



GOBIERNO  
DE ESPAÑA

VICEPRESIDENCIA  
CUARTA DEL GOBIERNO  
MINISTERIO  
PARA LA TRANSICIÓN ECOLÓGICA  
Y EL RETO DEMOGRÁFICO

MINISTERIO  
DE SANIDAD



## CONTROL MICROBIOLÓGICO EN AGUAS RESIDUALES COMO INDICADOR EPIDEMIOLÓGICO DE ALERTA TEMPRANA DE PROPAGACIÓN DE COVID-19

### NOTA TÉCNICA EXPLICATIVA SOBRE LA TÉCNICA Y VARIABILIDAD DE LOS RESULTADOS



## RESUMEN EJECUTIVO

El presente documento pretende exponer diferentes consideraciones técnicas a tener en cuenta a la hora de interpretar los resultados proporcionados dentro del proyecto VATar-COVID-19 en relación con la medida de las concentraciones del ARN de SARS-CoV-2 en las aguas residuales.

Esta nota técnica busca orientar a las personas receptoras de los resultados sobre cómo utilizar los mismos dentro de las limitaciones presentes en el análisis del virus en las aguas residuales, para que sean tratados y comunicados de la forma más correcta posible. Asimismo, se busca explicar qué factores influyen en los resultados que se aportan, y en qué medida y situaciones hay que tenerlos en cuenta a la hora de interpretar los datos.

Estas limitaciones derivan de las características de las redes de saneamiento, el desarrollo actual de la técnica y los métodos analíticos utilizados, el tratamiento y procesamiento de las determinaciones, su comparación con otros datos disponibles, y la comunicación de los resultados obtenidos.

Las principales conclusiones de la nota técnica son las siguientes:

1. Esta técnica ofrece dos aplicaciones principales:
  - a. Por un lado, nos permite detectar si el virus está en una comunidad en función de la presencia o ausencia del mismo en la red de saneamiento. De esta manera, una vez los niveles analizados superen el límite de detección de la técnica, podremos determinar que, como mínimo, hay personas infectadas en la comunidad que están excretando virus en las heces. Esta aplicación resulta fundamental cuando el virus llega a la comunidad, para anticiparse al inicio de una posible epidemia.
  - b. Por otro lado, en una fase más avanzada de la posible epidemia, este sistema nos permite conocer si el número de personas infectadas aumentan o disminuyen. Por tanto, saber cómo se está propagando el virus en la población, en función de la evolución de los niveles detectados.
2. No obstante lo anterior, aún hay incertidumbres científicas sobre, por ejemplo, el tiempo y la cantidad exacta de virus que cada persona excreta en las heces. Los estudios publicados hasta el momento apuntan que el virus se puede detectar en muestras de heces durante periodos más largos que en muestras nasofaríngeas, pero con una importante variabilidad, puesto que no todos los pacientes COVID-19 positivos excretan ARN del SARS-CoV-2 en las heces y, cuando ocurre, la cantidad de virus y la duración de la eliminación varían entre los individuos y a lo largo del tiempo. Además, esta técnica detecta el virus excretado por personas asintomáticas, presintomáticas, sintomáticas y mal diagnosticadas, de ahí la anticipación también.
3. Igualmente, debido al desarrollo actual de la técnica y a las incertidumbres que todavía rodean al SARS-CoV-2, hay que tener en cuenta algunas limitaciones a la hora de buscar relaciones entre los niveles detectados en las aguas residuales y los casos clínicos de COVID-19 diagnosticados. Asimismo, hay que buscar estas relaciones con cautela, tanto por la forma de identificar los casos reales como por la influencia de otras variables ambientales (por ejemplo, eventos de lluvia, temperatura, tiempo de retención hidráulica en alcantarillas, transporte de muestras al laboratorio, hora de muestreo, etc.) que aumentan las incertidumbres vinculadas a la correlación entre la detección del ARN del SARS-CoV-2 en muestras de aguas residuales.

4. Los resultados obtenidos tras el análisis de las muestras son cuantificados por los laboratorios como copias genómicas de SARS-CoV-2 por litro (cg/l). Debido a las magnitudes y rangos de los datos, las características de las poblaciones microbianas, y la forma de propagación de las enfermedades infecciosas en la población, estos datos se transforman a escala logarítmica ( $\log_{10}$  cg/l). Para calcular la evolución respecto a la semana anterior, la variación se calcula a partir de la diferencia de cuantificación en unidades logarítmicas respecto a la semana anterior, clasificándose en las siguientes categorías:

Variación	Unidades logarítmicas de diferencia
<b>Aumento significativo</b>	Más de +1 unidad logarítmica
<b>Aumento</b>	Entre +0,4 y +1
<b>Estable</b>	Entre -0,4 y +0,4
<b>Disminución</b>	Entre -0,4 y -1
<b>Disminución significativa</b>	Más de -1

5. Como resultado de lo anteriormente expuesto, es esencial tomar con cautela un dato puntual de la presencia de SARS-CoV-2 en las aguas residuales, siendo lo más importante analizar las tendencias semanales y la correlación global con la evolución de la situación sanitaria. Pequeñas variaciones semanales pueden ser debidas a las variaciones tanto por condicionantes ambientales como por pequeñas variaciones en relación con las variaciones en las horas de muestreo, por lo que la tendencia semana tras semana es la variable que mejor informa de la evolución de la situación. Se espera afinar el sistema de alerta temprana con el desarrollo progresivo de la técnica, disminuyendo la influencia de estos condicionantes y las incertidumbres presentes.



## Contenido

Resumen ejecutivo .....	2
I. Información relativa a la interpretación de resultados .....	5
1) Breve introducción a la epidemiología del SARS-CoV-2 en aguas residuales.....	5
2) Dianas utilizadas.....	6
3) Criterios de positividad de las muestras.....	8
4) Cálculos de concentración de copias genómicas.....	8
5) Conceptos relativos a la técnica utilizada .....	9
6) Presentación de los resultados y escalas utilizadas.....	10
a) Escalas utilizadas .....	10
b) Cálculo de la variación entre semanas y categorías utilizadas.....	10
7) Análisis de variantes de SARS-CoV-2.....	11
a) Análisis de variantes conocidas de SARS-CoV-2 mediante RT-qPCR dúplex. ....	11
b) Búsqueda de variantes de SARS-CoV-2 mediante secuenciación masiva .....	12
II. Relación con los casos clínicos notificados .....	16
1) Planteamiento teórico .....	16
2) Aplicaciones prácticas .....	17
3) Limitaciones para relacionar los datos epidemiológicos con las dianas del virus .....	18
III. Factores que influyen en los niveles detectados en aguas residuales.....	21
1) Factores ambientales .....	21
2) Características de la cuenca tributaria de la EDAR y de la red de saneamiento .....	21
3) Variaciones en las características estándar del agua residual .....	21
4) Muestreo en EDAR.....	22
5) Análisis del momento óptimo de muestreo .....	23
6) Logística .....	24
7) Variaciones en la excreción del virus por parte de las personas infectadas .....	24
8) Degradación del material genético del virus.....	25
9) Factores relativos a los métodos analíticos .....	26

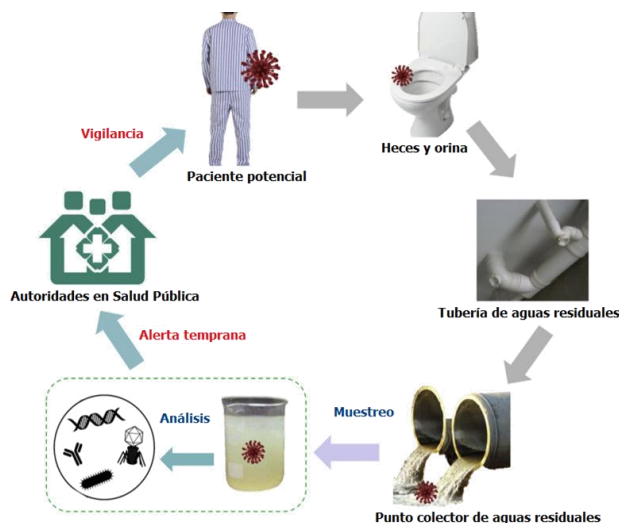
## I. Información relativa a la interpretación de resultados

### 1) Breve introducción a la epidemiología del SARS-CoV-2 en aguas residuales

La epidemiología basada en aguas residuales (*WBE: Wastewater-based epidemiology*, en inglés) es una técnica que, en el caso que nos ocupa, busca conocer la presencia y circulación del SARS-CoV-2 en una comunidad, a partir del material genético del virus que se detecta en las aguas residuales. A partir de esos niveles detectados y su variación a lo largo del tiempo podemos evidenciar la circulación del virus dentro de la comunidad.<sup>1</sup> Esto es posible porque un porcentaje elevado de personas infectadas, ya sean sintomáticas, asintomáticas, presintomáticas o postsintomáticas, o incluso sintomáticas sin diagnosticar, excreta en sus heces partículas víricas que se vierten a la red de saneamiento.<sup>2</sup>

Dentro de esta red de saneamiento se establecen puntos de muestreo, que pueden situarse en diferentes localizaciones hasta la entrada de la Estación Depuradora de Aguas Residuales (EDAR), donde se recogen muestras de estas aguas residuales. Estas muestras son enviadas al laboratorio, donde se determina la cantidad de virus presente analizando el material genético (ARN en este caso) que se detecta mediante la técnica denominada "Reacción en cadena de la polimerasa cuantitativa con transcripción inversa" (RT-qPCR, por sus siglas en inglés).

Los niveles detectados nos permiten conocer si el número de infectados en la comunidad aumenta o disminuye. Sin embargo, estos niveles detectados también pueden variar debido a multitud de factores a lo largo del proceso, que pueden distorsionarnos la información, no solo por el número de personas infectadas. La influencia de estos factores se va a explicar a lo largo del documento.



**Figura 1.** Resumen gráfico sobre la epidemiología basada en aguas residuales del SARS-CoV-2. Modificado de Mao *et al.*<sup>3</sup>

<sup>1</sup> Aguiar-Oliveira ML, Campos A, R Matos A, Rigotto C, Sotero-Martins A, Teixeira PFP, Siqueira MM. (2020). Wastewater-Based Epidemiology (WBE) and Viral Detection in Polluted Surface Water: A Valuable Tool for COVID-19 Surveillance-A Brief Review. *Int J Environ Res Public Health*. 17 (24): 9251 <https://doi.org/10.3390/ijerph17249251>

<sup>2</sup> Chavarria-Miró G, Anfruns-Estrada E, Martínez-Velázquez A, Vázquez-Portero M, Guix S, Paraira M, Galofré B, Sánchez G, Pintó RM, Bosch A. (2021). Time Evolution of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Wastewater during the First Pandemic Wave of COVID-19 in the Metropolitan Area of Barcelona, Spain. *Appl Environ Microbiol*. 87 (7): e02750-20. <https://doi.org/10.1128/aem.02750-20>

<sup>3</sup> Mao K, Zhang K, Du W, Ali W, Feng X, Zhang H. (2020). The potential of wastewater-based epidemiology as surveillance and early warning of infectious disease outbreaks. *Curr Opin Environ Sci Health*. 17:1-7. <https://doi.org/10.1016/j.coesh.2020.04.006>

## 2) Dianas utilizadas

Para realizar la detección y cuantificación del SARS-CoV-2 en las muestras de agua residual se emplea la técnica denominada "Reacción en cadena de la polimerasa cuantitativa con transcripción inversa" (RT-qPCR, por sus siglas en inglés).

Esta técnica se basa en amplificar regiones del material genético del virus, las denominadas dianas. Estas dianas son secuencias específicas idealmente presentes únicamente en ese virus. Las dianas del SARS-CoV-2 utilizadas dentro del proyecto VATar COVID-19 son:

- **IP4:** Región de la secuencia del material genético que codifica para la proteína "ARN polimerasa dependiente de ARN" (se abrevia como RdRp por sus siglas en inglés, y así se puede encontrar en otras fuentes). Se cuantifica de acuerdo a protocolos propios de los laboratorios, desarrollados para esta diana a partir del propuesto por el Instituto Pasteur<sup>4</sup>.
- **E:** Secuencia del material genético que codifica para la proteína de la envoltura del virus. Se cuantifica de acuerdo a protocolos propios de los laboratorios involucrados, desarrollados para esta diana a partir del propuesto por Corman VM *et al* (2020)<sup>5</sup>.
- **N1:** Secuencia del material genético que codifica para la nucleocápside del virus. Se cuantifica de acuerdo a protocolos propios de los laboratorios, desarrollados para esta diana a partir del propuesto por el CDC (Centro de control y prevención de enfermedades de Estados Unidos)<sup>6</sup>.

Para la detección de las variantes de preocupación ("VOC: *Variants of Concern*") se emplea la técnica de RT-PCR cuantitativa en tiempo real dúplex (*duplex RT-qPCR*, en inglés), la cual permite la amplificación y cuantificación simultáneas de dos secuencias dianas en un único análisis de RT-qPCR<sup>7</sup>. Además, la RT-PCR dúplex se basa en la detección y estimación de proporciones de mutaciones específicas de las principales variantes de preocupación, presentes concretamente en la diana S:

- **S:** secuencia del material genético que codifica para la glicoproteína espicular (*spike*), cuya función principal es mediar la entrada del virus en la célula hospedadora.<sup>8</sup>

<sup>4</sup> Institut Pasteur, Paris. Protocol: Real-time RT-PCR assays for the detection of SARS-CoV-2. 2020.

<https://www.who.int/docs/default-source/coronaviruse/real-time-rt-pcr-assays-for-the-detection-of-sars-cov-2-institut-pasteur-paris.pdf>

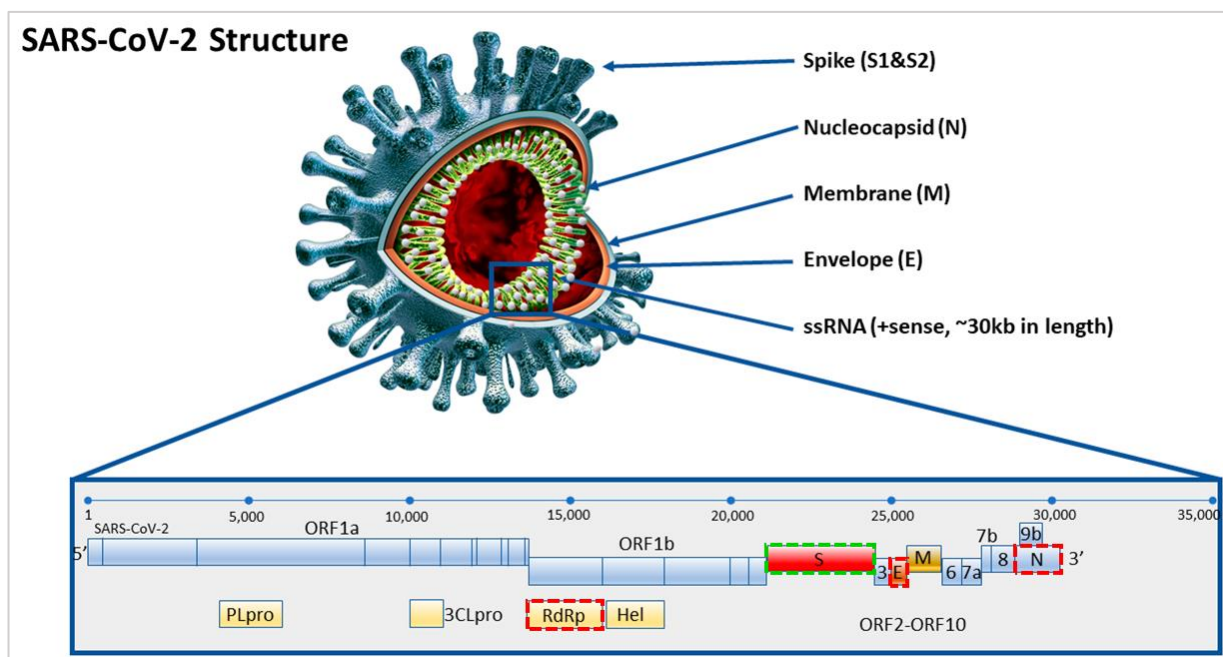
<sup>5</sup> Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DG, Haagmans BL, van der Veer B, van den Brink S, Wijsman L, Goderski G, Romette JL, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MP, Drosten C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill.* 25 (3): 2000045. <https://dx.doi.org/10.2807%2F1560-7917.ES.2020.25.3.2000045>

<sup>6</sup> CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel.

<https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html>

<sup>7</sup> Carcereny A, Martínez-Velázquez A, Bosch A, Allende A, Truchado P, Cascales J, Romalde JL, Lois M, Polo D, Sánchez G, Pérez-Cataluña A, Díaz-Reolid A, Antón A, Gregori J, Garcia-Cehic D, Quer J, Palau M, Ruano CG, Pintó RM, Guix S. (2021) Monitoring Emergence of the SARS-CoV-2 B.1.1.7 Variant through the Spanish National SARS-CoV-2 Wastewater Surveillance System (VATar COVID-19). *Environ Sci Technol.* 55 (17): 11756-11766. <https://dx.doi.org/10.1021%2Facs.est.1c03589>

<sup>8</sup> Almehdi AM, Khoder G, Alchakee AS, Alsayyid AT, Sarg NH, Soliman SSM. (2021). SARS-CoV-2 spike protein: pathogenesis, vaccines, and potential therapies. *Infection.* 49 (5): 855-876. <https://dx.doi.org/10.1007%2Fs15010-021-01677-8>



**Figura 2.** Principales dianas utilizadas para la detección del SARS-CoV-2 y sus variantes. Resaltado con líneas discontinuas de color rojo - - - -: Dianas utilizadas para la RT-qPCR y detectar la presencia del virus. Resaltado con líneas discontinuas de color verde - - - -: Diana utilizada para la RT-qPCR dúplex y detectar las variantes de preocupación. Figura obtenida de Kubina & Dziedzic<sup>9</sup>.

<sup>9</sup> Kubina R & Dziedzic A (2020). Molecular and Serological Tests for COVID-19 a Comparative Review of SARS-CoV-2 Coronavirus Laboratory and Point-of-Care Diagnostics. *Diagnostics*. 10 (6), 434. <https://doi.org/10.3390/diagnostics10060434>

### 3) Criterios de positividad de las muestras

El análisis de la presencia del ARN del virus a través de estas tres dianas nos permite determinar a las muestras como positivas, negativas o presuntas positivas. Dentro del proyecto VATar COVID-19, los criterios de positividad que se están aplicando son los siguientes:

- **Positivo**→Se determina como positivo aquella muestra que presenta un resultado positivo (por encima del límite de detección) para al menos dos dianas del virus (aunque la detección sea inferior al límite de cuantificación), o un positivo robusto para una de las dianas del virus.
- **Negativo**→Se determina como negativo aquella muestra con un resultado negativo para las dianas del virus analizadas.
- **Presunto positivo**→ Solo hay detección en una de las dianas, y esta no es robusta, siendo una diana positiva, y las otras dos negativas. Se entiende por resultado robusto cuando hay varias réplicas con detección dentro de la diana, los valores de cada determinación son coherentes entre sí, entre las distintas réplicas y con la evolución de la EDAR en semanas anteriores.

### 4) Cálculos de concentración de copias genómicas

- Se convertirán los valores de  $Cq < 40$  de cada pocillo a copias genómicas/L (cg/L) interpolando el  $Cq$  de la muestra problema a las rectas patrón externas construidas con diluciones seriadas en base 10 del material de referencia cuantificado. La recta patrón de cada gen debe tener un coeficiente de correlación mínimo ( $r^2$ ) de 0,98 y la pendiente de la recta debe estar entre 3,10 y 3,60 (corresponde a unas eficiencias de amplificación del 90 al 110 %) (ISO 15216-1, 2017).
- Cuando en el análisis de las 2 réplicas de alguna dilución una de las réplicas tiene un valor  $Cq < 40$  y la otra de "No  $Cq$ " o  $Cq \geq 40$ , a esta última réplica se le asignará una concentración igual al límite de detección (LD) teórico de 1 copia genómica/reacción.
- Para aquellas muestras para las que se ha obtenido valores de  $Cq < 40$ , se calcula el promedio y el error estándar de las concentraciones estimadas a partir de los resultados de las 4 réplicas (2 de ácido nucleico sin diluir y 2 de la dilución 1/10), siguiendo los siguientes criterios y escenarios:
  - Cuando únicamente se obtiene  $Cq < 40$  en una de las dos réplicas de una de las dos diluciones → se hará el promedio de las dos réplicas de esa dilución.
  - Si se obtienen valores de  $Cq < 40$  en, al menos, una réplica de cada una de las dos diluciones → se hará el promedio de las 4 réplicas siempre y cuando la diferencia entre la concentración de la muestra directa y la diluida -1 sea inferior a  $0,5 \log_{10}$ .
  - Cuando la concentración estimada a partir de la dilución -1 sea  $0,5 \log_{10}$  superior a la concentración estimada a partir de la directa, se considera que hay inhibición y se calcula la concentración utilizando únicamente los valores de  $Cq$  de la dilución -1.



## 5) Conceptos relativos a la técnica utilizada

Algunos conceptos recurrentes a la hora de interpretar los resultados son los siguientes:

**Límite de detección:** Concentración mínima a la cual se puede determinar el analito de manera confiable. Generalmente, se expresa como la concentración del analito a la cual el método lo puede detectar en un 95 % de los casos.

**Límite de cuantificación:** Es la mínima cantidad o concentración de analito que se puede cuantificar con precisión y exactitud, de manera reproducible y fiable. Los límites de cuantificación aplicados actualmente por los distintos laboratorios son los siguientes:

LABORATORIO/DIANAS	IP4 (log <sub>10</sub> cg/L)	E (log <sub>10</sub> cg/L)	N1 (log <sub>10</sub> cg/L)
IATA	4,47	4,88	4,23
CEBAS	4,29	4,43	4,26
USC	4,14	4,45	4,38
UB	4,47	4,74	4,28

**Tabla 1.** Límites de cuantificación de cada laboratorio para cada una de las dianas analizadas

### **Cq: Ciclo de cuantificación**

En la RT-qPCR se amplifica el material genético en repetidos ciclos (hasta 50). La cuantificación se produce cuando la cantidad de material genético supera un umbral marcado. El "umbral de ciclo" (Ct: *threshold cycle*) o "ciclo de cuantificación" (Cq: *quantification cycle*) indica el ciclo a partir del cual se comienza a detectar fluorescencia por encima del ruido de fondo, es decir, cuántos ciclos se necesitaron para detectar una señal real de la muestra analizada y, consecuentemente, que esta es positiva para lo que se buscaba (siendo SARS-CoV-2 en este caso). Por lo tanto, para un mismo marcador, valores menores de Cq se asocian a mayores cantidades de material genético.<sup>10</sup>

### Interpretación de los resultados:

- No Cq: Resultado negativo
- Cq mayor de 40: Resultado negativo
- Cq menor de 40: Positivo. Se relaciona con una cantidad determinada de virus en la muestra, especificada en la columna de Concentración (cg/L)

**Sensibilidad:** Se refiere a la cantidad mínima de ARN necesaria para que se produzca la amplificación. Se relaciona con los falsos negativos, ya que puede que una muestra sea positiva, pero sea dada como negativa porque no se ha amplificado por no tener suficiente cantidad de ARN.

**Especificidad:** Se refiere a la obtención de un solo producto amplificado. Viene determinada por los oligos y la especificidad con la que se unen al ARN molde. Se relaciona con los falsos positivos, ya que puede que una muestra sea negativa, pero sea dada como positiva porque se ha amplificado una región de ARN no diana o que no se buscaba amplificar.

<sup>10</sup> García Marín AM, Goig Serrano GA. (2020). Evaluación del ciclo umbral (Ct) de la qPCR como valor predictivo de la calidad de secuenciación. *SeqCOVID. CSIC*. <https://seqcovid.csic.es/es/evaluacion-del-ciclo-umbral-ct-de-la-qpcr-como-valor-predictivo-de-la-calidad-de-secuenciacion-2/>

## 6) Presentación de los resultados y escalas utilizadas

### a) Escalas utilizadas

Los resultados obtenidos tras el análisis de las muestras son cuantificados por los laboratorios como copias genómicas de SARS-CoV-2 por litro (cg/l). Debido a las magnitudes y rangos de los datos, las características de las poblaciones microbianas, y la forma de propagación de las enfermedades infecciosas en la población, estos datos se suelen utilizar en la escala logarítmica ( $\log_{10}$  cg/l). Esta escala utiliza el logaritmo en base 10 del valor determinado. Mientras que los niveles son detectados por el laboratorio en escala decimal.

**ADVERTENCIA:** Se recomienda prestar mucha atención la escala en la que se presentan los datos, antes de interpretarlos y comunicarlos. La recepción y tratamiento de los datos usando una u otra escala puede propiciar interpretaciones diferentes de los mismos datos.

Esto también es importante a la hora de representar los datos en las gráficas. Usando la escala decimal el espacio representado corresponde al valor absoluto de la determinación. Esto es, el espacio ocupado en la representación es el mismo para las variaciones de 10 a 20 que de 20 a 30.

En cambio, usando la escala logarítmica el gráfico se construye tomando variaciones relativas o porcentuales. De esta manera el espacio ocupado con una variación de 10 a 20, donde tenemos una variación del 100%, se representará como un espacio mayor que en el caso de 20 a 30, donde la variación es del 50%. Al usar los logaritmos decimales, con base 10, cada unidad de aumento representada con esta escala corresponde a un factor de 10.

Por tanto, aplicando la escala decimal la variación se representa de forma lineal, mientras que con la escala logarítmica se representa con el índice logarítmico (en este caso en base 10).

Dentro del proyecto VATar-COVID-19 se utilizan distintos tipos de gráficos usando las distintas escalas. Tanto el tipo de gráfico como la escala varía en función de la información que se quiera mostrar y el objetivo. Por ejemplo, la escala logarítmica que se suele utilizar para representar los niveles de copias genómicas detectadas se aplica, además de por lo ya comentado, cuando se pretenden visualizar las tendencias y evoluciones de forma más intuitiva.

### b) Cálculo de la variación entre semanas y categorías utilizadas

Para calcular la evolución respecto a la semana anterior, se utilizan aquellas dianas con detección en ambas semanas. La variación se calcula a partir de la diferencia de cuantificación en unidades logarítmicas respecto a la semana anterior. Así, se categoriza cualitativamente la variación teniendo en cuenta los umbrales de incertidumbre de la técnica. Para definir las categorías se aplican los intervalos que aparecen en la siguiente tabla:

Variación	Unidades logarítmicas de diferencia
<b>Aumento significativo</b>	Más de +1 unidad logarítmica
<b>Aumento</b>	Entre +0,4 y +1
<b>Estable</b>	Entre -0,4 y +0,4
<b>Disminución</b>	Entre -0,4 y -1
<b>Disminución significativa</b>	Más de -1

**Tabla 2.** Categorías de variación aplicadas e intervalos de variación entre semanas

## 7) Análisis de variantes de SARS-CoV-2

### a) Análisis de variantes conocidas de SARS-CoV-2 mediante RT-qPCR dúplex.

Para la detección de las variantes de preocupación de SARS-CoV-2 más relevantes en el momento del estudio se emplea la técnica de RT-qPCR dúplex. Esta permite la amplificación y cuantificación simultánea de dos secuencias diana en un único análisis RT-qPCR.

Los laboratorios del proyecto VATar-COVID-19 emplean la RT-qPCR dúplex para detectar mutaciones presentes en el gen S (que codifica para la glicoproteína de la espícula), las cuales sirven para la identificación de las variantes de preocupación de mayor relevancia en cada momento<sup>11, 12</sup>. Es decir, para cada muestra se realiza un análisis por cada variante que se quiere identificar, para el cual se busca una mutación específica e identitaria para la variante en cuestión, tras ello, se estima la proporción (%) en la que está dicha mutación (y por tanto dicha variante) en la muestra. Por ejemplo, se realizaría una RT-qPCR dúplex para detectar y conocer la proporción en la que está en la muestra la variante Delta, obteniéndose un resultado entre el 0 y el 100%, y otra RT-qPCR dúplex para detectar y conocer la proporción en la que está en la muestra el sublinaje BA.1 de la variante Ómicron, obteniéndose un resultado entre el 0 y el 100%. Hay que tener en cuenta que, como la cuantificación de cada variante se realiza mediante análisis independientes el cómputo total de la proporción de todas las variantes detectadas en una muestra puede ser superior o inferior al 100%, tomando con especial cautela aquellos porcentajes obtenidos a partir de datos inferiores al límite de cuantificación (marcados en rojo).

**ADVERTENCIA:** En el caso de no ser posible el análisis de mutaciones específicas e identitarias para una sola variante o sublinaje (en la mayoría de los casos debido a que dos o más variantes comparten las mutaciones presentes en el gen S), se realizará el análisis por RT-qPCR dúplex de la mutación presente en el menor número de variantes o sublinajes, indicándose de cuáles se trata. Se debe tener esto en cuenta a la hora de interpretar los resultados del análisis.

En función al porcentaje resultante del análisis, podemos hacer una estimación de la proporción en la que se encuentra la variante en cuestión en la muestra:

- 0 %: No se detecta la variante.
- 1-30 %: Se detecta la variante en una proporción baja.
- 31-75 %: Se detecta la variante en una proporción elevada.
- 76-100 %: Se detecta en una proporción predominante.
- Ne (No estimado): los valores del análisis para la detección de la mutación específica de la variante son inferiores al límite de cuantificación; por ello, aunque se encuentra esta mutación en la muestra, no se puede estimar la proporción en la que se está.
- En el caso de los resultados marcados en rojo (obtenidos a partir de datos inferiores al límite de cuantificación) no se realiza una estimación de la proporción, solo se indica su detección.

Como resultado de estos análisis, sabremos en qué proporción está(n) la(s) variante(s) de SARS-CoV-2 en el área de influencia de la EDAR de la que se obtuvo la muestra analizada, y si hay una clara predominancia de una variante en concreto.

<sup>11</sup> Ver referencia en nota 7.

<sup>12</sup> Carcereny A, Garcia-Pedemonte D, Martínez-Velázquez A, Quer J, Garcia-Cehic D, Gregori J, Antón A, Andrés C, Pumarola T, Chacón-Villanueva C, Borrego CM, Bosch A, Guix S, Pintó RM. (2022). Dynamics of SARS-CoV-2 Alpha (B.1.1.7) variant spread: The wastewater surveillance approach. *Environ Res.*208: 112720. <https://doi.org/10.1016/j.envres.2022.112720>

## b) Búsqueda de variantes de SARS-CoV-2 mediante secuenciación masiva

Se ha observado que SARS-CoV-2 ha ido evolucionando en diferentes variantes, las cuales se clasifican en las siguientes categorías<sup>13</sup>: «variantes de preocupación» (VOC: *Variants of Concern*), «variantes de interés» (VOI: *Variants of Interest*), «variantes bajo vigilancia» (VUM: *Variants under Monitoring*) y «variantes desescaladas».

Las características<sup>14</sup> que presentan cada una de las variantes para pertenecer a una u otra categoría son:

- Para ser considerada VOC tiene que:
  - tener cambios genéticos ya predichos o conocidos que afectan a las características del virus, como la transmisibilidad, gravedad de la enfermedad, escape inmunitario, diagnóstico inmunitario;
  - estar identificada para causar transmisión comunitaria significativa o múltiples grupos de COVID-19, en varios países con una prevalencia relativa creciente junto con un número creciente de casos a lo largo del tiempo, u otros impactos epidemiológicos aparentes que sugieren un riesgo emergente para la salud pública;
  - aumentar la transmisibilidad o cambiar perjudicialmente la epidemiología de COVID-19;
  - aumentar la virulencia o cambiar el cuadro clínico de la enfermedad;
  - disminuir la eficacia de las medidas sociales y de salud pública o de los diagnósticos, vacunas y tratamientos disponibles.
- Para ser considerada VOI tiene que:
  - tener cambios genéticos ya predichos o conocidos que afectan a las características del virus, como la transmisibilidad, gravedad de la enfermedad, escape inmunitario, diagnóstico inmunitario;
  - estar identificada para causar transmisión comunitaria significativa o múltiples grupos de COVID-19, en varios países con una prevalencia relativa creciente junto con un número creciente de casos a lo largo del tiempo, u otros impactos epidemiológicos aparentes que sugieren un riesgo emergente para la salud pública.
- Para ser considerada VUM tiene que tener cambios genéticos que se sospechan que afectan a las características del virus con alguna indicación de que puede representar un riesgo futuro, pero la evidencia del impacto fenotípico o epidemiológico actualmente no está clara, lo que requiere un seguimiento mejorado y una evaluación repetida en espera de nueva evidencia.
- Las «variantes desescaladas» son aquellas que ya no están bajo vigilancia, debido a alguno de los siguientes criterios:
  - la variante ya no circula a niveles de importancia para la salud pública mundial;
  - la variante ha estado circulando durante mucho tiempo sin ningún impacto en la situación epidemiológica general;
  - la evidencia científica demuestra que la variante no está asociada con ninguna propiedad preocupante.

<sup>13</sup> Información actualizada disponible en: <https://www.ecdc.europa.eu/en/covid-19/variants-concern>

<sup>14</sup> OMS - Variantes del SARS-CoV-2: <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>

Los laboratorios del proyecto VATar-COVID-19 han establecido los siguientes parámetros o decisiones que determinan si una muestra de agua residual recogida a la entrada de la EDAR es apta o no para su posterior análisis por secuenciación masiva:

- 1º) Se seleccionan las muestras que tengan un valor de Cq menor a 34, en base a lo publicado por Izquierdo-Lara, *et al.* (2021)<sup>15</sup>. Según este estudio, tras realizar una secuenciación masiva de nanoporos (*Oxford Nanopore Technologies*) en las muestras tomadas de aguas residuales comunitarias, los valores de Cq en los que se pudo obtener la mitad del genoma viral fueron de 34,6 para el gen N1 (que es la misma diana usada en este proyecto). De ahí que los laboratorios hayan establecido ese valor como el máximo permitido.
- 2º) Se procesan aquellas muestras cuya concentración tras la limpieza de amplicones<sup>16</sup> sea mayor a 4 ng/μL, para así conseguir una concentración equimolar y trabajar con 50 ng/μL por muestra, garantizando que cada muestra genere suficientes productos de amplificación para la secuenciación.
- 3º) Además, se comprueba la presencia de amplicón del tamaño esperado mediante análisis por *TapeStation*<sup>17</sup>.

**ADVERTENCIA:** Se ha de tener en cuenta que podría darse el caso de que ninguna de las muestras recogidas en una EDAR durante un mes cumpla estos criterios, no habiendo así ninguna muestra apta para ser secuenciada ese mes.

El análisis de las muestras aptas se realiza mediante secuenciación masiva (HTS: *High-Throughput Sequencing*) con la plataforma MinION de la empresa Oxford Nanopore Technologies. Esta tecnología se basa en aprovechar los perfiles de corriente eléctrica que genera cada molécula (nucleótidos) al pasar a través de un poro proteico de tamaño nanométrico («nanoporo») que sirve como un biosensor y está embebido en una membrana polimérica resistente a la electricidad, a cuyos lados hay establecido un diferencial de voltaje, pudiéndose detectar así la secuencia de una hebra de ADN o ARN (como es el caso del SARS-CoV-2) de manera directa; entonces, a medida que los nucleótidos pasan a través del nanoporo, se mide un cambio de corriente característico y se utiliza para determinar el tipo de nucleótido y por tanto, conocer la secuencia de material genético.<sup>18, 19</sup>

---

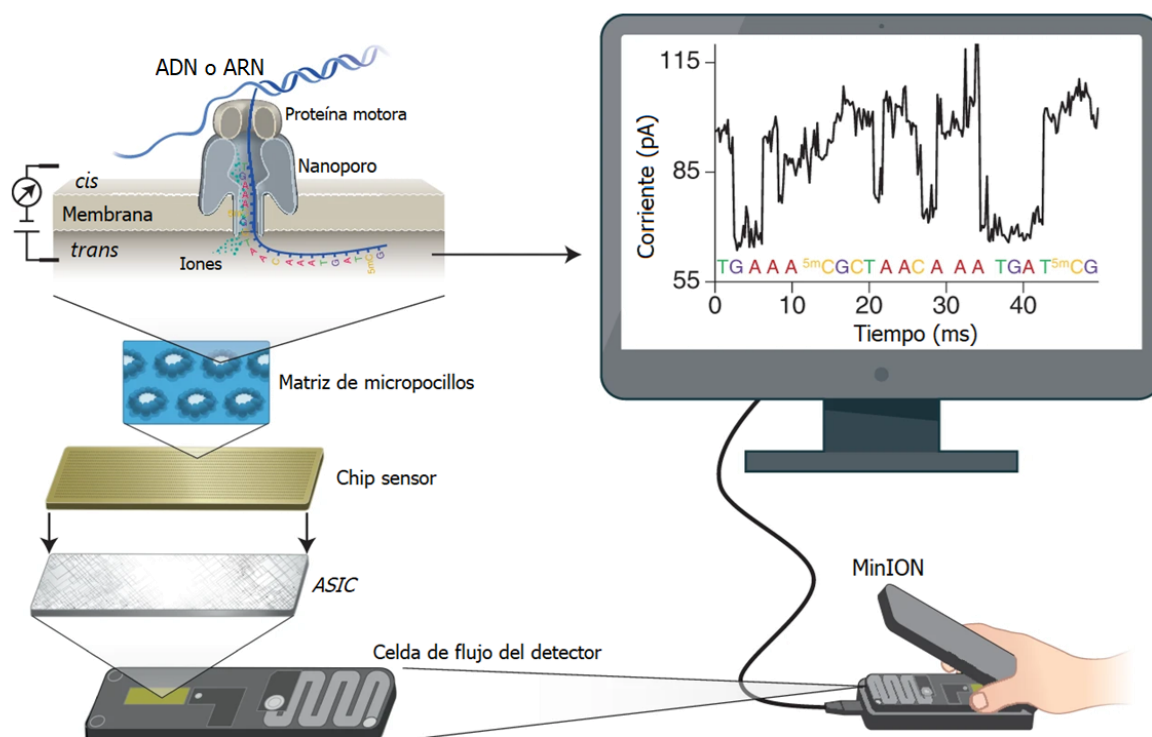
<sup>15</sup> Izquierdo-Lara R, Elsinga G, Heijnen L, Munnink BBO, Schapendonk CME, Nieuwenhuijse D, Kon M, Lu L, Aarestrup FM, Lycett S, Medema G, Koopmans MPG, de Graaf M. (2021). Monitoring SARS-CoV-2 Circulation and Diversity through Community Wastewater Sequencing, the Netherlands and Belgium. *Emerg Infect Dis.* 27 (5): 1405-1415. <https://doi.org/10.3201/eid2705.204410>

<sup>16</sup> Fragmento de material genético amplificado por PCR.

<sup>17</sup> Instrumento de la empresa Agilent. Es un sistema con soluciones de electroforesis automatizada para el control de calidad de material genético. Es adecuado para analizar el tamaño, la cantidad y la integridad de las muestras tanto de ADN como ARN. <https://www.agilent.com/en/product/automated-electrophoresis/tapestation-systems>

<sup>18</sup> Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggin M, Schloss JA. (2008). The potential and challenges of nanopore sequencing. *Nat Biotechnol.* 26 (10): 1146-1153. <https://doi.org/10.1038/nbt.1495>

<sup>19</sup> Feng Y, Zhang Y, Ying C, Wang D, Du C. (2015). Nanopore-based fourth-generation DNA sequencing technology. *Genomics Proteomics Bioinformatics.* 13 (1): 4-16. <https://doi.org/10.1016/j.gpb.2015.01.009>



**Figura 3.** Representación de la tecnología de secuenciación por nanoporos.

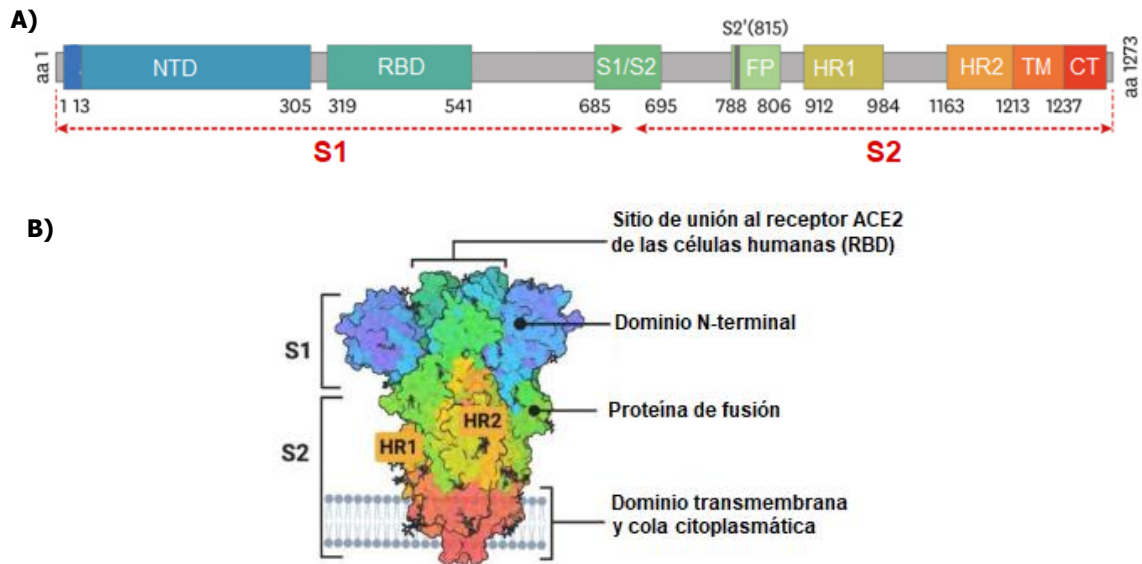
ASIC: *application specific integration circuit* (circuito integrado específico de la aplicación). Modificado de Wang Y *et al.* (2021)<sup>20</sup>

En las mencionadas muestras tomadas mensualmente en cada EDAR, los laboratorios buscan mutaciones presentes en el gen S, que codifica para la glicoproteína de la espícula (*spike*) y su función principal es mediar la entrada del virus en la célula hospedadora.

Las mutaciones que se detectan pueden clasificarse como una de las siguientes:

- **Mutaciones específicas** o identitarias de una de las variantes de preocupación (VOC) o de interés (VOI). Nos permiten saber **qué variante o variantes hay** en la muestra y con qué **probabilidad** están. Actualmente, se detectan mutaciones asociadas las siguientes variantes: B.1.1.7 (Alfa), P.1 (Gamma), B.1.351 (Beta), B.1.617.2 (Delta), C.37 (Lambda) y B.1.1.529 (Ómicron).
- **Misceláneas:** No son mutaciones específicas de una sola variante, pudiendo estar presente en dos o más, por ello no sirven para la identificación de variantes.
- **Mutaciones no asignadas.** Estas mutaciones aún no están asociadas a una variante y por tanto pueden servir como indicativo de la aparición de una nueva variante de SARS-CoV-2.

<sup>20</sup> Wang Y, Zhao, Y., Bollas, A., Wang, Y., & Au, K. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol*, 39(11): 1348-1365. <https://doi.org/10.1038/s41587-021-01108-x>



**Figura 4.**

**A)** Representación esquemática del gen S del SARS-CoV-2.

NTD: *N-terminal domain* (dominio N terminal); RBD: *receptor-binding domain* (dominio de unión al receptor); S2/S2: *protease cleavage site* (lugar de escisión de la proteasa); FP: *fusion peptide* (péptido de fusión); HR1, *heptad repeat 1* (motivo de 7 aminoácidos repetidos); HR2: *heptad repeat 2*; TM: *transmembrane domain* (dominio transmembrana); CT: *cytoplasmic tail* (cola citoplasmática). Figura obtenida de Jia & Gong (2021)<sup>21</sup>

**B)** Representación cristalográfica de la estructura de la glicoproteína de la espícula.

Modificado de Chilamakuri & Agarwal (2021)<sup>22</sup>

<sup>21</sup> Jia Z & Gong W. (2021). Will Mutations in the Spike Protein of SARS-CoV-2 Lead to the Failure of COVID-19 Vaccines? *J Korean Med Sci.* 36 (18): e124. <https://doi.org/10.3346/jkms.2021.36.e124>

<sup>22</sup> Chilamakuri R & Agarwal S. (2021). COVID-19: Characteristics and Therapeutics. *Cells.* 10 (2): 206. <https://doi.org/10.3390/cells10020206>

## II. Relación con los casos clínicos notificados

### 1) Planteamiento teórico

En el caso del SARS-CoV-2, los niveles detectados en aguas residuales se pueden utilizar como un sistema de alerta temprana debido a que las personas infectadas excretan el material genético del virus antes incluso de presentar síntomas clínicos y ser diagnosticado. Por lo tanto, puede ser detectado en las aguas residuales antes incluso de que la propia persona sea consciente de estar infectado, o cuando la persona curse como asintomática.<sup>23, 24</sup>

El tiempo y la cantidad exacta de material genético que cada persona excreta continúa siendo una incógnita. Aunque se necesita más investigación sobre este tema, los estudios publicados hasta el momento apuntan que el virus se puede detectar en muestras de heces durante periodos más largos que en muestras nasofaríngeas. Aun así, más evidencia se necesita para conocer con exactitud el patrón de excreción del virus en las heces durante el proceso infeccioso.<sup>25, 26</sup>

El SARS-CoV-2 se ha aislado de las heces y la orina en los tres días posteriores a la infección. En un estudio que evaluó la dinámica del virus en la provincia de Zhejiang en China, se observó que la duración media del virus en las heces fue de 22 días, que fue significativamente más larga que en las vías respiratorias (18 días) y las muestras de suero (16 días)<sup>27</sup>. Teniendo en cuenta que el material genético del virus se puede detectar hasta varios días fuera del cuerpo humano, su medición en aguas residuales podría proporcionar una herramienta válida y complementaria para el seguimiento preventivo y el diagnóstico de COVID-19 en las comunidades. También se están realizando esfuerzos para analizar muestras históricas de aguas residuales en busca de evidencia de circulación anterior de SARS-CoV-2.<sup>28, 29</sup>

<sup>23</sup> Bonanno Ferraro G, Veneri C, Mancini P, Iaconelli M, Suffredini E, Bonadonna L, Lucentini L, Bowo-Ngandji A, Kengne-Nde C, Mbaga DS, Mahamat G, Tazokong HR, Ebogo-Belobo JT, Njouom R, Kenmoe S, La Rosa G. (2021). A State-of-the-Art Scoping Review on SARS-CoV-2 in Sewage Focusing on the Potential of Wastewater Surveillance for the Monitoring of the COVID-19 Pandemic. *Food Environ Virol.* 2: 1-40. <https://doi.org/10.1007/s12560-021-09498-6>

<sup>24</sup> Mackul'ak T, Gál M, Špalková V, Fehér M, Briestenská K, Mikušová M, Tomčíková K, Tamáš M, Butor Škulcová A. (2021). Wastewater-Based Epidemiology as an Early Warning System for the Spreading of SARS-CoV-2 and Its Mutations in the Population. *Int J Environ Res Public Health.* 18 (11): 5629. <https://doi.org/10.3390/ijerph18115629>

<sup>25</sup> Kitajima M, Ahmed W, Bibby K, Carducci A, Gerba CP, Hamilton KA, Haramoto E, Rose JB. (2020). SARS-CoV-2 in wastewater: State of the knowledge and research needs. *Sci Total Environ.* 739: 139076. <https://doi.org/10.1016/j.scitotenv.2020.139076>

<sup>26</sup> Zhang Y, Cen M, Hu M, Du L, Hu W, Kim JJ, Dai N. (2021). Prevalence and Persistent Shedding of Fecal SARS-CoV-2 RNA in Patients With COVID-19 Infection: A Systematic Review and Meta-analysis. *Clin Transl Gastroenterol.* 12 (4): e00343. <https://doi.org/10.14309/ctg.0000000000000343>

<sup>27</sup> Zheng S, Fan J, Yu F, Feng B, Lou B, Zou Q, Xie G, Lin S, Wang R, Yang X, Chen W, Wang Q, Zhang D, Liu Y, Gong R, Ma Z, Lu S, Xiao Y, Gu Y, Zhang J, Yao H, Xu K, Lu X, Wei G, Zhou J, Fang Q, Cai H, Qiu Y, Sheng J, Chen Y, Liang T. (2020). Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. *BMJ.* 369: m1443. <https://doi.org/10.1136/bmj.m1443>

<sup>28</sup> Fongaro G, Stoco PH, Souza DSM, Grisard EC, Magri ME, Rogovski P, Schörner MA, Barazzetti FH, Christoff AP, de Oliveira LFV, Bazzo ML, Wagner G, Hernández M, Rodríguez-Lázaro D. (2021). The presence of SARS-CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. *Sci Total Environ.* 778: 146198. <https://doi.org/10.1016/j.scitotenv.2021.146198>

<sup>29</sup> La Rosa G, Mancini P, Bonanno Ferraro G, Veneri C, Iaconelli M, Bonadonna L, Lucentini L, Suffredini E. (2021). SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Sci Total Environ.* 750: 141711. <https://doi.org/10.1016/j.scitotenv.2020.141711>





Esto sería así siempre que el resto de factores que nos podrían influir en los niveles no varíen (hora y lugar de muestreo, precipitaciones, etc)<sup>32,33,34</sup>.

### 3) Limitaciones para relacionar los datos epidemiológicos con las dianas del virus

Debido al desarrollo actual de la técnica y a las incertidumbres que todavía rodean al SARS-CoV-2, por ser un virus nuevo, debemos tener en cuenta algunas limitaciones a la hora de buscar relaciones entre los niveles detectados en las aguas residuales y los casos clínicos de COVID-19 diagnosticados. Asimismo, hay que buscar estas relaciones con cautela.

Debemos tener en cuenta que no todos los pacientes COVID-19 positivos excretan ARN del SARS-CoV-2 en las heces y, cuando ocurre, la cantidad de virus y la duración de la eliminación varían entre los individuos y a lo largo del tiempo.

Como se va a comentar más adelante, otras variables ambientales (por ejemplo, eventos de lluvia, temperatura, tiempo de retención hidráulica en alcantarillas) aumentan las incertidumbres vinculadas a la correlación entre la detección del ARN del SARS-CoV-2 en muestras de aguas residuales.

Además, las características de los sistemas de notificación de casos clínicos, así como los indicadores escogidos para reportar la infección entre la población también aportan incertidumbres a esta relación. Algunas limitaciones derivadas de estos sistemas de notificación a la hora de relacionar los datos son:

- Los decalajes en el reporte de los datos.
- Los distintos criterios de cada comunidad autónoma para asignar el día de notificación (día de realización de la prueba o día de obtención de los resultados).
- Las pruebas diagnósticas utilizadas, la capacidad diagnóstica de cada comunidad y su variación desde el inicio de la epidemia.
- Los fines de semana donde no se reporta.
- La consolidación de los datos históricos.
- La población que vierte a la red o la depuradora. En muchas ocasiones no coincide con la totalidad de un municipio, ni de las zonas básicas de salud. Por ello habría que utilizar los casos notificados relativos a la zona exacta cuyas aguas acaban en la EDAR en cuestión, para una mejor aproximación.

Respecto a la utilización de los datos de las aguas residuales, hay que tener en cuenta las escalas en las que se están expresando (lineal o logarítmica), las variaciones en días puntuales debidos a factores externos como precipitaciones o caudal, las características de la red y el punto de muestreo, la posible degradación del SARS-CoV-2 a lo largo del recorrido y el decalaje respecto a los casos diagnosticados.

---

<sup>32</sup> Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. (2020). SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Res.* 181: 115942. <https://doi.org/10.1016/j.watres.2020.115942>

<sup>33</sup> Medema G, Been F, Heijnen L, Petterson S. (2020). Implementation of environmental surveillance for SARS-CoV-2 virus to support public health decisions: Opportunities and challenges. *Curr Opin Environ Sci Health.* 17: 49-71. <https://doi.org/10.1016/j.coesh.2020.09.006>

<sup>34</sup> Wu F, Zhang J, Xiao A, Gu X, Lee WL, Armas F, Kauffman K, Hanage W, Matus M, Ghaeli N, Endo N, Duvallet C, Poyet M, Moniz K, Washburne AD, Erickson TB, Chai PR, Thompson J, Alm EJ. (2020) SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *mSystems.* 5 (4): e00614-20. <https://doi.org/10.1128/mSystems.00614-20>

Debemos tener en cuenta que en los niveles detectados se estarían identificando también los casos asintomáticos, pre-sintomáticos y los casos mal diagnosticados a nivel del sistema sanitario, por lo que los niveles podrían ser más elevados que sus correspondientes desde el punto de vista clínico.

En epidemiología, se denomina **prevalencia** a la proporción de individuos de un grupo o una población (en medicina, persona), que presentan una característica o evento determinado (en medicina, enfermedades). Por lo general, se expresa como una fracción, un porcentaje o un número de casos por cada 10 000 o 100 000 personas.

La prevalencia de una enfermedad se calcula dividiendo el número total de los individuos que presentan un atributo o enfermedad en un momento o durante un periodo (evento), entre la población en ese punto en el tiempo o en la mitad del periodo (número de individuos). Cuantifica la proporción de personas en una población que tienen una enfermedad (o cualquier otro suceso) en un determinado momento y proporciona una estimación de la proporción de sujetos de esa población que tenga la enfermedad en ese momento.

$$\text{Prevalencia} = \frac{\text{Número existente de casos}}{\text{Población total en riesgo en un punto en el tiempo}}$$

Podemos distinguir dos tipos de prevalencia: **puntual y de periodo**.

- **Prevalencia puntual:** cuántas personas de un grupo definido están enfermas en un determinado momento. Ejemplo hipotético: 1% de los empleados están enfermos esta semana.
- **Prevalencia de periodo:** la proporción de personas que están o estarán enfermas en algún momento. Ejemplo hipotético: 10% de los habitantes de este pueblo tendrá un resfriado en algún momento durante su vida.

Es un parámetro útil porque permite describir un fenómeno de salud, identificar la frecuencia poblacional del mismo y generar hipótesis que lo expliquen. La utilizan normalmente los epidemiólogos, las personas encargadas de la política sanitaria, las agencias de seguros y en diferentes ámbitos de la salud pública.

La **incidencia** es el número de casos nuevos de una enfermedad en una población determinada y en un periodo determinado.

$$\text{Incidencia} = \frac{\text{Número de casos nuevos}}{\text{Población en riesgo en un punto en el tiempo}}$$

La prevalencia describe la proporción de la población que padece la enfermedad, que queremos estudiar, en un momento determinado, es decir es como una foto fija. En cambio, la incidencia va a contabilizar el número de casos nuevos, de la enfermedad que estudiamos, que aparecen en un período de tiempo previamente determinado; podemos equipararla a una película que refleja el flujo del estado de salud al de enfermedad en la población que estudiamos.

La **incidencia acumulada (IA)** se define como la proporción de individuos sanos que desarrollan la enfermedad a lo largo de un periodo determinado. Una proporción es el cociente de dos frecuencias absolutas en el que el numerador está incluido en el denominador. La incidencia acumulada proporciona una estimación de la probabilidad o el riesgo de que un individuo libre de una determinada enfermedad la desarrolle durante un período especificado de tiempo. Como cualquier proporción, suele venir dada en términos de porcentaje. No es una tasa porque el denominador no incluye unidad de tiempo, aunque es

preciso indicar cuál ha sido el periodo en el que se han producido esos nuevos casos de enfermedad, es decir, el periodo de observación, para poder ser interpretada.

La **tasa de incidencia (TI)** o densidad de incidencia es la relación entre el número de nuevos casos a lo largo de un periodo concreto y la suma de los períodos de riesgo de cada uno de los individuos a lo largo del período que se especifica. Es una tasa porque el denominador incluye unidad de tiempo. Las unidades en que se mide esta tasa son, por tanto, casos de enfermedad por cada persona-año.

Algunos indicadores que pueden utilizarse para buscar esta relación entre los niveles detectados en aguas residuales y los casos clínicos notificados se describen a continuación.

**Casos diarios notificados:** Número de casos clínicos diagnosticados cada día de COVID-19.

Es importante tomar como referencia la población que vierte a la EDAR, ya que en ocasiones no todo el municipio o zona básica de salud comparten la misma depuradora o red de saneamiento.

**Casos notificados acumulados:** Suma de los casos notificados durante un periodo de tiempo determinado. Se suele emplear de 7 o 14 días.

Como todos los indicadores que relacionan los casos de varios días, nos permiten obtener una foto más completa ya que contaríamos con los nuevos casos y los casos positivos detectados previamente que todavía excretan virus. Si bien todavía se desconoce si las cantidades de virus excretadas varían significativamente a lo largo de los días.

Hay que diferenciar los casos acumulados de los **casos activos**. Estos últimos hacen referencia a los casos acumulados sin contar aquellos que ya no estarían pasando la infección. Estos podrían ser las altas (con prueba diagnóstica negativa) o los fallecidos, y por tanto que ya no estarían pasando la infección.

**Media móvil de los casos diarios notificados:** Se hace el promedio de casos diarios durante los días anteriores. Se suele emplear la media móvil a 7 días. Esto se hace para evitar las alteraciones de días puntuales "anómalos" y suavizar la curva.

**Incidencia acumulada:** El número de casos nuevos de una enfermedad en una población determinada y en un periodo determinado, en relación al total de personas susceptibles de contraer la enfermedad. En el contexto del COVID-19, se suele utilizar la incidencia a 7 o 14 días, reportándose como casos diagnosticados por cada 100 000 habitantes.

Actualmente, todavía se desconoce qué indicador epidemiológico puede ser más adecuado para buscar estas relaciones. Se están realizando investigaciones en todo el mundo para esclarecer las incertidumbres que rodean estos aspectos.



### III. Factores que influyen en los niveles detectados en aguas residuales

Existen numerosos factores que pueden influir en los niveles de SARS-CoV-2 detectados en las aguas residuales. Algunos de ellos son:

#### 1) Factores ambientales

- Temperatura del agua residual: Los valores de temperatura del agua residual son en general en origen constantes, pero se ven influenciados por la variación de la temperatura ambiente dependiendo de la zona y de la época del año en la que se realice la medición. A mayor temperatura, mayor degradación del material genético del virus. Además, los valores de temperatura afectan al proceso biológico de depuración. Oscilan en un rango de entre 15 y 30 °C en casos ambientales extremos.
- Precipitaciones: Se trata de un factor de gran importancia ya que, ante un evento de lluvias, la escorrentía generada en la cuenca termina parcial o totalmente introduciéndose en la red de saneamiento, lo que hará que aumente el caudal circulante por la red y haya mayor dilución. Cuando se trata de redes de tipo unitario el aumento de la dilución conlleva una disminución en la concentración del material genético del virus a detectar. Además, dada la heterogeneidad de las lluvias en cuencas grandes es complicado determinar con cierta precisión la relación entre lluvia y caudal de llegada a la EDAR o entre lluvia y dilución producida en la red.

#### 2) Características de la cuenca tributaria de la EDAR y de la red de saneamiento

- Tamaño de la cuenca: La cuenca de cada EDAR, entendida como el área o extensión donde se generan las aguas residuales que son tratadas en una depuradora, tiene unas características únicas y diferentes de cualquier otra cuenca. A mayor tamaño de cuenca y en consecuencia mayor recorrido de la red de saneamiento, hay mayor degradación potencial del material genético del virus. Una solución a partir del conocimiento profundo de la red sería la sectorización de la misma y la identificación de colectores principales para muestrear en puntos más próximos a los focos de emisión (viviendas, hospitales, etc.)
- Red de saneamiento. Al igual que la cuenca tributaria de la EDAR, cada red de saneamiento presenta sus particularidades en cuanto a longitud, capacidad, etc. De acuerdo con el diseño de las redes de saneamiento, estas pueden asumir generalmente sin problema las aguas residuales generadas en tiempo seco. Sin embargo, ante eventos de lluvia, dependiendo de la capacidad de la red pueden producirse descargas al medio ambiente mediante o derivaciones a otras instalaciones que complican la cuantificación del caudal que llega a las EDAR y de la dilución real. Asimismo, se debería tener en cuenta la presencia de bombes, que retienen y laminan el caudal de aguas residuales circulante por la red, modificando los tiempos de recorrido en la misma y afectando a la degradación del material genético del virus.

#### 3) Variaciones en las características estándar del agua residual

- Infiltraciones en la red: La entrada en la red de saneamiento de agua presente en el terreno, debido al mal estado de los colectores y niveles freáticos elevados, aumentaría el caudal circulante en la red y se produciría mayor dilución. Esta situación podría identificarse con aumentos de caudal no

asociados a eventos de lluvia, pero exigiría un conocimiento exhaustivo de los caudales generados en tiempo seco.

- Intrusión salina: Fenómeno que se produce en zonas costeras cuando una masa de agua salada avanza en el terreno. De producirse infiltraciones de agua en salada en la red, aparte de reducir la capacidad de la misma, y afectar a la degradación de los materiales se produciría un aumento en la conductividad eléctrica. Valores elevados de este parámetro podrían influir en la degradación del material genético del virus y, también, afectaría al proceso biológico pudiendo generar problemas en los procesos de depuración.
- Los vertidos procedentes de industrias u otras instalaciones singulares podrían alterar las características fisicoquímicas de las aguas residuales que llegan a la EDAR y consecuentemente degradar el material genético del virus que se pretende detectar. Para detectar la posible influencia de estos vertidos, se puede realizar el control y seguimiento de pH y conductividad eléctrica mediante medición in situ en el momento de la toma de muestras. Además, convendría tener en cuenta el papel de la vigilancia de las aguas residuales para la detección del SARS-CoV-2 de origen animal (por ejemplo, sitios de producción animal, mercados húmedos...) y la consideración de otros métodos moleculares óptimos de detección (como por ejemplo la secuenciación de próxima generación) de cara a la detección de posibles mutaciones en la fuente animal para evitar su propagación a la comunidad.

#### 4) Muestreo en EDAR

- Tipo de muestra: La elección del tipo de muestra va a tener influencia en la detección y cuantificación del material genético del virus.
  - o Compuesta. Muestra integrada de varias muestras de un periodo de tiempo, normalmente 24 h y proporcional al caudal influente en el momento de toma de cada una de las muestras. El principal inconveniente es la potencial dilución de la muestra.
  - o Simple. Elección del momento de mayor carga fecal, atendiendo a indicaciones del explotador de la EDAR y/o mediante análisis de *Escherichia coli*, u otros microorganismos comensales o patógenos presentes en las aguas residuales y que estén reconocidos como indicadores de materia fecal en las aguas residuales.

Es el tipo de muestra seleccionado en el proyecto VATar-COVID-19 y su principal inconveniente es la incertidumbre en cuanto a la elección del momento óptimo de toma de muestra, así como a las variaciones que pudieran producirse. Se han realizado ejercicios de correlación entre *E. coli* y SARS-CoV-2 a diferentes horas.

- Punto de toma de la muestra: Posible influencia de otros procesos de la EDAR, tales como recirculación de fangos, retorno de agua de deshidratación de fangos, etc.
- Contaminación de las muestras debido a desviaciones en el protocolo de toma de muestra o errores humanos.

A la hora de conocer la evolución de la epidemia en una zona, para evitar la distorsión de la información obtenida del monitoreo de las aguas residuales, se recomienda mantener estable el mayor número de variables posibles. Comparar los niveles dentro del mismo punto de muestreo, utilizando las mismas

técnicas tanto de muestreo como analíticas, y hacerlo el mismo día y hora puede ayudar a reducir la influencia de estos factores en los niveles detectados.

### 5) Análisis del momento óptimo de muestreo

Debido al tipo de muestra simple que se recoge, el momento de muestreo es fundamental. Por ello es recomendable realizar un análisis para conocer el momento óptimo de realizarlo. Dentro del proyecto VATar-COVID-19, se estableció que este momento coincidiese con el de mayor carga fecal a la entrada a la EDAR durante la mañana, por ser representativo y, en teoría, no presentar variaciones significativas entre días distintos, aparte de entre días de diario/fin de semana.

Es necesario determinar la hora exacta de muestreo dentro de esa franja horaria, para garantizar que no haya intrusión de vertidos en ese momento, y que se realiza en el momento de mayor carga fecal, o cuando los niveles de SARS-CoV-2 son más elevados. Este último punto es importante ya que aumenta la sensibilidad del sistema de alerta temprana en aquellas fases de la epidemia donde haya ausencia de detección en las aguas residuales.

Para conocer dicho momento óptimo de muestreo de cada depuradora se han realizado múltiples estudios en distintas EDAR. Cabe señalar que la hora óptima es específica de cada EDAR, debido a las características de la misma, de su red de saneamiento y población conectada.

Estos estudios consisten en tomar muestras puntuales a distintas horas. De cada muestra se obtienen y comparan los niveles de distintos parámetros físico-químicos y microbiológicos. Dentro de los parámetros físico-químicos se analizan, mediante una sonda multiparamétrica en el punto de muestreo, los siguientes:

- pH
- Temperatura del agua
- Conductividad eléctrica
- Oxígeno disuelto
- % saturación O<sub>2</sub>

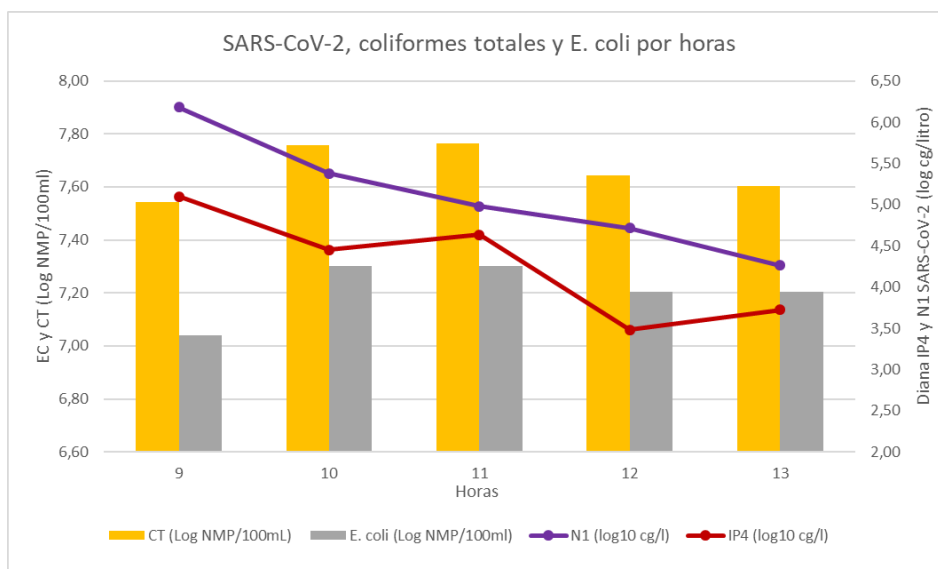
Asimismo, también se compara el caudal registrado en la muestra tomada en cada hora de estudio.

Dentro de los parámetros microbiológicos se realizan determinaciones para:

- *E. coli*
- Coliformes totales
- SARS-CoV-2, usando las dianas utilizadas en el proyecto (IP4, E y N1)

Con este análisis se intenta acotar el momento de máxima carga fecal a la entrada de la EDAR, así como el momento óptimo de muestreo. Aunque más investigaciones y evidencias son necesarias, la hora de máxima carga fecal no siempre coincide exactamente con la de mayores niveles de SARS-CoV-2.

Como ejemplo de una de las EDAR analizadas, a continuación, se exponen los resultados microbiológicos del estudio realizado en Torrejón de Ardoz:



**Figura 6.** Resultados microbiológicos del análisis de la hora óptima de muestreo en la EDAR de Torrejón de Ardoz

En base a la información que se tenía al comienzo del proyecto se habían establecido las 12:00 h como momento de muestreo. Como se puede apreciar, se detectó una diferencia significativa en los niveles de SARS-CoV-2 entre la muestra de las 12 h y la de las 9 h. Por ello, se determinó modificar el momento de muestreo a dicha hora a pesar de que los indicadores fecales (*E. coli* y coliformes) indicaban una mayor carga fecal a las 10 h. Los valores físico-químicos y de caudal no presentaron modificaciones significativas en las horas analizadas.

## 6) Logística

- Retrasos en las entregas al laboratorio. Generan problemas de mantenimiento de temperatura de conservación de las muestras. Si la temperatura de las muestras no se mantiene por debajo de la temperatura especificada en el protocolo de muestreo, se produciría degradación del material genético del virus

## 7) Variaciones en la excreción del virus por parte de las personas infectadas

Como se ha comentado con anterioridad, en base al conocimiento disponible hasta el momento, se desconocen diferentes factores relativos a la interacción del virus con su hospedador. Algunos de estos factores afectan a los niveles detectados en las aguas residuales. Entre ellos, se desconoce el porcentaje de personas infectadas que excretan el virus en sus heces, pues no todas lo hacen. Asimismo, se desconocen los patrones de excreción o la cantidad excretada a lo largo del proceso infeccioso. A esto hay que sumarle posibles variaciones en estos factores entre los genotipos que están circulando en la comunidad.<sup>35</sup>

A nivel social también existen factores que pueden influir en los niveles detectados, sobre todo cuando la metodología de muestreo incluye la toma de muestras puntuales. Los días festivos, la movilidad, o los confinamientos domiciliarios de la población, pueden afectar a los patrones conductuales de la comunidad que se está monitoreando, influyendo en los niveles detectados en las aguas residuales y en su uso.

<sup>35</sup> van Doorn AS, Meijer B, Frampton CMA, Barclay ML, de Boer NKH. (2020). Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Aliment Pharmacol Ther.* 52 (8): 1276-1288. <https://doi.org/10.1111/apt.16036>



## 8) Degradación del material genético del virus

Desde que el virus es excretado por la persona infectada hasta que se recoge la muestra, hay multitud de factores que le afectan y que pueden acelerar o disminuir la degradación del mismo. Esto tiene una influencia directa en los niveles que se detectan a la entrada de la EDAR o en los colectores.<sup>36</sup>

Un factor fundamental es el tiempo que tarda desde su excreción a la detección, ya que determina el mayor o menor impacto del resto de factores. Como se ha comentado con anterioridad, un mayor o menor tiempo de viaje está producido por las características de la cuenca tributaria y de la red de saneamiento.

Otros factores que afectan a la degradación del material genético, y que hay que tener en cuenta son:

- Los pH extremos. Aunque los rangos usuales de las aguas residuales no parecen tener un gran impacto en las determinaciones.<sup>37</sup>
- La temperatura. Como ya se ha comentado, es un factor fundamental. Diversos estudios muestran como la estabilidad del ARN vírico varía en gran medida según la temperatura. Mientras que a 4 °C el virus se puede detectar pasados varios días, este tiempo se reduce a temperatura ambiente.<sup>38</sup>
- La composición de las aguas residuales. Al ser una matriz compleja, compuesta por multitud de sustancias siendo muchas de origen biológico, la interacción de estas con el virus acelera la degradación del material genético del mismo.<sup>39</sup>
- Sólidos en suspensión. Aunque investigaciones adicionales son necesarias sobre este punto, diversos estudios apuntan a la posibilidad de que el SARS-CoV-2 tenga gran afinidad por los sólidos en suspensión de las aguas residuales, de tal manera que los niveles detectados pueden variar en función de la mayor o menor presencia de estos.<sup>36</sup>
- Vertidos: Como ya se ha comentado, determinadas sustancias que se pueden verter desde las industrias o los hogares pueden favorecer la degradación del material genético del virus.

---

<sup>36</sup> Mandal P, Gupta AK, Dubey BK. (2020). A review on presence, survival, disinfection/removal methods of coronavirus in wastewater and progress of wastewater-based epidemiology. *J Environ Chem Eng.* 8 (5): 104317. <https://doi.org/10.1016/j.jece.2020.104317>

<sup>37</sup> Chan KH, Sridhar S, Zhang RR, Chu H, Fung AY, Chan G, Chan JF, To KK, Hung IF, Cheng VC, Yuen KY. (2020). Factors affecting stability and infectivity of SARS-CoV-2. *J Hosp Infect.* 106 (2): 226-231. <https://doi.org/10.1016/j.jhin.2020.07.009>

<sup>38</sup> Hart OE, Halden RU. (2020). Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Sci Total Environ.* 730: 138875. <https://doi.org/10.1016/j.scitotenv.2020.138875>

<sup>39</sup> Amoah ID, Kumari S, Bux F. (2020). Coronaviruses in wastewater processes: Source, fate and potential risks. *Environ Int.* 143: 105962. <https://doi.org/10.1016/j.envint.2020.105962>

## 9) Factores relativos a los métodos analíticos

Dentro de este apartado, debido a que la técnica se está perfeccionando paralelamente al desarrollo del proyecto, pueden producirse cambios en el protocolo a lo largo del proyecto que hagan que los resultados de distintas fechas tengan que compararse con cautela.

Asimismo, se pueden producir incidencias puntuales relativas a la técnica analítica que afecten a los niveles detectados en las aguas residuales. Para solventar este tipo de situaciones, se realizan distintos controles a lo largo del proceso.

En la fase de concentración y de extracción, se realizan controles añadiendo una cantidad conocida de otros virus a la muestra, que sirven como virus control de proceso. Estos se cuantifican al finalizar la determinación, para comparar el resultado con la cantidad que se introdujo y así conocer la eficiencia del proceso y garantizar la fiabilidad de los resultados. Asimismo, durante la cuantificación se realizan múltiples controles para garantizar la calidad de las determinaciones.

Para reducir incertidumbres se realizan varias réplicas de cada muestra, aplicando distintas diluciones, como se detalla en los protocolos de los laboratorios.

Por último, para garantizar que la calidad de los resultados aportados por todos los laboratorios integrados dentro del proyecto VATar COVID-19, se realizó un ejercicio de intercomparación a partir de la misma muestra, obteniendo resultados satisfactorios para todos los laboratorios.

Actualmente se están llevando a cabo diferentes estudios en el campo de la inferencia estadística con el fin de establecer correlaciones ajustadas a las titulaciones del virus detectadas a nivel de EDAR. En concreto, la agencia estadounidense del medio ambiente (USEPA, *The United States Environmental Protection Agency*) está llevando a cabo modelos matemáticos sobre redes Bayesianas para establecer una correlación de los niveles detectados de SARS-CoV-2 y los datos epidemiológicos disponibles a nivel poblacional. Además, se han llevado a cabo estudios matemáticos sobre "*Regression and Susceptible-Exposed-Infected-Removed modelling*":

- *Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study (TWu J., Leung K., and Leung G.M.)*
- *Early dynamics of transmission and control of COVID19: a mathematical modelling study (Kucharski A.J et al)*
- *Projecting social contact matrices in 152 countries using contact surveys and demographic data (Prem K., Cook A.R. et al.)*
- *Bayesian inference modelling: Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2) (Li R., Pei S. et al)*

Otros factores a tener en cuenta son los siguientes: la persistencia de fragmentos de ARN del SARS-CoV-2 en las aguas residuales, el establecimiento de un modelo de cultivo celular para evaluar la infectividad del virus a partir de muestras ambientales, así como el potencial de supervivencia de virus infecciosos en aguas residuales tratadas y no tratadas, entre otros.

Todos estos factores que se han comentado hacen que, en el momento actual de desarrollo de la técnica, la comparación de los niveles detectados entre distintas EDAR o redes de saneamiento no sea recomendable. Esto se debe a la influencia de factores determinantes como la longitud de la red, la población que vierte, factores ambientales puntuales o los laboratorios y técnicas implicadas en la determinación. El hecho de que se detecten niveles más elevados en una localización u otra, no implica necesariamente que haya más personas infectadas en esa población. Esto no implica que con el desarrollo de modelos matemáticos esto sea posible en un futuro.

