier-Rembert, A., Frenois-Veyrat, G., Mouchel, J. M., Boni, M., Maday, Y., Marechal, V., & Moulin, L. (2021). Several forms of SARS-CoV-2 RNA can be detected in wastewaters: Implication for wastewater-based epidemiology and risk assessment. *Water Research*, 198. <https://doi.org/10.1016/J.WATRES.2021.117183>

Ye, Y., Ellenberg, R. M., Graham, K. E., & Wigginton, K. R. (2016). Survivability, Partitioning, and Recovery of Enveloped Viruses in Untreated Municipal Wastewater. *Environmental Science & Technology*, 5077–5085. <https://doi.org/10.1021/acs.est.6b00876>

Zhang, N., Gong, Y., Meng, F., Bi, Y., Yang, P., & Wang, F. (2020). Virus shedding patterns in nasopharyngeal and fecal specimens of COVID-19 patients. *MedRxiv*, 2020.03.28.20043059. <https://doi.org/10.1101/2020.03.28.20043059>

Zhou, P., Yang, X. lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. di, Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., ... Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270. <https://doi.org/10.1038/S41586-020-2012-7>

Zhu, C., He, G., Yin, Q., Zeng, L., Ye, X., Shi, Y., & Xu, W. (2021). Molecular biology of the SARs-CoV-2 spike protein: A review of current knowledge. *Journal of Medical Virology*, 93(10), 5729. <https://doi.org/10.1002/JMV.27132>

