

Acceso al agua y COVID-19: un estudio de regresión discontinua para áreas periurbanas de Lima Metropolitana, Perú

Andrés Gómez-Lobo
Mauro Gutiérrez
Sandro Huamaní
Diego Marino
Tomás Serebrisky
Ben Solís

Acceso al agua y COVID-19: un estudio de regresión discontinua para áreas periurbanas de Lima Metropolitana, Perú

Andrés Gómez-Lobo^{*}
Mauro Gutiérrez^{**}
Sandro Huamaní^{**}
Diego Marino^{**}
Tomás Serebrisky^{***}
Ben Solís^{***}

^{*} Departamento de Economía de la Universidad de Chile

^{**} Superintendencia Nacional de Servicios de Saneamiento (Sunass)

^{***} Sector de Infraestructura y Energía del BID

Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del

Banco Interamericano de Desarrollo

Acceso al agua y COVID-19: un estudio de regresión discontinua para áreas
periurbanas de Lima Metropolitana, Perú / Andrés Gómez-Lobo, Mauro Gutiérrez,
Sandro Huamaní, Diego Marino, Tomás Serebrisky, Ben Solís.

p. cm. — (Documento de trabajo del BID ; 1351)

Incluye referencias bibliográficas.

1. Water-supply-Peru-Regression analysis. 2. Infrastructure (Economics)-Peru-
Regression analysis. 3. Water utilities-Peru-Regression analysis. 4. Coronavirus
infections-Social aspects-Peru-Regression analysis. I. Gómez-Lobo, Andrés. II.
Gutierrez, Mauro. III. Huamaní, Sandro. IV. Marino, Diego. V. Serebrisky, Tomás. VI.
Solís, Ben. VII. Banco Interamericano de Desarrollo. Sector de Infraestructura y
Energía. VIII. Serie.
IDB-WP-1351

<http://www.iadb.org>

Copyright © 2022 Banco Interamericano de Desarrollo. Esta obra se encuentra sujeta a una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas (CC-IGO 3.0 BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando el reconocimiento respectivo al BID. No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI (UNCITRAL). El uso del nombre del BID para cualquier fin distinto al reconocimiento respectivo y el uso del logotipo del BID, no están autorizados por esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Después de un proceso de revisión por pares, y con el consentimiento previo y por escrito del BID, una versión revisada de esta obra podrá reproducirse en cualquier revista académica, incluyendo aquellas referenciadas por la Asociación Americana de Economía a través de EconLit, siempre y cuando se otorgue el reconocimiento respectivo al BID, y el autor o autores no obtengan ingresos de la publicación. Por lo tanto, la restricción a obtener ingresos de dicha publicación sólo se extenderá al autor o autores de la publicación. Con respecto a dicha restricción, en caso de cualquier inconsistencia entre la licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas y estas declaraciones, prevalecerán estas últimas.

Note que el enlace URL incluye términos y condiciones adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.



Dirección de correspondencia: bsolissosa@iadb.org

Resumen

Este trabajo presenta los resultados de un estudio cuasi-experimental para el cual se recolectó información mediante una encuesta realizada en zonas periurbanas de Lima Metropolitana entre octubre y noviembre de 2021. La encuesta se realizó a hogares que residen cerca y a ambos lados de la frontera de la cobertura de la red pública de agua. Nuestro trabajo encuentra que el acceso al agua mediante red se asoció a una reducción en la probabilidad de infección por COVID-19. Asimismo, una extensión al modelo, utilizando efectos heterogéneos, sugiere que no basta con que un hogar esté conectado a la red, sino que además debe garantizarse una dotación mínima de consumo. Los resultados deben interpretarse tomando en consideración las limitaciones en la información. Estos resultados ponen de relevancia la necesidad de la inversión en infraestructura para el cierre de brechas de acceso, y la importancia de asegurar servicios de calidad y asequibles para la población.

Abstract

This paper presents the results of a quasi-experimental study using information collected through a survey conducted in peri-urban areas of Metropolitan Lima between October and November 2021. The survey was applied to households residing close to and on both sides of the geographic boundary of piped water supply. Our work finds that access to piped water was associated with a reduction in the probability of a COVID-19 infection. Furthermore, the model shows heterogeneous effects that suggest that it is not enough for a household to be connected to the water supply network, but that a minimum consumption endowment must also be guaranteed. The results should be interpreted by taking into consideration the limitations of the information. These results highlight the need for investment in infrastructure to close access gaps and the importance of ensuring quality and affordable services for the population.

Códigos JEL: L95, I14, I15, I10, I18

Palabras clave: agua, saneamiento, COVID-19, salud, regresión discontinua, Lima, Perú.

Tabla de contenido

1.	INTRODUCCIÓN	3
2.	REVISIÓN DE LA LITERATURA	5
3.	CONTEXTO.....	16
3.1.	COVID-19: TRANSMISIÓN, CONDICIONES DE RIESGO Y PREVENCIÓN	16
3.2.	EL COVID-19 EN PERÚ.....	17
3.3.	SITUACIÓN ACTUAL DEL ACCESO AL AGUA EN PERÚ	21
4.	CARACTERÍSTICAS DE LOS DATOS.....	25
5.	METODOLOGÍA.....	27
6.	RESULTADOS	31
6.1.	MODELO LOGIT RD: EL ROL DEL ACCESO AL AGUA.....	31
6.2.	MODELO LOGIT CON EFECTOS HETEROGÉNEOS POR NIVEL DE CONSUMO: MÁS ALLÁ DEL ACCESO, ENTENDIENDO LA IMPORTANCIA DE UNA DOTACIÓN MÍNIMA	35
6.3.	TEST DE FALSIFICACIÓN.....	36
7.	CONSIDERACIONES PREVIAS A LA DISCUSIÓN DE RESULTADOS	39
8.	DISCUSIÓN DE RESULTADOS.....	42
9.	RECOMENDACIONES DE POLÍTICA	43
	REFERENCIAS	44
	ANEXO 1 – FICHA TÉCNICA DE LA ENCUESTA.....	49
	ANEXO 2 – TEST DE MEDIAS PARA LA PROPORCIÓN DE INDIVIDUOS CON PRUEBA COVID-19	51
	ANEXO 3 – RESULTADOS DEL TEST DE MEDIAS (TESTEADOS VS NO TESTEADOS).....	52
	ANEXO 4 – MODELO CON CORRECCIÓN POR POSIBLE SESGO DE SELECCIÓN	53
	ANEXO 5 – PROBABILIDAD DE INFECCIÓN POR COVID-19 SEGÚN DISTANCIA A LA FRONTERA	55

Acceso al agua y COVID-19: un estudio de regresión discontinua para áreas periurbanas de Lima Metropolitana - Perú¹

1. Introducción

Dentro de los múltiples retos que enfrenta la región de América Latina y el Caribe se encuentra el cierre de la brecha de infraestructura de servicios públicos que permita mejorar significativamente la calidad de vida de sus habitantes. Resulta necesario que los países realicen un esfuerzo adicional de inversión pública y privada para garantizar el acceso a los servicios básicos. La brecha de inversión en agua, energía, telecomunicaciones y transporte hasta 2030 se ha estimado en 2,2 billones de dólares (Brichetti *et al.*, 2021), equivalente a 43% del PBI de 2019 de la región. En particular, para el caso de los servicios de agua y saneamiento, se estima que se requieren 374 mil millones de dólares para brindar acceso universal a agua potable y saneamiento gestionados de forma segura y garantizar el tratamiento de las aguas residuales en áreas urbanas. Ello permitiría construir los principales componentes de infraestructura para el cumplimiento del Objetivo de Desarrollo Sostenible N° 6 pero requeriría, por parte de la región, un esfuerzo de inversión promedio anual de 0,5% del PBI hasta 2030.

En el caso de Perú se estima que, a 2020, el 91,2% de la población contaba con acceso a agua a través de red pública, existiendo una marcada diferencia entre las áreas urbanas (94,8%) y rurales (77,6%). En Lima Metropolitana², que concentra alrededor de un tercio de la población nacional, la cobertura alcanza el 95,6% (INEI, 2021) y las zonas sin acceso se encuentran en las áreas periurbanas de la ciudad.

La falta de acceso a servicios de agua y saneamiento cobró mayor relevancia en el contexto de la pandemia por la enfermedad por coronavirus 2019 (en adelante, COVID-19), debido a que una de las primeras recomendaciones de la Organización Mundial de la Salud (OMS) fue el frecuente lavado de manos como estrategia contra el virus (OMS, 2020). Sin embargo, considerando que la recomendación requiere lavarse las manos con jabón, agua corriente limpia y durante al menos 20 segundos, esta resultó de difícil implementación para aquellos hogares que

¹ Los autores agradecen al Dr. Mateo Prochazka, MD MSc, por sus valiosos aportes a una versión preliminar del documento, a Guillermo Cruces y Clara Pasmán por sus comentarios a las últimas versiones del documento y a Beremiz Rojas por su apoyo en la sistematización de información estadística.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo, ni de los países que representa. Los errores u omisiones corren por exclusiva cuenta de los autores. Los interesados pueden remitir comentarios a la siguiente dirección de correo electrónico: bsolissosa@iadb.org

² A lo largo del presente estudio, siguiendo la nomenclatura del Instituto Nacional de Estadística e Informática de Perú, utilizaremos la denominación Lima Metropolitana para incluir la región conformada por la provincia de Lima y la provincia constitucional del Callao (Encuesta Nacional de Hogares 2021).

no están conectados a la red de los prestadores del servicio de agua potable. Asimismo, como se explica más adelante en este documento, los medios alternativos para el abastecimiento de los usuarios no conectados —principalmente camiones cisterna— generan interacciones sociales con miembros de otros hogares (como formar largas colas o compartir instalaciones) que pueden incrementar la probabilidad de infección por COVID-19 en esta población, por lo cual resulta de interés analizar si la población sin acceso al agua se ha visto significativamente más expuesta a la enfermedad.

Perú es un caso de estudio de relevancia debido a que, a febrero de 2022, es el país que registraba el mayor número de muertes oficiales a causa del COVID-19 por mil habitantes a nivel mundial (*The Economist*, 2022). En particular, la región de Lima Metropolitana (que incluye la capital de Perú) ha sido una de las más afectadas tanto a nivel de casos como de muertes. La alta mortalidad por COVID-19 observada en el Perú es parcialmente atribuible a factores estructurales (Gianella *et al.*, 2020; Schwalb y Seas, 2020). Entre estos destacan factores relacionados al sistema de salud (carente de infraestructura y personal especializado, además de sobrecargado y sin capacidad adecuada para pruebas moleculares) así como altos niveles de pobreza y bajo acceso a servicios básicos como agua y saneamiento. El acceso a estos servicios es de particular importancia ante una infección altamente transmisible que requiere de hábitos de higiene como el lavado de manos para prevenir su transmisión.

Este trabajo contribuye a la creciente literatura que busca explicar las diferencias en el impacto del COVID-19 en la salud de distintas poblaciones. A la fecha de redacción de este trabajo no conocemos estudios que hayan aproximado, con el individuo como unidad de análisis, la relación causal entre el acceso y calidad del servicio de agua y la probabilidad de infección por COVID-19. Para el presente estudio se ha llevado a cabo una encuesta en zonas periurbanas de Lima Metropolitana. Los hogares encuestados fueron aquellos que residían cerca de y a ambos lados de la frontera de cobertura de la red pública de agua potable en diferentes zonas periurbanas de la ciudad. Este diseño reduce las diferencias observables y no observables entre los hogares conectados y no conectados a la red pública, mejorando la posibilidad de inferir causalidad entre el acceso a la red pública y la probabilidad de infección por el virus. La encuesta recolectó información sobre acceso al agua y saneamiento, características del hogar y de la vivienda, hábitos de higiene y morbilidad por COVID-19, entre otras variables. Con dicha información, se realiza una evaluación bajo el método de Diseño de Regresión Discontinua (*Regression Discontinuity Design*) con el fin de establecer las diferencias existentes en morbilidad por COVID-19 que resulten atribuibles al acceso al agua de los hogares.

2. Revisión de la literatura

Existe abundante literatura que documenta la relación existente entre agua y saneamiento (acceso, calidad, higiene) y la salud humana. Waddington y Snilstveit (2009) realizaron una revisión sistemática de evaluaciones conducidas en las últimas tres décadas en 35 países de ingreso medio y bajo, estableciendo la relación entre agua, saneamiento, higiene y diarrea infantil. Otros trabajos en la misma línea son Darvesh *et al.* (2017) y Wolf *et al.* (2018). También existen revisiones sistemáticas de estudios donde se comprueba la relación entre calidad del agua y cáncer pulmonar (Celik *et al.*, 2008) y los impactos de intervenciones de agua, saneamiento e higiene en el cólera (Taylor *et al.*, 2015). Al respecto, Cárdenas (2022) realiza una revisión extensiva de la bibliografía escrita sobre la relación entre agua y salud.

Con respecto al COVID-19, existe un creciente número de investigaciones que resaltan que los factores sociales, económicos y demográficos pueden contribuir a disparidades en el impacto de esta enfermedad. La falta de infraestructura adecuada genera dinámicas sociales (aglomeraciones en transporte público, en puntos de abastecimiento de agua) que incrementan el riesgo de transmisión; mientras que otros factores socioeconómicos pueden limitar severamente la capacidad de los hogares para implementar las recomendaciones de prevención (por ejemplo, la falta de ingresos o trabajo formal puede dificultar el acatamiento de medidas de confinamiento). Asimismo, la desigualdad en el acceso a los sistemas de salud puede ser determinante para los pacientes que requieren oxígeno, hospitalización o una cama en Unidad de Cuidados Intensivos.

Las características de los documentos y publicaciones revisadas deben ser tomadas en cuenta en la interpretación de los resultados encontrados por los autores. La gran mayoría de los trabajos corresponden al análisis de los primeros meses de la pandemia. Por tanto, recogen los patrones iniciales de transmisión en un contexto en el que se conocía poco sobre el virus, no se contaba con vacunas disponibles y existía una o más variantes del virus que predominaban en ese momento³. Casi la totalidad de los trabajos revisados tienen las características de estudios epidemiológicos ecológicos; es decir, estudios en los que la unidad de estudio es la población (ya sea a nivel de países, estados, condados o distritos) y no el individuo. Por la naturaleza de los estudios, no se puede atribuir causalidad en las relaciones encontradas.

Para medir el impacto del COVID-19, los trabajos revisados utilizan principalmente indicadores de incidencia (número de casos confirmados), mortalidad (número de muertes) y

³ La vigilancia genómica varía de acuerdo con el país. En Perú, fueron variantes de preocupación las variantes Alfa, Lambda, Gamma (ver Vargas-Herrera *et al.*, 2022) y recientemente Ómicron.

la tasa de letalidad (ratio entre muertes y casos confirmados). Para efectos de la presente investigación, estos trabajos pueden dividirse en dos grupos. Los trabajos del primer grupo analizan la relación entre la afectación de la población por COVID-19 y las características socioeconómicas y demográficas de las localidades. Los trabajos del segundo grupo incluyen explícitamente la falta de acceso o calidad en los servicios de agua y saneamiento en el análisis de las disparidades geográficas en el impacto del COVID-19.

Dentro del primer grupo de trabajos, la mayoría de los autores reconoce que existen factores que hacen que sea más frecuente diagnosticar la enfermedad en cierta población o que esta evolucione desfavorablemente, tales como la edad, el sexo y la existencia de comorbilidades. Estos factores han sido incluidos en estudios en Brasil (Chauvin, 2021), China (You *et al.*, 2020); España (Marí-Dell’Olmo’ *et al.*, 2021), Estados Unidos (Ahmad *et al.*, 2020; Clouston, Natale y Link, 2021; Hyde, 2021; Kamis *et al.*, 2021; Rozenfeld *et al.*, 2020; Stokes *et al.*, 2021; Strully *et al.*, 2021; Xu *et al.*, 2021), Inglaterra (Nicodemo *et al.*, 2020) y en un análisis multi-país (Chauvin *et al.*, 2020). Un gran número de los estudios reseñados también incluyen variables de raza, etnicidad y educación⁴ (ver **Tabla 1**). A su vez, un número importante de estudios identificaron factores geográficos asociados con un mayor impacto de la enfermedad, dentro de los cuales se encuentran el acceso físico a la localidad y su conectividad con carreteras o aeropuertos. Zhang *et al.* (2021) encuentran que, a inicios de la pandemia, la cercanía a Wuhan estaba asociada con mayor infección por COVID-19 en China. Scarpone *et al.* (2020) encuentran que una mayor interconexión de la comunidad predice la incidencia de la enfermedad en Alemania. Fortaleza *et al.* (2020) realizan un análisis del tiempo de introducción del COVID-19 en São Paulo y encuentran que la introducción temprana de la enfermedad se observaba en distritos con mayor conectividad y que la distancia en carretera hasta la capital tuvo un efecto protector. En la misma línea, Chauvin (2021) encuentra que, en Brasil, la cercanía a un aeropuerto conectado internacionalmente exhibe correlación con el número de muertes *per cápita* en la primera ola⁵. Los resultados de estos trabajos son consistentes con el patrón de transmisión esperable en cualquier brote o epidemia. Con excepción del caso de China, en el resto de los países la transmisión se inició con la llegada de los primeros casos del extranjero. La transmisión se dio inicialmente desde personas infectadas hacia personas que tuvieron contacto directo con ellas (transmisión focalizada), lo cual explica la importancia de la conectividad y cercanía a aeropuertos.

⁴ Stokes *et al.* (2021) encuentran que el impacto desigual de la pandemia en población de ciertas razas y características socioeconómicas es exacerbado cuando se considera el exceso de muertes no asignadas al COVID-19. La interpretación de los resultados de etnicidad y raza podría encontrarse en factores culturales, comportamientos y actitudes frente a la pandemia.

⁵ Para Chauvin (2021), la primera ola llega a su punto máximo en julio 2020.

A medida que avanzó el proceso de transmisión comunitaria, otras variables de tipo demográfico como la densidad poblacional (habitantes por kilómetro cuadrado), la presencia de barrios precarios e informales, los patrones de movilidad y el uso de transporte público aumentaron el número y la cercanía de los contactos sociales, incrementando el riesgo de infección. Fortaleza *et al.* (2020) encuentran una relación positiva entre densidad poblacional de distritos de São Paulo y las variables de impacto de COVID-19 (introducción temprana, incidencia y mortalidad); mientras que Chauvin (2021), también para Brasil, encuentra una correlación positiva entre la densidad poblacional de las ciudades y el número de muertes *per cápita*. You *et al.* (2020) encuentran, en un análisis para los distritos de Wuhan, una relación positiva entre tasa de morbilidad y variables como densidad poblacional y proporción de áreas construidas⁶. Asimismo, Chauvin (2021) encuentra que las ciudades con un alto porcentaje de hogares ubicados en *favelas* sufrieron de un número desproporcionadamente mayor de muertes *per cápita*, pues en este tipo de barrios resulta difícil seguir las recomendaciones de distanciamiento social, en línea con lo afirmado en Brotherhood *et al.* (2020). Los patrones de movilidad también juegan un rol importante en la transmisión de la enfermedad. Chauvin (2021) encuentra que existió una mayor afectación por COVID-19 en las ciudades de mayor ingreso en Brasil⁷, las cuales se caracterizaron por una mayor movilidad, debido a que la actividad económica mantuvo su dinamismo en dichas zonas a pesar de la pandemia. Esto fue determinado a partir de un análisis complementario de una base de datos con información de movilidad de 60 millones de celulares. El estatus socioeconómico de los hogares o PBI de las localidades en algunos casos puede ser una variable *proxy* del nivel de actividad comercial y movilidad (Zhang *et al.*, 2021; Clouston *et al.*, 2021)⁸. You *et al.* (2020) encuentra una relación entre el impacto de la pandemia y ventas minoristas en China. Asimismo, el transporte público puede llegar a contribuir de manera importante a la transmisión, como fue el caso de Nueva York durante el despegue inicial de la pandemia (Harris, 2020). El uso o tiempo en el transporte público también son utilizados en estudios en Brasil (Chauvin, 2021; Freire de Souza *et al.*, 2020) y Estados Unidos (Dasgupta *et al.*, 2020; Hyde, 2021; Rozenfeld *et al.*, 2020; Strully, 2021).

⁶ En el análisis realizado en Zhang *et al.* (2021) para distintas ciudades de China, los autores encuentran una asociación inversa; es decir, mayor incidencia en zonas de menor densidad. Este patrón resulta único y contraintuitivo, pero los autores sugieren que podría deberse a patrones de migración interna temporal en el contexto de festividades locales y de ciudades grandes a ciudades pequeñas y medianas en el contexto de la pandemia.

⁷ Desmet y Wacziarg (2021) encuentran el resultado opuesto para el caso de Estados Unidos.

⁸ Clouston *et al.* (2021) encontraron que, al inicio de la pandemia, un mayor estatus socioeconómico se asociaba con una mayor transmisión del virus. Conforme se producía el distanciamiento social, una política que podía ser más acatada por hogares de mayor estatus, los condados más pobres comenzaron a exhibir mayores tasas de incidencia y mortalidad. Esto es consistente con lo encontrado por De Groot y Lemanski (2021) para Sudáfrica.

La literatura también señala que variables de tipo laboral pueden ser importantes para explicar el impacto del COVID-19. Los estudios revisados incluyen variables como el empleo en actividades esenciales, manuales o de atención al público y el desempleo (Hawkins, 2020; Lewis *et al.*, 2020; Niedzwiedz *et al.*, 2020; Scarpone *et al.*, 2020; Strully *et al.*, 2021). Chauvin *et al.* (2020) sugieren que, entre otros factores, niveles altos de empleo informal podrían contribuir a que los adultos jóvenes y de mediana edad en países en desarrollo tengan más probabilidades de morir de COVID-19.

Otra variable utilizada en gran parte de los estudios revisados es el hacinamiento al interior de la vivienda, medido usualmente como el número de personas por habitación. Este no es solamente un indicador de la cercanía en las interacciones intradomiciliarias, sino de la capacidad de confinar a uno de los miembros del hogar en caso se encuentre infectado. Chauvin (2021) encuentra una conexión fuerte y estadísticamente significativa entre el hacinamiento y la incidencia y mortalidad por COVID-19 en Brasil. Kamis *et al.* (2021) encuentran que inicialmente las disparidades de mortalidad entre condados estadounidenses eran bajas, pero que luego estas se hicieron más evidentes en desmedro de los condados con mayor hacinamiento. También para Estados Unidos, Dasgupta *et al.* (2020) encuentran que aquellos condados con una alta vulnerabilidad social, incluido un alto índice de hacinamiento, tenían una mayor probabilidad de ser considerados como *hotspots* de infección. Kamis *et al.* (2021) encuentran que, controlando por pobreza, el hacinamiento predice de manera importante la mortalidad por COVID-19. En la misma línea se encuentran los resultados encontrados por Ahmad *et al.* (2020); Lewis *et al.* (2020), Niedzwiedz *et al.* (2020), y Strully *et al.* (2021).

La **Tabla 1** presenta un resumen de los principales estudios revisados que analizan la relación entre características sociales y demográficas y el COVID-19.

Tabla 1- Estudios sobre la relación entre características sociales y demográficas y COVID-19

Nº	País	Autor(es)	Año	Periodo de análisis	Principal variable de impacto del COVID-19	VARIABLES para correlación, explicativas y/o de control	Principal resultado
1	Multipaís (18 países de ingreso alto y 13 países en desarrollo)	Chauvin, Fowler y Herrera	2020	Depende del país (hasta mayo o agosto de 2020)	Tasa de mortalidad por COVID-19	Edad, género, minoría étnica, condiciones de la salud pública (pre pandemia).	Los adultos jóvenes y de mediana edad en países en desarrollo tienen más probabilidades de morir de COVID-19 en comparación a sus pares en países desarrollados.
2	Alemania	Scarpone <i>et al.</i>	2020	Enero 2020 – marzo 2020	Casos confirmados de COVID-19	Conectividad, empleo, densidad de iglesias, proporción de visitantes extranjeros que pernoctan.	Los predictores más fuertes de la incidencia de COVID-19 a escala de condado estaban relacionados con la interconexión de la comunidad, la ubicación geográfica, la infraestructura de transporte y la estructura del mercado laboral.
3	Brasil	Chauvin	2021	Febrero 2020 – febrero 2021	Número de muertes <i>per cápita</i> .	Densidad poblacional, tiempo en transporte público, aeropuerto conectado internacionalmente, mayor ingreso, población de mayor edad, hacinamiento domiciliario, presencia de <i>favelas</i> .	Las ciudades con mayor ingreso (mayor movilidad) fueron más afectadas por el COVID-19. Un alto porcentaje de población viviendo en favelas volvía particularmente vulnerable a la ciudad ante el COVID-19.
4	Brasil	Fortaleza <i>et al.</i>	2020	A mayo 2020	Tiempo de introducción del COVID-19 y tasas de incidencia y mortalidad.	Densidad demográfica, distancia a la capital del estado, proporción de población en áreas urbanas, índice de Gini para desigualdades en el ingreso.	El grado de influencia/conectividad de los municipios, así como la densidad demográfica se asociaron positivamente a la introducción temprana del COVID-19 y una mayor incidencia de la enfermedad y mortalidad.
5	Brasil	Freire de Souza, Dornels <i>et al.</i>	2020	Febrero 2020 – mayo 2020	Número de casos por 100,000 habitantes, muertes por millón de habitantes, tasa de letalidad.	17 indicadores asociados a la transmisión del COVID-19: ingreso <i>per cápita</i> menor a medio salario mínimo y tardar más de 1 hora para ir a trabajar, abandono educativo y laboral, entre otros.	La expansión inicia en municipalidades con mayor índice de desarrollo, pero afecta también particularmente a municipalidades de alta vulnerabilidad social.
6	China	You <i>et al.</i>	2020	A febrero 2020	Tasa de morbilidad por COVID-19	Densidad poblacional, proporción de espacio construido, ventas minoristas	En un análisis para los 13 distritos de Wuhan, variables como densidad poblacional, proporción

						por área de tierra, población adulta mayor, PBI por área de tierra, densidad hospitalaria.	de área construida y edad están asociadas positivamente a morbilidad por COVID-19.
7	China	Zhang <i>et al.</i>	2021	A marzo 2020	Número de casos confirmados.	PBI, densidad poblacional, recursos sanitarios (personal, camas, instituciones), distancia a Wuhan.	Mayor riesgo de infección por COVID-19 en ciudades con PBI más alto, recursos de salud limitados y distancia corta a Wuhan. Relación inversa entre incidencia y densidad poblacional.
8	España	Marí-Dell'Olmo <i>et al.</i>	2021	Marzo 2020 – noviembre 2020	Casos acumulados de COVID-19	Grupos de edad, género, rangos de ingresos personales.	Se identificaron inequidades sociales en la incidencia de COVID-19 en Barcelona por grupo de edad, género, área geográfica e ingreso. Se identificaron diferencias entre la primera y la segunda ola..
9	Estados Unidos	Clouston, Natale y Link	2021	Enero 2020 – mayo 2020	Casos diarios confirmados y mortalidad diaria relacionada a COVID-19	Estatus socioeconómico, población mayor de 65 años, afroamericana, hispana, urbanidad.	Al inicio de la pandemia, un mayor estatus socioeconómico se asociaba con el surgimiento de la enfermedad, Posteriormente, los condados de menor estatus terminaron siendo más afectados.
10	Estados Unidos	Dasgupta <i>et al.</i>	2020	Junio – Julio 2020	Clasificación de <i>hotspot</i> de infección por COVID-19 a nivel de condado	Raza, etnia, tipo de vivienda, transporte	Condados con vulnerabilidad social (particularmente minorías raciales y étnicas y con hacinamiento residencial) tenían mayor probabilidad de ser considerados <i>hotspots</i> de infección.
11	Estados Unidos	Hawkins	2020	Enero 2020 – junio 2020	Número de casos positivos de COVID-19	Ingreso promedio; porcentaje de habitantes sin seguro de salud, incidencia de pobreza, desempleo; y porcentaje de trabajadores empleados en servicios de transporte y asistencia sanitaria y social.	Mayor incidencia de COVID-19 en localidades con más pobreza, menores ingresos, menor cobertura de seguro, más desempleo y un mayor porcentaje de la fuerza laboral empleada en servicios esenciales.
12	Estados Unidos	Kamis, Stolte <i>et al.</i>	2021	Abril 2020 – octubre 2020	Número de muertes por COVID-19	Porcentaje de hogares con más de una persona por habitación, porcentaje de familias debajo de la línea de pobreza, población afro, hispana, mayores a 65 años, etc.	Controlando por pobreza, el hacinamiento predice de manera importante la mortalidad por COVID-19.
13	Estados Unidos	Lewis <i>et al.</i>	2020	Marzo 2020 – julio 2020	Casos confirmados de COVID-19	Etnicidad, raza, hacinamiento, tipo de trabajo, inseguridad alimentaria y de atención médica.	Las probabilidades de infección por COVID-19 en áreas de muy alta privación de Utah fueron tres

							veces más altas que aquellas en áreas de muy baja privación.
14	Estados Unidos	Rozenfeld <i>et al.</i>	2020	Febrero – abril 2020	Positividad en la prueba realizada (estudio a nivel de individuo)	Edad, sexo, raza, etnicidad, idioma distinto al inglés, seguridad financiera del vecindario, calidad del aire, inseguridad de la vivienda y del transporte.	Se encontró un mayor riesgo de infección por COVID-19 asociado a variables clínicas, pero también sociodemográficas (raza, etnicidad, condición de la vivienda, del vecindario y del transporte, entre otras).
15	Estados Unidos	Stokes <i>et al.</i>	2021	Enero 2020 – diciembre 2020	Número de muertes por COVID-19, exceso de muertes (incluidas no atribuidas al COVID-19)	Población mayor a 65 años, hispana, afroamericana, variables de comorbilidades (diabetes, obesidad, fumador), ruralidad, mediana de ingreso.	Las muertes totales exceden las oficialmente registradas por COVID-19, y cuando estas son consideradas el impacto en población de ciertas razas y características socioeconómicas es más desigual.
16	Estados Unidos	Strully <i>et al.</i>	2021	A mayo 2020	Número de casos confirmados de COVID-19	Porcentaje de población afroamericana, asiática, latina y nacida en el extranjero. Densidad poblacional, composición de edad y sexo, tasa de desempleo, mediana del ingreso del hogar, número de miembros del hogar, hacinamiento, condiciones de la vivienda (instalaciones incompletas de cocina o fontanería).	Los condados estadounidenses con más residentes inmigrantes, así como con más residentes centroamericanos o negros, tienen más casos de COVID-19. También resultan significativas variables como el porcentaje de viviendas con problemas y el uso de transporte público,
17	Estados Unidos	Xu <i>et al.</i>	2021	Enero 2020 – diciembre 2020	Número de muertes por COVID-19	Sexo, raza, etnicidad, rango de edad, condiciones de salud subyacentes (diabetes, influenza, neumonía).	Los hombres tienen mayores tasas de mortalidad dentro de casi todos los distintos estratos de raza y edad. También se identifica disparidad entre población hispana, afroamericana y blanca.
18	Inglaterra	Nicodemo <i>et al.</i>	2020	Enero 2020 – junio 2020	Muertes por COVID-19	Índice de vulnerabilidad ante el COVID-19 (prevalencia de enfermedades de alto riesgo, urbanidad, disponibilidad de recursos en el sistema de salud, etc).	El índice de vulnerabilidad comunitaria ante el COVID-19 está correlacionado con medidas de privación social. El norte de Inglaterra (mayor privación) se identifica como particularmente vulnerable a la enfermedad.
19	Inglaterra	Niedzwiedz <i>et al.</i>	2020	Marzo 2020 – mayo 2020	Infección confirmada, caso que requiere hospitalización y positividad (estudio a nivel de individuo)	Etnicidad, país de nacimiento, privación socioeconómica (desempleo, propiedad de automóvil y vivienda, hacinamiento), educación, tipo de trabajo, urbano.	Ciertos grupos étnicos tienen considerablemente mayor probabilidad de infección por COVID-19. El menor nivel de educación y las privaciones socioeconómicas estaban asociadas a mayor riesgo de infección y hospitalización.

Fuente: Elaboración propia.

Un segundo grupo más reducido de estudios analiza la relación entre el impacto del COVID-19 (incidencia, mortalidad, letalidad) y variables de acceso o calidad de los servicios de agua y saneamiento (ver **Tabla 2**).

Para los países que conforman la región del África Sub Sahariana, Amankwaa y Fischer (2020) realizan un análisis simple de correlaciones y encuentran que la letalidad por COVID-19 se encuentra negativamente correlacionada con el acceso seguro al agua y al saneamiento. En la misma línea, para Brasil, Silva *et al.* (2021) analizaron los datos de morbilidad y mortalidad y encontraron que las tasas altas de incidencia de la enfermedad se encontraban asociadas significativamente a baja cobertura del servicio de agua y a valores fuera del rango del índice de coliformes fecales en el agua potable. Con respecto a la mortalidad por COVID-19, los autores verificaron una asociación significativa con bajas tasas de cobertura de saneamiento y tratamiento de aguas residuales⁹. En India, Das *et al.* (2020) encontraron que la incidencia de COVID-19 fue 90% más alta en aquellos distritos de la megaciudad de Chennai con poca disponibilidad de servicios de agua, saneamiento e higiene.

Para el caso de Estados Unidos, Ahmad *et al.* (2020) estimaron el riesgo relativo de incidencia y mortalidad relacionados con las malas condiciones de vivienda, controlando por variables sociales y demográficas a nivel de condado. Se considera que una vivienda tiene malas condiciones cuando, entre otros criterios, carece de acceso a agua entubada, inodoro con descarga de agua o una bañera o ducha¹⁰. Los autores encontraron que con cada aumento del 5% en el porcentaje de hogares con malas condiciones de vivienda, hubo un 50% más de riesgo de infección y 42% mayor riesgo de mortalidad por COVID-19. Por su parte, Hyde (2021) realizó un análisis de la letalidad de la enfermedad, tomando en cuenta las infracciones en la calidad del agua a las cuales han estado expuestos los condados de Estados Unidos. El autor encontró que la tasa de letalidad por COVID-19 fue más alta en aquellos condados con más infracciones: 18% más alta en condados más afectados por infracciones graves en materia de salud y 15% más alta en condados más afectados por infracciones que involucran contaminantes asociados a enfermedades cardiovasculares¹¹.

⁹ También se encontraron *clusters* en las regiones Norte y Noreste que se caracterizan por ser las más pobres de Brasil, presentando bajos ingresos, asentamientos humanos y un saneamiento deficiente sistema de saneamiento básico.

¹⁰ Los otros criterios incluyen hacinamiento, alta carga del costo de la vivienda e instalaciones de la cocina incompletas. Se considera que un hogar tiene hacinamiento cuando el ratio de habitantes por habitación es mayor a 1. Se considera que un hogar afronta una carga alta de costo de vivienda cuando este (incluido el pago de servicios públicos) es mayor al 50% de sus ingresos. Un hogar tiene instalaciones de cocina incompletas cuando carece de un lavabo con agua corriente, estufa o cocina o refrigerador..

¹¹ Se consideran como infracciones graves en materia de salud aquellas que suponen una amenaza inmediata para la salud de las personas expuestas, mientras que las infracciones que involucran plomo, arsénico, cadmio y cobre se consideran como aquellas asociadas a un mayor riesgo de enfermedades cardiovasculares.

Finalmente, en un estudio reciente, Revollo-Fernandez *et al.* (2022) encuentran que el número de muertes por COVID-19 a nivel municipal en México se encuentra estadísticamente asociado a determinantes socioeconómicas y de salud, dentro de las cuales se encuentra el acceso al agua. Los autores encuentran que un incremento de 20% en el porcentaje de personas sin acceso al agua está asociado a un aumento promedio de 12,3% en el número acumulado de muertes por COVID-19.

Los estudios presentados analizan la relación entre COVID-19 y las características de la vivienda (incluido el acceso a agua potable y saneamiento) mediante correlaciones que no permiten inferir causalidad. Este es un aspecto que el presente estudio intenta mejorar realizando un aporte metodológico a la literatura existente.

Tabla 2- Estudios sobre la relación entre agua y COVID-19

N°	País/región	Autor(es)	Año	Periodo de análisis	Principal variable de impacto del COVID-19	Variables para correlación, explicativas y/o de control	Principal resultado
1	África Sub Sahariana	Amankwaa y Fischer	2020	Enero 2020 – mayo 2020	Letalidad por COVID-19	Acceso seguro a agua, acceso seguro a saneamiento	Los autores encontraron una correlación entre una mayor tasa de letalidad y bajas tasas en el acceso seguro al agua y al saneamiento.
2	Brasil	Silva <i>et al.</i>	2021	Febrero 2020 – mayo 2020	Incidencia y mortalidad por COVID-19	Cobertura de agua, cobertura de saneamiento, tratamiento de aguas residuales, índice de coliformes fecales en el agua potable.	Las tasas altas de incidencia por COVID-19 se encontraban asociadas significativamente a baja cobertura del servicio de agua y valores fuera del rango del índice de coliformes fecales en el agua potable. Las altas tasas de mortalidad por COVID-19 se encontraban asociadas a cobertura de saneamiento y tratamiento de aguas residuales bajos.
3	Estados Unidos	Ahmad <i>et al.</i>	2020	Data acumulada a marzo y abril 2020	Incidencia y mortalidad por COVID-19	Porcentaje de hogares con condiciones de vivienda pobres (hacinamiento, elevado costo de la vivienda, instalaciones de cocina incompletas, instalaciones de fontanería incompletas), comorbilidades, étnicas y raciales, género, educación, ingreso.	Con cada aumento del 5% en el porcentaje de hogares con malas condiciones de vivienda hubo un 50% más de riesgo de infección y 42% mayor riesgo de mortalidad por COVID-19.
4	Estados Unidos	Hyde	2021	Enero 2020 – setiembre 2020	Mortalidad por COVID19	Infracciones en la calidad del agua, sexo, raza, etnicidad, tasa de pobreza, densidad poblacional, población mayor a 65 años, uso de transporte público, número de miembros del hogar, educación, contaminación del aire.	Los condados con mayor probabilidad de exposición a contaminantes del agua (particularmente los que imponen un riesgo de enfermedad cardiovascular) tienen una mayor tasa de fatalidad de casos, en relación con áreas comparables.

5	India	Das <i>et al.</i>	2020	Mayo 2020	Casos confirmados de COVID-19	Mala condición de vivienda; servicios de agua, saneamiento e higiene; disparidad de género; baja posesión de activos.	Las privaciones socioeconómicas están asociadas significativamente con mayor incidencia de COVID-19 en la megaciudad de Chennai.
6	México	Revollo-Fernandez <i>et al.</i>	2022	Febrero 2020 – setiembre 2021	Mortalidad por COVID-19	Falta de acceso a agua, porcentaje de pobreza, privación social, población vulnerable según ingresos, población masculina, población mayor a 60 años, educación, tasa de mortalidad por neumonía, tasa de mortalidad por diabetes, tasa de mortalidad por hipertensión, número acumulado de infecciones por COVID-19, falta de acceso a servicios de salud, densidad.	A nivel municipal, la mortalidad por COVID-19 está asociada con un mayor porcentaje de población sin agua, mayor porcentaje de población mayor a 60 años, y mayor mortalidad por diabetes.

Fuente: Elaboración propia.

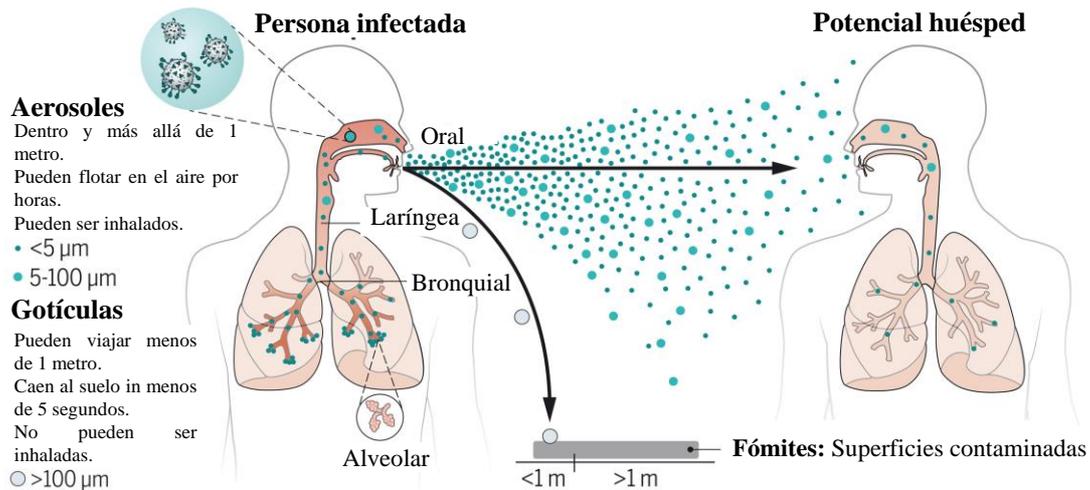
3. Contexto

3.1. COVID-19: transmisión, condiciones de riesgo y prevención

El COVID-19 es una enfermedad transmisible causada por el coronavirus del síndrome respiratorio agudo severo 2 (SARS-CoV-2). Los pacientes pueden presentar síntomas (fiebre, tos, dolores de cabeza, cansancio, dificultad para respirar, pérdida del gusto y el olfato), aunque un porcentaje importante de los pacientes son asintomáticos.

La forma dominante de transmisión del SARS-CoV-2 es por vía respiratoria de una persona infectada a una persona susceptible. La transmisión por vía respiratoria incluye transmisión por aerosoles y transmisión por gotículas de secreciones respiratorias. Al respirar, hablar o toser, las personas infectadas exhalan partículas virales que se encuentran suspendidas en el aire a manera de aerosoles. Estas partículas pueden viajar más de un metro e incluso mantenerse suspendidas en el aire por horas, facilitando la transmisión a distancia (incluso mayor a dos metros) especialmente en ambientes poco ventilados o con poco intercambio de aire. El virus es también excretado en gotículas, las cuales son más grandes (tamaño mayor a 100 micras) y caen rápidamente al suelo o superficies, debido a su peso, contribuyendo a posible transmisión por fómites. Las gotículas pueden caer en las manos, las cuales actúan como fómites especialmente cuando las personas se tocan la cara, la nariz o la boca. Las gotículas tienen un alcance corto (menor a un metro), por lo cual solo contribuyen a transmisiones luego de contacto cercano entre una persona infectada y una persona susceptible a la infección (ver Figura 1). Las personas asintomáticas o presintomáticas excretan el virus en aerosoles y gotículas de manera inadvertida, contribuyendo a la transmisión incluso ante la ausencia de síntomas.

Figura 1- Fases en la transmisión aérea de virus respiratorios



Fases en la transmisión aérea de los virus respiratorios. Los aerosoles cargados de virus ($<100 \mu\text{m}$) son generados en primer lugar por un individuo infectado mediante actividades espiración, a través de las cuales son exhalados y transportados en el ambiente. Pueden ser inhalados por un huésped potencial para iniciar una nueva infección, siempre que sigan siendo infecciosos. A diferencia de las gotículas ($> 100 \mu\text{m}$), los aerosoles pueden permanecer en el aire durante horas y desplazarse más allá de 1 o 2 m del individuo infectado que los exhala, causando nuevas infecciones a corto y largo alcance.

Fuente: Wang *et al.* (2021). Traducción propia.

Debido a la transmisión por aerosoles y gotículas, la proximidad física y la permanencia en lugares cerrados o con pobre ventilación por períodos prolongados de tiempo incrementan el riesgo de transmisión. Ante ello, la mayoría de los países coincide en recomendar el distanciamiento social, el lavado de manos frecuente, el uso de mascarillas y la ventilación de ambientes cerrados como medidas de prevención primaria. La capacidad de los hogares para implementar estas recomendaciones se encuentra en gran medida limitada por sus características socio económicas (precariedad de la vivienda, hacinamiento, empleo informal, falta de acceso a servicios de agua y saneamiento, entre otras).

3.2. El COVID-19 en Perú

El primer caso de COVID-19 en el Perú se confirmó el 6 de marzo de 2020. Nueve días después, mediante Decreto Supremo N° 044-2020-PCM, se declaró el Estado de Emergencia Nacional por las graves consecuencias de la propagación de la enfermedad. La principal medida en este decreto (y prorrogada por sucesivas normas) fue la inmovilización social obligatoria, salvo para actividades esenciales como la compra de alimentos y medicinas. Las medidas incluyeron el cierre de las fronteras, la prohibición del transporte entre regiones, y la reducción en un 50% del transporte urbano. Se cerraron todos los centros educativos y se

dispuso la modalidad remota en los ámbitos educativo¹² y laboral en entidades públicas y privadas. Adicionalmente, se suspendió totalmente la operación de los locales comerciales, incluidos los restaurantes. Se prohibieron las visitas y reuniones sociales y el uso de espacios públicos, y se estableció un toque de queda que inicialmente comenzaba a las 20:00 horas y terminaba a las 05:00 horas del día siguiente. Todas estas restricciones fueron fiscalizadas por las fuerzas armadas y policiales¹³. A fin de incentivar el cumplimiento de las medidas de restricción, el gobierno peruano dispuso la entrega de bonos a familias cuyo ingreso se vio afectado¹⁴; sin embargo, la implementación enfrentó dificultades debido a la baja tasa de bancarización, el alto nivel de informalidad laboral y las carencias en los sistemas de información del Estado. De esta forma, la vulnerabilidad de ciertos grupos sociales ante el COVID-19 se exacerbó en vista de las deficiencias en el sistema sanitario y en los mecanismos de focalización de subsidios o transferencias, así como por la ausencia de ayudas a migrantes y refugiados, entre otros factores (Vásquez-Rowe y Gandolfi, 2020). Especialmente durante el inicio de la pandemia —que se caracterizó por medidas restrictivas de aislamiento social— la precariedad socioeconómica (empleo informal con generación de ingresos día por día o ausencia de un refrigerador en casa) impidió a gran parte de la población cumplir con las medidas de aislamiento por períodos prolongados (Diaz-Cassou *et al.*, 2020; Jaramillo y López, 2021). De este modo, las medidas de cuarentena demostraron ser sub óptimas (Brown *et al.*, 2020).

La evolución del número de casos en el período de estudio de este documento¹⁵ se caracteriza por dos olas de transmisión, que se distinguen claramente en el **Gráfico 1**¹⁶. Particularmente durante la primera ola de transmisión, existió una escasez generalizada de pruebas diagnósticas.

¹² Para ver el impacto del COVID-19 en estudiantes vulnerables de educación superior en Perú, puede revisarse el trabajo de Elacqua *et al* (2022).

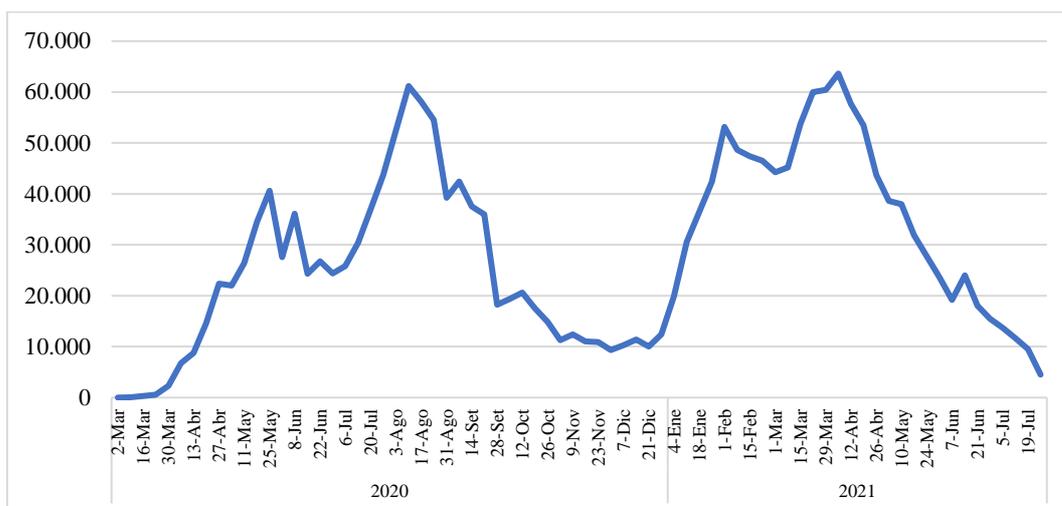
¹³ El Gobierno del Perú implementó temporalmente algunas otras medidas tales como la restricción parcial y total de tránsito vehicular y la implementación de salidas segregadas por género según días, entre otras. En normas sucesivas, se modificaron los horarios de toque de queda, se abrieron las fronteras y progresivamente los locales comerciales y la circulación en espacios públicos, diferenciando las restricciones entre regiones.

¹⁴ Para un inventario y descripción de los principales programas de apoyo al ingreso en América Latina y el Caribe frente al COVID-19, puede revisarse Cejudo *et al.* (2021). De acuerdo con Stampini *et al.* (2021), los bonos “Yo me quedo en casa”, “Bono independiente”, “Bono rural” y “Bono familiar universal” alcanzaron una cobertura conjunta del 38% de la población.

¹⁵ A la fecha de elaboración del presente documento (inicios de 2022), se registraba en Perú la tercera ola de transmisión, con la variante Ómicron como variante predominante y con una tasa de letalidad considerablemente menor a causa de la vacunación y a la menor letalidad inherente de esta variante.

¹⁶ Para revisar las variantes dominantes y la variación espacial de las mismas, pueden revisarse Vargas-Herrera *et al.* (2022) y Romero *et al.* (2021).

Gráfico 1- Número de nuevos infectados semanales por COVID-19 en Perú

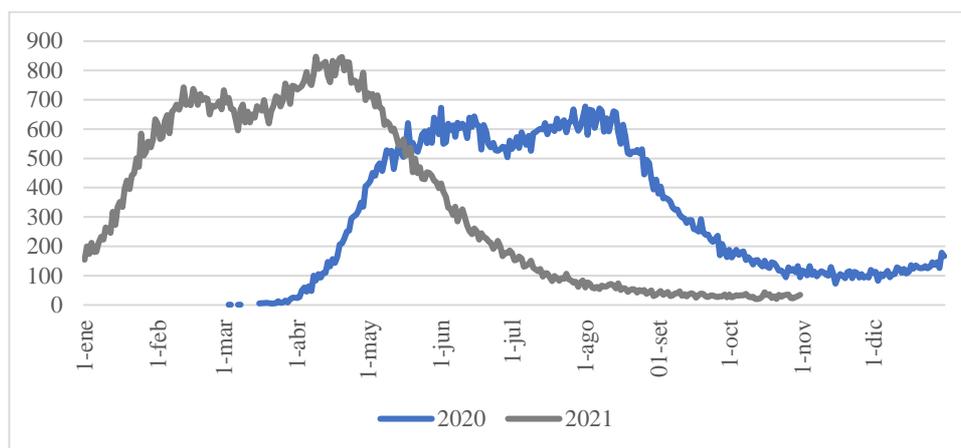


Fuente: Ministerio de Salud. Elaboración propia.

Debido a que el número de infectados puede verse influenciado por la disponibilidad y frecuencia de realización de las pruebas diagnósticas, la revisión de la cifra de fallecidos por COVID-19 da un mejor indicador del impacto de la enfermedad en la población y permite hacer mejores comparaciones con otros países. De marzo 2020 a octubre 2021, se registraron en Perú un total de 201.388 fallecidos por COVID-19. Es importante mencionar que estas cifras ya incluyen las mejoras que se hicieron al sistema de vigilancia de mortalidad por COVID-19 en mayo 2021¹⁷. El **Gráfico 2** presenta el número de fallecidos diarios por COVID-19 donde se pueden observar claramente ambas olas de transmisión.

¹⁷ Mediante Resolución Ministerial N° 095-2021-PCM, se creó el Grupo de Trabajo Técnico de naturaleza temporal con el objeto de proponer los criterios para actualizar la cifra de fallecidos por la COVID-19. El informe final de dicho Grupo de Trabajo Técnico puede encontrarse en <https://cdn.www.gob.pe/uploads/document/file/1920118/Informe%20final%20del%20grupo%20de%20trabajo%20de%20vigilancia%20de%20mortalidad%20por%20COVID-19.pdf>

Gráfico 2- Número de fallecidos diarios por COVID-19 en Perú



Fuente: Plataforma Nacional de Datos Abiertos. Elaboración propia.

De acuerdo con la información publicada por los países, Perú es el país que registra más muertes oficiales por COVID-19 en el mundo (618 por cada 100.000 habitantes)¹⁸. Al considerar la estimación de exceso de muertes —que permite eliminar la subestimación en las cifras oficiales— Perú es el país con mayor número de fallecidos en América Latina, (ver **Tabla 3**)¹⁹.

Tabla 3- Estimación del exceso de muertes en América Latina

País	Exceso de muertes (por 100.000 habitantes)	Diferencia de estimación vs cifra oficial
Perú	640 - 670	+6%
México	490-560	+100%
Bolivia	450-540	+200%
Honduras	240-530	+100%
El Salvador	180-500	+500%
Nicaragua	160-470	+12.000%
Argentina	320-470	+50%
Ecuador	390-430	+100%
Colombia	330-410	+50%
Brasil	330-380	+20%

Fuente: *The Economist* (2022).

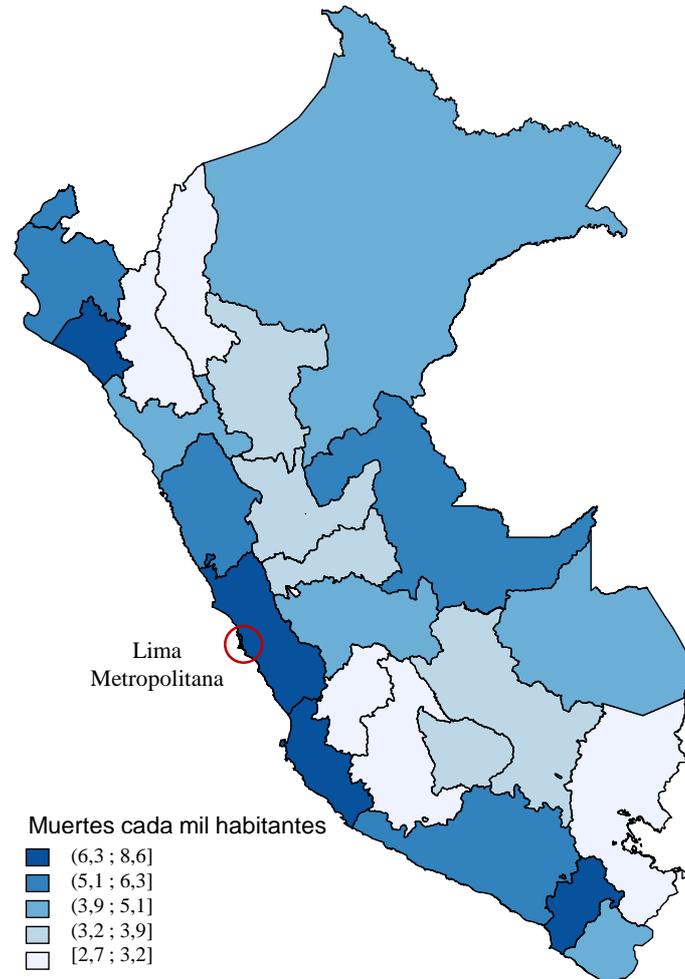
A nivel departamental dentro de Perú (ver **Figura 2**), tomando el indicador de número de muertes por COVID-19 por cada mil habitantes, observamos que los departamentos más afectados han sido: Callao (8,56), Ica (8,22) y Lima (8,10). Esto refuerza la relevancia de

¹⁸ <https://www.economist.com/graphic-detail/coronavirus-excess-deaths-tracker> [revisado el 3 de febrero de 2022]

¹⁹ Cabe destacar que, en comparación con el resto de los países, Perú exhibe una mínima discrepancia entre el número oficial de fallecidos por COVID-19 y el indicador de exceso de muertes, pues el gobierno hizo un esfuerzo por sincerar las cifras, mejorando los criterios para la determinación del número de fallecidos por COVID-19.

realizar el estudio en Lima Metropolitana (que, como se ha mencionado, está conformada por la provincia de Lima y la provincia constitucional del Callao), debido a que se verifica que ha sido una región especialmente afectada por la pandemia.

Figura 2- Muertes por COVID-19 por cada mil habitantes según departamento



Nota: Datos al 31/07/2021. Fuente: Ministerio de Salud. Elaboración propia.

3.3. Situación actual del acceso al agua en Perú

Como en otros países de América Latina y el Caribe (ALC), en Perú persiste una brecha de infraestructura importante en materia de servicios de agua y saneamiento. De acuerdo con el Plan Nacional de Infraestructura para la Competitividad publicado por el Gobierno de Perú (MEF, 2019), las brechas de inversión de corto plazo en agua y saneamiento se estiman en 1.543 y 7.390 millones de dólares²⁰, respectivamente. En el largo plazo, se requerirán inversiones por el monto

²⁰ Tipo de cambio utilizado: 3,9 PEN/USD.

de 6.217 y 18.345 millones de dólares para alcanzar niveles de acceso similares a los de países desarrollados²¹. En un estudio más reciente, Brichetti *et al.* (2021) han estimado que Perú necesitaría invertir 20.901 millones de dólares hasta 2030 a fin de brindar acceso universal al agua y saneamiento gestionados de manera segura y garantizar el tratamiento de aguas residuales en zonas urbanas. Ello permitiría que el país construya los principales componentes de infraestructura para el cumplimiento del Objetivo de Desarrollo Sostenible N° 6.

De acuerdo con información oficial, en los últimos años ha existido un incremento ligero en la cobertura de agua: en 2013 el 86,1% de la población se abastecía a través de red pública, mientras que en 2020 dicho porcentaje ascendió a 91,2% (ver **Tabla 4**).

Tabla 4- Hogares con abastecimiento de agua por red pública, 2013-2020 (%)

Ámbito	2013	2014	2015	2016	2017	2018	2019	2020
Nacional	86,1	87,6	88,2	89,2	89,4	90,7	90,8	91,2
Urbano	93,4	93,6	93,9	94,5	94,4	95,3	94,9	94,8
Rural	63,2	68,3	69,5	71,2	72,2	74,4	75,6	77,6

Fuente: INEI (2021). Elaboración propia.

Para el año 2020, de acuerdo con información de la Encuesta Nacional de Hogares (ENAH), 90% de los hogares a nivel nacional contaban con acceso a través de red pública, con gran variabilidad entre áreas geográficas. Como puede verse en la **Tabla 5**, en el caso particular de Lima Metropolitana, un 5,7% no cuenta con este servicio, recurriendo mayoritariamente al abastecimiento a través de camiones cisterna (4,6%). Considerando que Lima Metropolitana concentra casi un tercio de la población nacional, ello significa que más de medio millón de personas en esta región no cuentan con el servicio de agua potable.

²¹ En el mismo sentido, la brecha estimada a 2030 por el Ministerio de Vivienda, Construcción y Saneamiento es de alrededor de 25.641 millones de dólares (Ministerio de Vivienda, Construcción y Saneamiento, 2021).

Tabla 5- Acceso al servicio de agua por tipo de abastecimiento y dominio geográfico en 2020 (%)

Tipo de abastecimiento de agua	Costa norte	Costa centro	Costa sur	Sierra norte	Sierra centro	Sierra sur	Selva	Lima Metrop.	Total
Red pública dentro de la vivienda	87,2	84,0	85,5	85,9	82,9	75,1	75,1	87,4	83,2
Red pública fuera de la vivienda (dentro del edificio)	1,3	1,4	1,6	3,8	7,2	7,6	4,1	4,4	4,4
Pilón o pileta de uso público	1,5	3,2	4,9	0,3	1,0	1,9	1,5	2,5	1,9
Camión-cisterna u otro similar	4,1	6,3	4,1	-	0,2	2,0	1,5	4,6	3,0
Pozo (agua subterránea)	1,4	1,5	1,2	0,5	0,4	6,2	3,6	0,4	1,8
Manantial	0,1	0,1	0,1	5,3	4,3	2,3	2,6	0,0	1,6
Otra forma	3,7	2,5	1,5	1,8	1,8	4,1	4,1	0,8	2,4
Río, acequia, lago, laguna	0,9	0,9	1,1	2,4	2,1	0,9	7,5	-	1,6
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

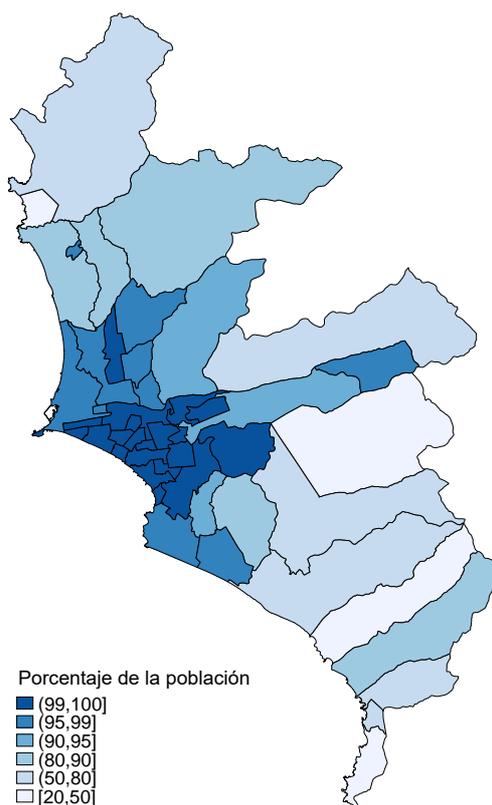
Nota: El dominio de Lima Metropolitana incluye los distritos de la provincia de Lima y los distritos que componen la provincia constitucional del Callao. Fuente: Base de datos ENAHO 2020. Elaboración propia.

Los hogares que no están conectados, y se abastecen principalmente mediante camiones cisterna, racionan su consumo debido al alto precio que pagan y a que no se les garantiza un suministro continuo, por lo cual deben almacenar el agua en contenedores. De acuerdo con una encuesta realizada en Lima Metropolitana (Sunass, 2015), los hogares no conectados consumen en promedio 4,8 metros cúbicos al mes (cifra muy inferior al consumo promedio de un hogar conectado a la red de Sedapal, que se encuentra alrededor de los 15 metros cúbicos al mes) y pagan en promedio 6 veces más por cada metro cúbico. Además, este tipo de abastecimiento implica riesgos a la salud debido a la calidad del agua recibida. En el contexto de transmisión de COVID-19, el abastecimiento de agua a través de camiones cisterna puede aumentar la exposición al virus debido a que, para obtener agua, las personas a menudo deben salir de sus casas, formar colas o aglomerarse alrededor de un camión cisterna, lo cual puede incrementar las interacciones sociales cercanas y favorecer la transmisión del virus. Asimismo, como se discutirá más adelante, el acceso al agua está fuertemente correlacionado con el acceso al saneamiento. Las personas que habitan en viviendas que no están conectadas a la red de alcantarillado utilizan instalaciones como pozos sépticos, letrinas compartidas o inclusive podrían realizar la defecación al aire libre.

En Lima Metropolitana, la empresa prestadora encargada es Servicio de Agua Potable y Alcantarillado de Lima (Sedapal). La cobertura que brinda esta empresa varía entre los 43 distritos que componen la provincia de Lima y los 7 distritos contenidos en la provincia Constitucional del

Callao²². Como puede verse en la **Figura 3**, el crecimiento de la ciudad ha ocasionado que los mayores retos para expandir la cobertura de agua potable se encuentren en la periferia. Nuestro análisis está centrado precisamente en estos distritos, donde coexisten usuarios con y sin acceso al agua. El objetivo de nuestro documento es evaluar la asociación de la variable de acceso al agua con la infección por COVID-19, controlando por otras variables que pudieran incidir en el resultado.

Figura 3 - Distritos de Lima Metropolitana, según porcentaje de acceso al agua mediante red pública



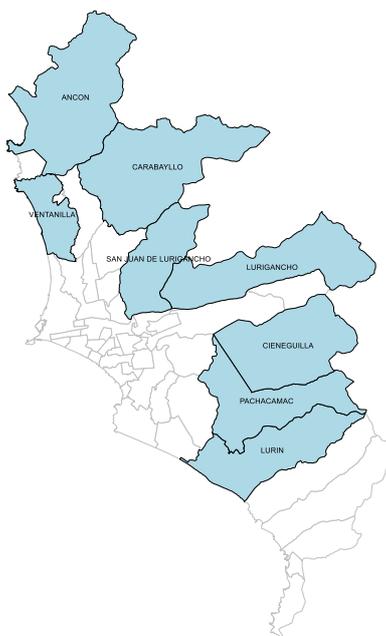
Fuente: INEI – Sistema de Consulta de Base de Datos de los Censos Nacionales 2017.
Elaboración propia.

²² Distritos de la provincia de Lima: Ancón, Ate, Barranco, Breña, Carabayllo, Chaclacayo, Chorrillos, Cieneguilla, Comas, El Agustino, Independencia, Jesús María, La Molina, La Victoria, Lima, Lince, Los Olivos, Lurigancho, Lurín, Magdalena Del Mar, Miraflores, Pachacamac, Pucusana, Pueblo Libre, Puente Piedra, Punta Hermosa, Punta Negra, Rímac, San Bartolo, San Borja, San Isidro, San Juan De Lurigancho, San Juan De Miraflores, San Luis, San Martín De Porres, San Miguel, Santa Anita, Santa María Del Mar, Santa Rosa, Santiago De Surco, Surquillo, Villa El Salvador y Villa María Del Triunfo. Distritos de la provincia constitucional del Callao: Bellavista, Callao, Carmen De La Legua Reynoso, La Perla, La Punta, Mi Perú y Ventanilla.

4. Características de los datos

Los datos utilizados provienen de una encuesta diseñada y aplicada específicamente para este estudio. El trabajo de campo fue realizado entre octubre y noviembre de 2021, en modalidad presencial, en áreas periurbanas de los siguientes distritos de Lima Metropolitana: Ancón, Carabayllo, Cieneguilla, Lurigancho, Lurín, Pachacamac, San Juan de Lurigancho y Ventanilla (ver **Figura 4**). Se recolectó información para un total de 1.121 hogares y en las preguntas relacionadas a salud, se cuenta con información para 3.330 miembros del hogar²³. El proceso de muestreo fue el siguiente: se seleccionaron aleatoriamente un total de 634 hogares sin acceso al servicio de agua (grupo de control) en zonas con discontinuidad geográfica en la prestación del servicio de agua potable. Luego, se identificaron los hogares que tuvieran acceso al agua y estuvieran más cercanos a los hogares sin acceso (grupo de tratados). La ficha técnica de la encuesta puede revisarse en el **Anexo 1** de este documento.

Figura 4- Distritos seleccionados para la aplicación de la encuesta



Nota: Datos al 31/07/2021. Fuente: Ministerio de Salud. Elaboración propia.

El 46,4% de la muestra se encontraba conectada a la red pública de agua, mientras que el 53,6% tenía tipos de abastecimiento alternativos, principalmente camiones cisterna y pilones de uso público (ver **Tabla 6**). Si bien el agua extraída de los pilones de uso público proviene de la red pública de la empresa de agua, los hogares con este tipo de acceso no se consideran dentro de

²³ Como se verá más adelante, las estimaciones se han realizado tomando en cuenta a los hogares ubicados, a una distancia máxima de 2 kilómetros, a uno y otro lado de la discontinuidad geográfica. Por tanto, esta sección presenta información descriptiva de los 996 hogares y 3.015 individuos que cumplen con dicho criterio.

las observaciones conectadas (grupo de tratados) para los efectos del presente estudio. Esto debido a que –como mencionan Howard, Bartram *et al.* (2020)– las fuentes de agua compartida pueden asociarse a mayor potencial de transmisión del virus, debido a que dificultan el aislamiento y la distancia social, e involucran la manipulación de llaves de agua por múltiples personas.

Tabla 6- Hogares por tipos de abastecimiento de agua en la muestra (%)

	Tipo de abastecimiento	%
Conectado a red pública (Grupo de tratados)	Red pública por tubería que llega hasta dentro de la vivienda	45,2
	Red pública por tubería que llega hasta fuera de la vivienda pero dentro de la edificación	1,3
Sin conexión a red pública (Grupo de control)	Camión cisterna	36,7
	Pilón de uso público	9,6
	Solicitada al vecino con acceso a agua por red pública	4,3
	Pozo de agua subterránea	2,7
	Otros	0,3

Fuente: Encuesta en zonas periurbanas de Lima Metropolitana. Elaboración propia tomando en consideración factores de expansión.

Del total de miembros del hogar de la muestra, 28,7% se realizó una prueba diagnóstica de COVID-19 (ver **Tabla 7**).

Tabla 7- Porcentaje de miembros del hogar con prueba diagnóstica de COVID-19 (%)

	Total	Conectados a la red pública de agua (Grupo de tratados)	No conectados a la red pública de agua (Grupo de control)
Sí se hizo una prueba	28,7	30,1	27,5
No se hizo una prueba	71,3	69,9	72,5

Fuente: Encuesta en zonas periurbanas de Lima Metropolitana. Elaboración propia tomando en consideración factores de expansión.

Dentro de aquellos miembros que se realizaron una prueba diagnóstica, 34,2% corresponden a pruebas moleculares, 33,9% a pruebas rápidas (muestra de sangre del dedo), 27,2% a pruebas de antígenos y 4,7% a pruebas de Elisa (muestra de sangre obtenida del brazo)²⁴.

²⁴ Los encuestadores utilizaron ayuda visual (infografías) con las principales características de la toma de muestras de los distintos tipos de prueba a fin de ayudar a los encuestados a responder esta pregunta.

Tabla 8- Tipo de pruebas diagnósticas de COVID-19 realizadas (%)

	Total	Conectados a la red pública de agua (Grupo de tratados)	No conectados a la red pública de agua (Grupo de control)
Molecular	34,2	34,7	33,8
Rápida	33,9	32,6	35,1
Antígeno	27,2	28,0	26,4
Elisa	4,7	4,8	4,7

Nota. Miembros del hogar que se realizaron una prueba COVID-19. Fuente: Encuesta en zonas periurbanas de Lima Metropolitana. Elaboración propia tomando en consideración factores de expansión.

Con respecto a los resultados obtenidos en la prueba diagnóstica, el índice de positividad es de 33,6%, con una diferencia entre quienes se encuentran conectados a la red pública de agua (grupo de tratados) y quienes no se encuentran conectados (grupo de control), como se ve en la **Tabla 9**. Esta diferencia de alrededor de diez puntos porcentuales en la positividad entre ambos grupos no debe ser interpretada automáticamente como el impacto del acceso en la probabilidad de infección. Si bien los grupos de control y tratamiento deberían ser, por construcción, similares en el resto de las características, la metodología desarrollada a continuación nos permitirá afinar la estimación teniendo en consideración las diferencias que se han encontrado entre las observaciones de control y tratamiento, a pesar de encontrarse geográficamente cerca.

Tabla 9- Resultados de la prueba diagnóstica de COVID-19

	Total	Conectados a la red pública de agua	No conectados a la red pública de agua
Positiva	33,6	28,0	38,9
Negativa	66,4	72,0	61,1

Nota. Miembros del hogar que se realizaron una prueba COVID-19. Fuente: Encuesta en zonas periurbanas de Lima Metropolitana. Elaboración propia tomando en consideración factores de expansión.

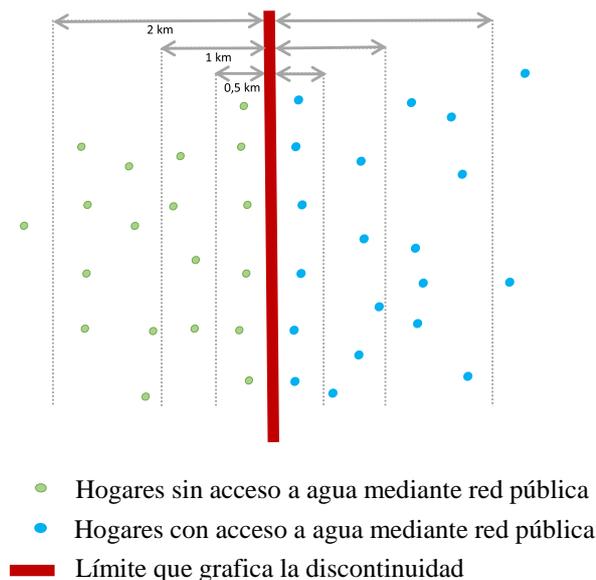
5. Metodología

El presente documento es un estudio transversal de diseño cuasi-experimental con un grupo comparador, cuyo objetivo es identificar la asociación entre acceso a agua mediante red (variable independiente, de intervención o de tratamiento) e infección por COVID-19 (variable resultado, dependiente o de impacto). La unidad de análisis es el individuo.

El método utilizado es el de “Regresión Discontinua” (RD, por sus siglas en inglés). Los estudios con diseño RD permiten determinar el efecto causal de una intervención o tratamiento, considerando que éste ha sido asignado teniendo como criterio un punto de corte o umbral en determinada variable observable. Este tipo de estudios ha sido ampliamente utilizado en economía (Lee and Lemieux, 2010) y es recomendado de manera creciente en estudios epidemiológicos y

tienen²⁵, trazada en el punto equidistante entre ambos tipos de hogares. Asimismo, el análisis incluye bandas de proximidad a la frontera o límite, como puede verse en la **Figura 6**. Al respecto, debe considerarse que el efecto marginal de la variable de tratamiento puede aproximarse mejor en la medida que los hogares se encuentren más cerca a la discontinuidad geográfica. Además, la gran mayoría de las observaciones (alrededor del 82%) se encuentran a una distancia máxima de 0,5 km de la frontera.

Figura 6 – Definición de la discontinuidad geográfica



Fuente: Elaboración propia.

Los hogares más cercanos a la frontera deberían compartir características socioeconómicas, a excepción de su acceso a agua mediante red. Por tanto, en teoría, las diferencias en la variable de impacto o resultado debería atribuirse al acceso al agua. En la práctica, sin embargo, es posible que existan diferencias en algunas características relevantes para el estudio, para lo cual se incorporan variables de control.

Para la estimación se sigue el enfoque desarrollado en Dell (2010) que consiste en incorporar como controles las variables de longitud y latitud de manera polinomial. Dado que la variable dependiente es de naturaleza dicotómica (1 si el individuo se infectó por COVID-19 y 0 en caso contrario), se plantea el siguiente modelo logit para la ecuación principal:

²⁵ La encuesta piloto realizada antes de la recolección de información de la muestra total permitió advertir que en algunos casos las viviendas cuentan con instalaciones sanitarias intradomiciliarias, pero no se abastecen de Sedapal. Suelen emplear soluciones comunales que les brindan servicios de red y grifos de agua. Para los efectos del presente estudio, dichas observaciones son consideradas como observaciones de tratamiento, pues permiten un suministro relativamente continuo de agua que discurre y no requiere que los hogares salgan de sus viviendas para abastecerse de agua.

$$Pr[Y_i = 1] = \Phi(\beta + \beta T_i + \theta f(\text{ubicación geográfica}) + X'\gamma)$$

Donde:

Y_i es una variable dicotómica que toma el valor de 1 si el individuo “ i ” resultó positivo en una prueba diagnóstica de COVID-19 en el periodo comprendido entre marzo de 2020 y la fecha de recolección de información.

T_i es la variable dicotómica de tratamiento que indica si el individuo “ i ” tiene acceso al servicio de agua.

El término $f(\dots)$ es un polinomio de grado 2 que utiliza los datos de latitud y longitud de cada hogar.

El término X' representa un vector de variables de control.

El término Φ es la función de densidad acumulada de la distribución logística estándar.

Debido a que la variable dependiente requiere conocer el antecedente de infección por COVID-19 de las personas, la muestra utilizada corresponde a aquellos individuos que reportaron haberse realizado una prueba diagnóstica. Si bien esto reduce la muestra total, consideramos que es la alternativa más rigurosa para recoger el antecedente de infección por COVID-19 de manera adecuada²⁶. Al respecto, no existen diferencias estadísticamente significativas en el porcentaje de individuos con prueba diagnóstica entre observaciones de control y tratamiento (ver test de medias en el **Anexo 2**). Es decir, no existieron diferencias significativas en el acceso a pruebas diagnósticas entre uno y otro lado de la frontera de prestación del servicio de agua²⁷. Al analizar las principales variables recogidas para el presente estudio (ver test de medias en el **Anexo 3**), tampoco se verifica que existan diferencias significativas entre los individuos que accedieron a una prueba diagnóstica y los que no accedieron a una. Debido a ello, consideramos que utilizar únicamente las observaciones de individuos que cuentan con una prueba diagnóstica no introduce sesgo de selección y que, por tanto, la metodología propuesta resulta adecuada para estimar el efecto del tratamiento (acceso al agua) en la probabilidad de registrar antecedente de infección por COVID-19²⁸.

²⁶ Otras alternativas descartadas fueron asumir supuestos con respecto a la positividad de los individuos sin prueba diagnóstica o utilizar la información auto reportada de síntomas para inferir la positividad. En este último caso, debe considerarse que los síntomas no son una condición suficiente (los síntomas pueden ser similares a los de otras enfermedades respiratorias, entre otras) ni necesaria (un número importante de infectados puede ser asintomático).

²⁷ Esto puede atribuirse a que —sobre todo al inicio de la pandemia— la escasez de pruebas diagnósticas era generalizada, lo cual causó que no existieran diferencias significativas en términos de acceso entre la población. Posteriormente, el Ministerio de Salud realizó campañas de detección en lugares como mercados o estaciones de transporte. En meses posteriores, los laboratorios privados ofrecieron pruebas diagnósticas, aunque es esperable que su elevado precio haya constituido una barrera al acceso para la mayoría de los hogares (de control y tratamiento) comprendidos en este estudio.

²⁸ Como se verá más adelante, si la estimación se realizara considerando sesgo de selección, el efecto marginal resulta menor, pero continúa estando dentro del intervalo de confianza de los resultados primarios presentados en este estudio.

6. Resultados

Esta sección presenta los resultados de la estimación del modelo logit previamente planteado, así como una extensión de este, incluyendo efectos heterogéneos relacionados a la dotación per cápita de agua de los hogares. Asimismo, el análisis se complementa con un test de falsificación o placebo a fin de verificar la validez de las estimaciones. Se reportan en paréntesis los errores estándar robustos clusterizados a nivel distrital.

6.1. Modelo Logit RD: el rol del acceso al agua

La **Tabla 10** presenta los resultados de la estimación en cuatro columnas que corresponden a cuatro rangos de distancia a uno y otro lado de la frontera geográfica de prestación: 0,5 km; 1 km; 1,5km y 2km. Como se ha mencionado anteriormente, debe considerarse que, en promedio, el 82% de las observaciones se encuentran a una distancia menor o igual a 0,5 km de la discontinuidad y que el efecto marginal de la variable de tratamiento puede aproximarse mejor en la medida que los hogares se encuentren más cerca a la discontinuidad geográfica.

Como puede verse, la variable de *tratamiento* resulta significativa y el signo del coeficiente es negativo. Es decir, el acceso al agua por red está asociado a una reducción en la probabilidad de reportar antecedente de infección por COVID-19.

La variable *edad* también resulta significativa (para distancias mayores a 0,5km) y con un coeficiente pequeño, aunque con el signo esperado. Si bien no puede aseverarse que una mayor edad aumente *per se* la probabilidad de infección, esta variable puede estar recogiendo la mayor exposición al virus por participación en el mercado laboral, realización de tareas tales como compras domésticas, y el hecho de que los niños y adolescentes tuvieron que acatar un estricto confinamiento, habiéndose retomado las clases presenciales dos años después. Esta variable también podría recoger el hecho de que la probabilidad de realizarse una prueba diagnóstica se incrementa con la edad.

La variable *primera ola*, que es una variable dicotómica que toma el valor 1 si la prueba se hizo entre marzo y diciembre 2020 y 0 si se hizo después, también resulta significativa. Existen distintas explicaciones posibles para este resultado. Una primera es que, según los antecedentes revisados del caso peruano, la disponibilidad de pruebas al inicio de la pandemia era limitada. Por tanto, estas fueron administradas dando prioridad a personas que reportaban síntomas y/o habían tenido contacto con pacientes infectados, por lo cual la probabilidad de un resultado positivo era mayor en la primera ola. Otras explicaciones no excluyentes son que las medidas de restricción en la primera ola fueron más estrictas (y pudieron favorecer la transmisión), el poco conocimiento sobre las vías de transmisión del virus al inicio de la

pandemia y otros aspectos relacionados a las variantes predominantes del virus en la primera ola.

La presencia de *comorbilidades*, definida como una variable dicotómica que recoge si el individuo padece de sobrepeso, diabetes, enfermedades cardíacas, enfermedades pulmonares o sistema inmunitario debilitado también resulta significativa y con el signo esperado. Como se ha revisado, las comorbilidades están asociadas a una mayor severidad de la enfermedad, lo que puede haber incrementado las probabilidades de buscar y acceder a una prueba diagnóstica.

Finalmente, la variable *bono* resulta significativa y con un signo negativo para los tres primeros rangos de distancia presentados. Esta variable dicotómica indica si el hogar recibió alguno de los bonos del Ministerio de Desarrollo e Inclusión Social (MIDIS), tales como el Bono Yo me Quedo en Casa o el Bono Familiar Universal. Esto indicaría que haber recibido un bono reducía la probabilidad de infección por COVID-19, presumiblemente debido a que esta ayuda financiera favorecía el aislamiento social y, por tanto, reducía la exposición de los miembros del hogar a la infección por COVID-19.

Otras variables que se incluyen en la estimación, aunque no resultan estadísticamente significativas son el *hacinamiento* en la vivienda (medido como el número de miembros del hogar por habitación), así como variables dicotómicas que recogen si el hogar registra como mínimo *secundaria incompleta de la madre*, si la persona que aporta los ingresos principales tiene un *trabajo presencial con atención al público*, si el hogar habita en una *vivienda precaria* (es decir, si no está construida con material noble) y si el hogar tiene en promedio mensual *ingresos mayores a 1000 soles*. Si bien estas variables resultaron significativas en muchos de los trabajos revisados, debe considerarse que probablemente fueron factores que favorecieron la transmisión inicial del SARS-CoV-2; mientras que la variable dependiente del presente trabajo recoge los antecedentes de infección por COVID-19 acumulados a octubre - noviembre 2021; es decir, para la primera y segunda ola de transmisión.

Tabla 10- Resultados Modelo Logit

	<0,5 km (1)	<1 km (2)	<1,5 km (3)	<2 km (4)
<i>Tratamiento</i>	-0,505*** (0,1740)	-0,336* (0,1890)	-0,350** (0,1700)	-0,302* (0,1590)
<i>Edad</i>	0,004 (0,0060)	0,010* (0,0050)	0,010* (0,0050)	0,009* (0,0060)
<i>Secundaria incompleta de la madre</i>	-0,265 (0,1740)	-0,243 (0,2150)	-0,248 (0,2060)	-0,195 (0,1970)
<i>Trabajo presencial con atención al público</i>	-0,091 (0,1750)	-0,019 (0,1190)	-0,062 (0,1290)	-0,044 (0,1040)
<i>Vivienda precaria</i>	0,081 (0,1930)	0,213 (0,2170)	0,195 (0,2040)	0,306* (0,1850)
<i>Hacinamiento</i>	-0,063 (0,1740)	-0,022 (0,1340)	-0,015 (0,1280)	-0,023 (0,1160)
<i>Ingresos mayores a 1000 soles</i>	-0,454 (0,2830)	-0,279 (0,2270)	-0,219 (0,1830)	-0,208 (0,1580)
<i>Primera ola</i>	0,915*** (0,1920)	0,846*** (0,1060)	0,900*** (0,1170)	0,886*** (0,1210)
<i>Comorbilidades</i>	0,754*** (0,1710)	0,727*** (0,1900)	0,752*** (0,2040)	0,721*** (0,2230)
<i>Bono</i>	-0,300* (0,1730)	-0,390** (0,1860)	-0,391** (0,1880)	-0,287 (0,2460)
<i>Constante</i>	-0,289 (0,4620)	-0,830** (0,3810)	-0,833** (0,3830)	-0,910** (0,3680)
Incluye variables de geolocalización	Sí	Sí	Sí	Sí
Observaciones	612	707	728	749
Log Likelihood	-362,645	-420,849	-430,603	-445,965
Akaike Inf. Crit.	765,289	881,698	901,205	931,929

Los niveles de significancia son: ***p<0,01, **p<0,05, *p<0,1

Nota: Entre paréntesis se muestran errores estándar robustos clusterizados a nivel distrital.

Fuente: Elaboración propia.

Debido a que el modelo estimado es un modelo logit, los coeficientes no pueden ser interpretados directamente, por lo cual es necesario calcular los efectos marginales²⁹. La **Tabla 11** presenta los efectos marginales de la variable de tratamiento, calculados para los hogares ubicados a distancias de 0,5 km, 1 km, 1,5 km y 2 km de la discontinuidad geográfica en la

²⁹ Al realizar las estimaciones con un Modelo de Probabilidad Lineal, el coeficiente asociado a la variable de tratamiento es el mismo que el efecto marginal calculado para el Modelo Logit.

prestación. De acuerdo con los efectos marginales calculados, contar con agua de red está asociado a una reducción de entre 6,2 y 9,9 puntos porcentuales en la probabilidad de infectarse de COVID-19. Considerando el nivel de positividad de los usuarios no conectados, esto significaría que la probabilidad de infectarse de COVID-19 se reduce entre 15% y 25% cuando el hogar cuenta con acceso al agua mediante red³⁰.

Tabla 11- Efectos Marginales de la variable de tratamiento según distancia

Distancia	AME	SE	z	p	lower	upper
0,5 km	-0,0999	0,0345	-2,8939	0,0038	-0,1676	-0,0323
1 km	-0,0646	0,0364	-1,7760	0,0757	-0,1359	0,0067
1,5 km	-0,0673	0,0336	-2,0067	0,048	-0,1331	-0,0016
2 km	-0,0618	0,0329	-1,8805	0,0600	-0,1262	0,0026

Fuente: Elaboración propia.

En el **Anexo 5** se puede ver la representación gráfica de la probabilidad de infección por COVID-19 de los individuos en función de la distancia que los separa de discontinuidad geográfica. En el panel de la izquierda se encuentran graficadas las observaciones de control (no conectados a la red), mientras que en la parte derecha se encuentran las observaciones de tratamiento (conectados a la red). El gráfico permite apreciar el salto en la probabilidad de infección por COVID-19³¹.

³⁰ El Anexo 4 presenta la estimación para casos de sesgo por selección. En estos casos su corrección puede realizarse mediante la metodología propuesta por Heckman (1979). Sin embargo, Greene (2006) señala que esta metodología no resulta apropiada cuando la ecuación principal a estimar es no lineal. Para estos casos, propone un método general para corregir el problema de sesgo de selección mediante la maximización de la función de verosimilitud por simulaciones.

Este método consiste en la estimación de dos ecuaciones; la ecuación de selección (que en este caso estima la probabilidad de realizarse una prueba diagnóstica) y la ecuación principal (que en este caso estima la probabilidad de registrar antecedente de infección por COVID-19).

Como puede verse, el efecto marginal resulta menor, pero continúa estando dentro del intervalo de confianza de los resultados primarios presentados en este estudio.

³¹ El gráfico presenta polinomios de grado dos a ambos lados de la discontinuidad. Debe recordarse que en este gráfico se presenta la probabilidad de antecedente de infección por COVID-19 únicamente en función de la distancia, por lo que la atención debe centrarse en el salto existente en la frontera de prestación del servicio de agua.

6.2. Modelo Logit con efectos heterogéneos por nivel de consumo: más allá del acceso, entendiendo la importancia de una dotación mínima

En esta sección se presentan los resultados de una extensión al modelo previamente definido: un modelo con efectos heterogéneos, estratificando por nivel de consumo de los usuarios conectados. Esta especificación divide a los individuos que tienen acceso al agua de red en dos grupos: aquellos que consumen menos de 3 metros cúbicos *per cápita* mensuales (o el equivalente a 150 litros por habitante por día)³² y aquellos que consumen más de ese nivel³³.

La **Tabla 12** muestra los resultados para la estimación del modelo logit con efectos heterogéneos por dotación de consumo de agua. Por motivos de simplicidad de incluyen dos rangos de distancia (0,5 km y 1,5 km). Bajo el grupo de bajo consumo, se comparan observaciones de control y tratamiento cuyo consumo mensual per cápita no excede los 3 m³. Como puede verse, en este escenario la variable de tratamiento no resulta significativa. Bajo el grupo de alto consumo, se comparan observaciones de control con un consumo mensual per cápita que no excede los 3 m³ y observaciones de tratamiento con un consumo superior a los 3 m³. En este caso, el coeficiente asociado a la variable de tratamiento resulta significativo. Esto estaría indicando que no basta con estar conectado a la red, sino que además debe garantizarse una dotación mínima de agua. Posibles explicaciones para estos resultados son que se necesita una cantidad mínima para cumplir con las recomendaciones de higiene, pero también que los hogares conectados a la red (pero con poca calidad) podrían tener que complementar su provisión con agua proveniente de camiones cisterna o pilones de uso público, lo cual los expone a un mayor riesgo de infección por COVID-19. Los efectos marginales de este modelo, para el grupo de alto consumo, pueden revisarse en la **Tabla 13**.

³² De acuerdo con la Organización Mundial de la Salud (2017), para considerarse acceso óptimo, el volumen de agua requerido se encuentra entre los 100 y 200 litros por habitante por día

³³ Para hallar el consumo per cápita, se estima el consumo de agua a nivel de hogar y se divide entre el número de miembros del hogar. Para los hogares conectados, el consumo total a nivel de hogar se infiere a partir del monto mensual que pagan por agua y saneamiento y la estructura tarifaria vigente de la empresa Sedapal S.A en la fecha del trabajo de campo. Para los hogares no conectados, el consumo total del hogar se estima a través de las preguntas incluidas en el cuestionario sobre recipientes, contenido y frecuencia de almacenamiento de agua.

Tabla 12- Resultados Modelo Logit con efectos heterogéneos

	Distancia <0,5 km		Distancia <1,5 km	
	Bajo consumo (Menor a 3m ³ per cápita)	Alto consumo (Mayor o igual a 3m ³ per cápita)	Bajo consumo (Menor a 3m ³ per cápita)	Alto consumo (Mayor o igual a 3m ³ per cápita)
<i>Tratamiento</i>	-0,365 (0,4240)	-0,754** (0,3110)	-0,14 (0,3570)	-0,495* (0,2590)
<i>Edad</i>	0,005 (0,0090)	0,003 (0,0050)	0,011 (0,0080)	0,010*** (0,0040)
<i>Secundaria incompleta de la madre</i>	-0,147 (0,2150)	0,005 (0,2290)	-0,25 (0,2060)	-0,096 (0,2120)
<i>Trabajo presencial con atención al público</i>	0,369* (0,2130)	-0,018 (0,3420)	0,241 (0,1580)	0,06 (0,2190)
<i>Vivienda precaria</i>	0,074 (0,2990)	-0,325* (0,1910)	0,233 (0,2750)	-0,11 (0,2160)
<i>Hacinamiento</i>	-0,201 (0,2250)	-0,116 (0,2010)	-0,155 (0,1620)	-0,067 (0,1420)
<i>Ingresos mayores a 1000 soles</i>	-0,058 (0,6660)	-0,582 (0,4020)	0,044 (0,4410)	-0,266 (0,2920)
<i>Primera ola</i>	0,687*** (0,1870)	0,947*** (0,3250)	0,766*** (0,1370)	0,917*** (0,2030)
<i>Comorbilidades</i>	0,763** (0,3660)	0,308 (0,2230)	0,707*** (0,2360)	0,332* (0,1980)
<i>Bono</i>	-0,054 (0,3440)	-0,314 (0,2870)	-0,146 (0,3610)	-0,453** (0,2260)
<i>Constante</i>	-0,398 (0,4480)	0,029 (0,6180)	-0,803* (0,4480)	-0,621 (0,4310)
Incluye variables de geolocalización	Sí	Sí	Sí	Sí
Observations	298	416	376	510
Log Likelihood	-180,26	-248,525	-227,233	-304,392
Akaike Inf. Crit.	400,52	537,05	494,466	648,785

Los niveles de significancia son: ***p<0,01, **p<0,05, *p<0,1

Notas: En el grupo de bajo consumo de agua se consideran hogares tratados y controles con un consumo menor a 3 m³ per cápita. En el grupo de alto consumo de agua, se consideran hogares controles con un consumo menor a 3 m³ per cápita y hogares tratados con un consumo mayor o igual a 3 m³ per cápita. Entre paréntesis se muestran errores estándar robustos clusterizados a nivel distrital.

Fuente: Elaboración propia.

Tabla 13- Efectos marginales Modelo Logit con efectos heterogéneos (alto consumo)

Distancia	AME	SE	z	p	lower	upper
0,5 km	-0,1493	0,0618	-2,4160	0,0157	-0,2704	-0,0282
1,5 km	-0,1016	0,0538	-1,8871	0,0592	-0,2071	0,0039

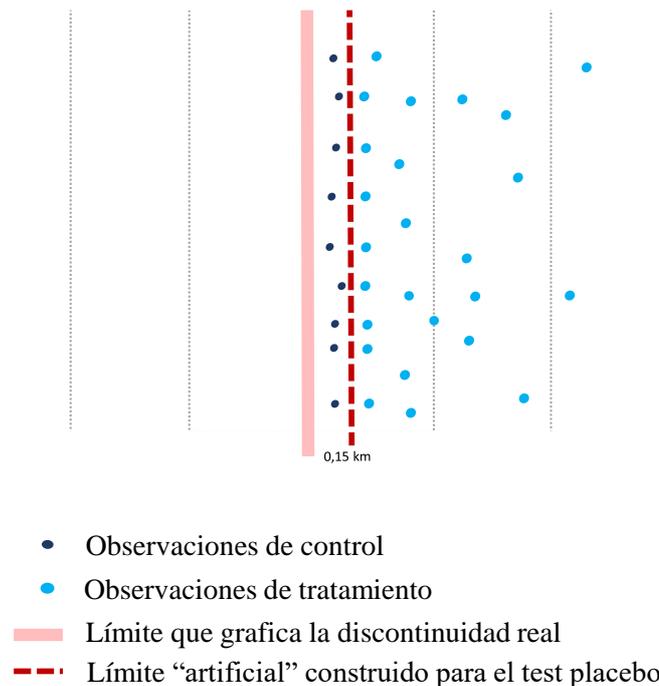
Fuente: Elaboración propia.

6.3. Test de falsificación

A fin de verificar la validez de las estimaciones, se realiza un test de falsificación o placebo. Este consiste en desplazar la frontera de acceso una distancia de 0,15 km. Para ello se

retiran los hogares de control (no conectados) de la muestra y se reclasifican como nuevas observaciones de control a aquellas observaciones de tratamiento que se encuentran a 0,15 km de la antigua frontera que delimitaba la discontinuidad en la prestación del servicio de agua. Con ello, se construye una nueva frontera “artificial” que divide a los nuevos hogares de control y tratamiento, trazada en el punto equidistante entre ambos tipos de hogares (ver **Figura 7**).

Figura 7 – Definición de la frontera artificial para test placebo



Fuente: Elaboración propia.

Al replicar la regresión considerando las nuevas observaciones de control y tratamiento, se obtiene que el tratamiento (acceso a agua) no tiene un coeficiente significativo (ver **Tabla 14**). Es decir, no se puede rechazar la hipótesis nula de que el coeficiente asociado al tratamiento “artificial” no es estadísticamente diferente de cero, lo cual apoya la validez de los resultados del presente trabajo.

Tabla 14- Resultados Logit Test Placebo (traslado de frontera)

	<0,5 km (1)	<1 km (2)	<1,5 km (3)	<2 km (4)
<i>Tratamiento</i>	0,06 (0,4100)	-0,02 (0,3350)	0,028 (0,3110)	0,005 (0,3100)
<i>Edad</i>	0,01 (0,0130)	0,009 (0,0110)	0,008 (0,0090)	0,008 (0,0090)
<i>Secundaria incompleta de la madre</i>	0,378 (0,4670)	-0,029 (0,3720)	-0,371 (0,3280)	-0,342 (0,3260)
<i>Trabajo presencial con atención al público</i>	0,096 (0,4540)	-0,17 (0,3510)	-0,337 (0,3090)	-0,313 (0,3070)
<i>Vivienda precaria</i>	0,447 (0,5290)	0,36 (0,4510)	0,479 (0,3920)	0,47 (0,3910)
<i>Hacinamiento</i>	-0,553 (0,4290)	-0,656* (0,3400)	-0,684** (0,3150)	-0,704** (0,3130)
<i>Ingresos mayores a 1000 soles</i>	1,868*** (0,4120)	1,300*** (0,3170)	1,009*** (0,2800)	0,973*** (0,2780)
<i>Primera ola</i>	0,268 (0,4380)	0,171 (0,3660)	0,229 (0,3400)	0,167 (0,3360)
<i>Comorbilidades</i>	0,28 (0,5140)	0,777** (0,3920)	0,742** (0,3650)	0,715** (0,3640)
<i>Bono</i>	-0,007 (0,4880)	-0,06 (0,3910)	-0,309 (0,3660)	-0,316 (0,3640)
<i>Constante</i>	-2,091** (0,8620)	-1,292* (0,6950)	-0,703 (0,6080)	-0,631 (0,6040)
Incluye variables de geolocalización	Sí	Sí	Sí	Sí
Observations	199	258	302	303
Log Likelihood	-94,859	-133,989	-163,821	-165,258
Akaike Inf. Crit.	229,719	307,978	367,642	370,516

Los niveles de significancia son: ***p<0,01, **p<0,05, *p<0,1

Nota: Entre paréntesis se muestran errores estándar robustos clusterizados a nivel distrital.

Fuente: Elaboración propia.

7. Consideraciones previas a la discusión de resultados

Para la elaboración de este documento, se diseñó un estudio cuasi-experimental que incluyó la recolección de información mediante encuestas a fin de hacer un estudio con el individuo como unidad de análisis. Esto nos permite afirmar que, metodológicamente, los resultados van más allá de un estudio ecológico y permiten aproximarse a la medición del efecto del acceso al agua en la infección. Sin embargo, consideramos importante desarrollar en esta sección algunas de las limitaciones, cómo estas han sido atendidas y sus implicaciones en la interpretación de los resultados.

Un primer grupo de limitaciones está relacionado a los sesgos que podrían introducirse debido a la información utilizada. El sesgo de información es una diferencia sistemática con respecto al fenómeno que se intenta medir que surge en la recopilación de la información. En nuestro estudio, todas las variables —incluyendo las preguntas relacionadas a la prueba diagnóstica y las comorbilidades— fueron recolectadas por auto reporte, por lo cual están influenciadas por la capacidad de recordar del encuestado o por una presión que lo lleva a responder en línea con lo que considera socialmente aceptable. Asimismo, las estimaciones realizadas consideran que una única prueba diagnóstica representa de manera válida el antecedente de infección por COVID-19; sin embargo, no puede descartarse que un individuo con una prueba diagnóstica con resultado negativo haya tenido una infección antes o después de la prueba.

Otro sesgo para considerar es el posible sesgo de selección, que consiste en un error sistemático por el cual los individuos incluidos en la muestra con que se estima el modelo difieren de los no incluidos, en especial en variables de importancia para el estudio. En el presente documento se han incluido tests de medias que señalan que no existieron diferencias significativas en el acceso a pruebas diagnósticas entre uno y otro lado de la frontera de prestación del servicio de agua y que tampoco se verifica que existan diferencias significativas entre los individuos que accedieron a una prueba diagnóstica y los que no accedieron a una. Al no existir indicios de un posible sesgo de selección, consideramos que la estimación con la metodología propuesta resulta adecuada. Sin embargo, incluso si la estimación se realizara corrigiendo por sesgo de selección (ver **Anexo 4**), aunque el efecto marginal resulta menor, continúa estando dentro del intervalo de confianza de los resultados primarios presentados en este estudio.

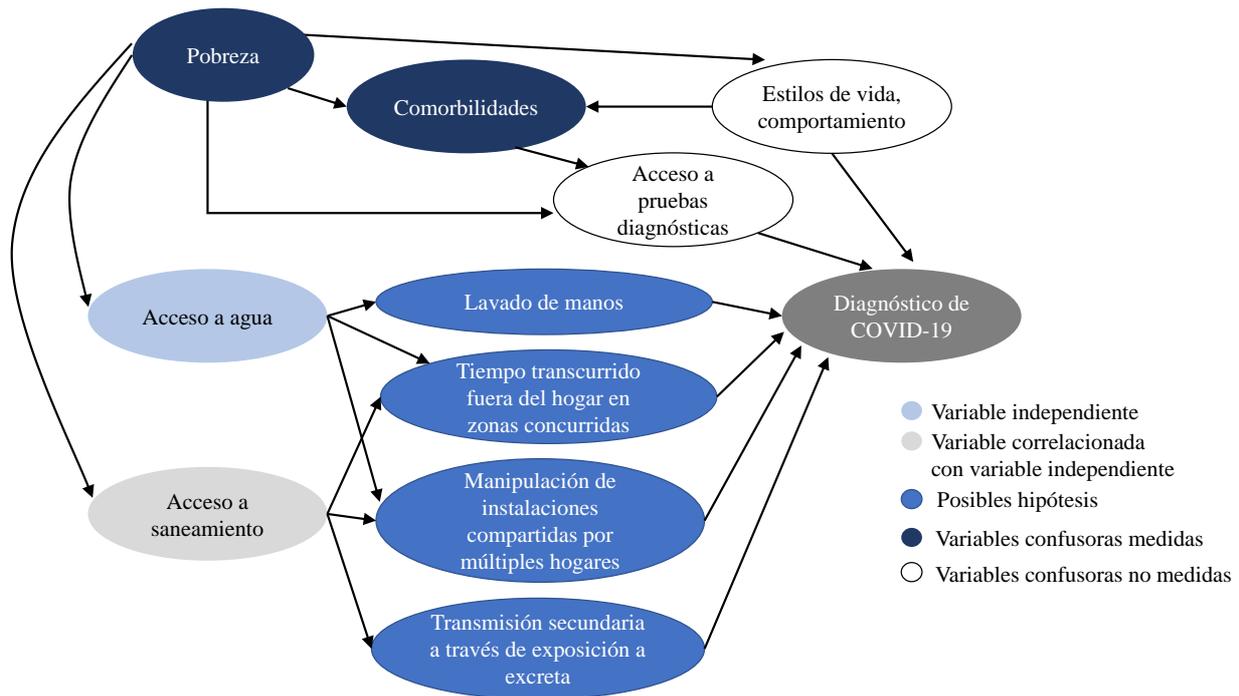
Un segundo grupo de limitaciones está relacionado al fenómeno de confusión y las implicancias para la inferencia de relaciones de causalidad. En primer lugar, el diseño de este estudio cuasi-experimental asumió *a priori* que las observaciones de control y tratamiento serían equivalentes para un grupo de variables confusoras, sobre todo en distancias muy cercanas a la frontera de discontinuidad en la prestación del servicio de agua. Ello con el fin de emular el fenómeno de aleatorización. Sin embargo, el trabajo de campo y el análisis posterior de los

resultados de tests de medias han permitido identificar diferencias en ciertas variables. A pesar de la proximidad física de las viviendas, por ejemplo, es posible que aquellas que no cuenten con acceso a agua por red no cuenten con título de propiedad y estén ubicadas en terrenos irregulares, lo que podría estar asociado a algunas de las diferencias identificadas. A pesar de la proximidad física, también se verifica, por ejemplo, diferencias en los ingresos. Esto ha sido parcialmente atendido mediante la inclusión de variables de control en la regresión.

Para explicar el fenómeno de confusión, resulta útil emplear un diagrama acíclico dirigido (DAG, por sus siglas en inglés). Esta es una representación visual utilizada frecuentemente en epidemiología, donde se plasman los supuestos de causalidad y que ayuda a identificar la presencia de confusión en las variables (Suttorp *et al.*, 2015).

Las características socioeconómicas (agrupadas bajo el concepto Pobreza en la **Figura 8**) son variables confusoras fundamentales pues tienen relación con el acceso al agua y la infección por COVID-19. Por ejemplo, las personas en condición de pobreza tienden a tener comorbilidades que se asocian con una mayor severidad del COVID-19, lo cual incrementa la probabilidad de buscar una prueba diagnóstica. Esto es atendido mediante la variable de comorbilidades en el modelo de estimación. Las personas en condición de pobreza también pueden tener estilos de vida o comportamientos que incrementan su riesgo de infección. Por ejemplo, pueden tener hábitos alimenticios que contribuyen a la existencia de comorbilidades o pueden tener comportamientos individuales que aumentan la transmisión del virus como el incumplimiento de las medidas de inmovilidad social para buscar ingresos. Estos pueden verse potenciados por factores de riesgo estructurales como el mayor hacinamiento. Este asunto es atendido en el presente estudio mediante la inclusión de variables socioeconómicas como el ingreso o el hacinamiento en el hogar.

Figura 8 – Diagrama acíclico dirigido



Fuente: Elaboración propia.

Con respecto a la inferencia de relaciones de causalidad, si bien hemos encontrado una asociación estadísticamente significativa entre la falta de acceso al agua y la infección por COVID, es necesario mencionar que existen diversas hipótesis que pueden explicar este resultado. Una primera está relacionada al lavado de manos, en línea con las recomendaciones de las autoridades sanitarias; sin embargo, se considera menos plausible en línea con la evidencia epidemiológica que señala que la transmisión aérea es el principal tipo de transmisión. Una segunda hipótesis es que los hogares no conectados a la red pública de agua deben incurrir en interacciones sociales más frecuentes y cercanas. Como se ha desarrollado previamente, para obtener el agua, las personas a menudo deben salir de sus casas, formar colas o aglomerarse alrededor de un camión cisterna, lo cual, incrementa el riesgo de infección por COVID-19. En tercer lugar, en el presente estudio, dentro de los hogares de control se encuentran aquellos que recurren a pilones de uso público, un tipo de abastecimiento que incrementa el tiempo transcurrido fuera de las viviendas en zonas concurridas e involucra la manipulación de instalaciones compartidas por múltiples hogares, lo cual puede aumentar el riesgo de transmisión del virus.

Otro aspecto a tener en cuenta es que el acceso al agua está fuertemente correlacionado con el acceso al saneamiento. Es decir, las personas que carecen de acceso al agua utilizan instalaciones de saneamiento compartidas con otros hogares y/o que no les permiten una separación higiénica de las excretas. El virus SARS-CoV-2 ha sido detectado en las excretas de pacientes (Amahmid *et al.*, 2021) y se ha documentado su persistencia por días en aguas

residuales (Godini *et al.*, 2021), lo cual podría crear una vía secundaria de transmisión (Khumar Takur *et al.*, 2021). Si bien se ha considerado que la posibilidad de transmisión fecal/de orina – oral/ocular es baja, los riesgos podrían incrementarse en el contacto persona a persona y en países en desarrollo con saneamiento pobre (Jones *et al.*, 2020). En ese sentido, debido a que no se puede incluir el acceso al saneamiento en la regresión a fin de evitar problemas de multicolinealidad, tampoco podemos descartar que el acceso al agua esté recogiendo los efectos anteriormente mencionados.

8. Discusión de resultados

El presente documento de trabajo contribuye al acervo de literatura existente que identifica la relación entre acceso a servicios de agua y saneamiento e indicadores de salud. En particular, los resultados primarios de nuestro estudio indican que el acceso al agua mediante red se asoció a una reducción de entre 6,2 y 9,9 puntos porcentuales en la probabilidad de registrar antecedente de infección por COVID-19. Considerando el nivel de positividad de los usuarios no conectados, esto significaría que la probabilidad de infectarse de COVID-19 se reduce entre 15% y 25% cuando el hogar cuenta con acceso al agua mediante red.

Asimismo, los resultados también sugieren que no basta con que un hogar esté conectado, sino que además debe garantizarse una dotación mínima de consumo. El documento encuentra que el acceso al agua se asocia significativamente a una reducción en la probabilidad de registrar antecedente de infección por COVID-19, siempre que se garantice una dotación mínima de 3 metros cúbicos de agua *per cápita* mensuales (o 150 litros por habitante por día). Si bien esto podría apoyar la hipótesis de que la calidad en el servicio es importante para cumplir con las recomendaciones de higiene, también podría significar que los hogares conectados a la red pero con poca calidad deben complementar su provisión con agua proveniente de camiones cisterna o pilones de uso público, lo cual los expone a un mayor riesgo de infección por COVID-19.

Consideramos que los resultados contribuyen a la hipótesis de que la falta de acceso a servicios básicos debe considerarse como un problema que afecta la salud pública. Y que más allá del acceso a la infraestructura física, el abastecimiento suficiente y asequible también son determinantes para la salud pública. Por tanto, ponen de relevancia la necesidad de que tanto Perú como los demás países de América Latina y el Caribe inviertan en infraestructura para el cierre de brechas de acceso e implementen políticas para asegurar un consumo mínimo del servicio. La falta de acceso a agua mediante red reduce la calidad de vida de la población, pues además de obligarles a recurrir a fuentes de escasa calidad y reducir su dotación de agua, generan condiciones que facilitan la transmisión de enfermedades, particularmente en el contexto de la pandemia por COVID-19.

9. Recomendaciones de política

Los trabajos especializados señalan que la alta mortalidad por COVID-19 en Perú es parcialmente atribuible a factores estructurales, tales como un sistema de salud débil y los altos niveles de pobreza y carencia de servicios básicos (Gianella *et al.*, 2020; Schwalb y Seas, 2020). En el presente estudio, identificamos que la falta de acceso al servicio de agua se encuentra asociada con una mayor probabilidad de registrar antecedente de infección por COVID-19. Al evaluar las formas alternativas de provisión de agua a las cuales deben recurrir los hogares no conectados a la red, se han podido identificar distintas situaciones y conductas que incrementaron el riesgo de transmisión del COVID-19 para estos hogares. Estos resultados ponen en evidencia la importancia de incrementar la inversión en el sector de agua y saneamiento para lograr una cobertura universal de estos servicios, con un estándar gestionado de forma segura, conforme a lo establecido por los ODS. Esta recomendación es aplicable para Perú, pero también para el resto de América Latina y el Caribe, donde aún persiste una importante brecha de acceso a servicios públicos (Brichetti *et al.*, 2021).

La segunda recomendación de política se desprende de los resultados del modelo de efectos heterogéneos. Si bien es importante brindar acceso universal, también es importante garantizar una dotación mínima de agua a los hogares. El consumo de agua de un hogar puede verse dificultado o impedido por un nivel deficiente en la calidad del servicio³⁴ o por tarifas altas. La falta de una dotación mínima puede comprometer los hábitos de higiene de los hogares u obligarlos a recurrir a otras formas de acceso al agua que, en el contexto de la pandemia, implicaron una mayor exposición a la infección por COVID-19. Por tal motivo, es importante que los países realicen las inversiones e implementen las medidas de gestión operativa necesarias para mejorar la continuidad y presión del servicio de agua. Del mismo modo, deben implementar políticas destinadas a garantizar la asequibilidad del servicio para los hogares en situación de vulnerabilidad económica.

³⁴ Al analizar la información recopilada en la encuesta sobre el nivel de satisfacción y continuidad (en horas), se observa que ambos se incrementan a medida que los usuarios conectados a la red pública de agua se encuentran más lejos de la discontinuidad geográfica en la prestación.

Referencias

Ahmad, K., Erqou, S., Shah, N., Nazir, U., Morrison, A. R., Choudhary, G., and Wu, W. C. (2020). *Association of poor housing conditions with COVID-19 incidence and mortality across US counties*. PLoS ONE, 15(11 November):1–13.

Amahmid, O., El Guamri, Y., Rakibi, Y., Ouizat, S., Yazidi, M., Razoki, B., Kaid Rassou, K., Asmama, S., Bouhoum, K. y D. Belghyti (2021). *Occurrence of SARS-CoV-2 in excreta, sewage, and environment: epidemiological significance and potential risks*. International Journal of Environmental Health Research. Marzo 2021.

Amankwaa, Godfred, and Christian Fischer (2020). *Exploring the Correlation between COVID-19 Fatalities and Poor WASH (Water, Sanitation and Hygiene) Services*. MedRxiv N° June.

Brichetti, J.P; Mastronardi, L; Rivas Amiassorho, M.E., Serebrisky, T. y B. Solís (2021). *La brecha de infraestructura en América Latina y el Caribe. Estimación de las necesidades de inversión para progresar hacia el cumplimiento de los Objetivos de Desarrollo Sostenible al 2030*. Banco Interamericano de Desarrollo. Monografía 962.

Brotherhood, L., Cavalcanti, T., Da Mata, D., and Santos, C. (2020). *Slums and Pandemics*. Cambridge Working Papers in Economics: 2076.

Brown, Patrick, Greenwald, Zoë R., Salinas, Luis Ernesto, Aguirre Martens, Gabriel, Newcombe, Leslie, Gelband, Hellen, Veillard, Jeremy y Prabhat Jha (2020). *Mortality from COVID in Colombia and Peru: Analyses of Mortality Data and Statistical Forecasts*. COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv.

Cárdenas, Helena (2022). *Review on Urban Water & Sanitation on Three Topics: Subsidy Incidence, Price Elasticities, and Relations to Health: Framework and Research Paths for Future Studies*. Banco Interamericano de Desarrollo. Nota técnica 2407.

Cejudo, Guillermo M., De los Cobos, Pablo, Michel, Cynthia L. y Diana L. Ramírez (2021). *Inventario y caracterización de los programas de apoyo al ingreso en América Latina y el Caribe frente a COVID-19*. Banco Interamericano de Desarrollo. Nota técnica 2334.

Celik, Ismail, Lisa Gallicchio, Kristina Boyd, Tram K. Lam, Genevieve Matanoski, Xuguang Tao, Meredith Shiels, et al. (2008). *Arsenic in Drinking Water and Lung Cancer: A Systematic Review*. Environmental Research 108 (1): 48–55.

Chauvin, Juan Pablo (2021). *Why Does COVID-19 Affect Some Cities More than Others?: Evidence from the First Year of the Pandemic in Brazil*. Banco Interamericano de Desarrollo. Documento de trabajo 1251.

Chauvin, Juan Pablo, Fowler, Annabelle y Nicolás Herrera L (2020). *The Younger Age Profile of COVID-19 Deaths in Developing Countries*. Banco Interamericano de Desarrollo. Documento de trabajo 1154.

Clouston, Sean A.P., Ginny Natale, and Bruce G. Link (2021). *Socioeconomic Inequalities in the Spread of Coronavirus-19 in the United States: A Examination of the Emergence of Social Inequalities.* Social Science and Medicine 268.

Darvesh, Nazia, Jai K. Das, Tyler Vaivada, Michelle F. Gaffey, Kumanan Rasanathan, and Zulfiqar A. Bhutta (2017). *Water, Sanitation and Hygiene Interventions for Acute Childhood Diarrhea: A Systematic Review to Provide Estimates for the Lives Saved Tool*. BMC Public Health 17 (Suppl 4).

- Das, A., Ghosh, S., Das, K., Basu, T., Das, M., & Dutta, I. (2020). *Modeling the effect of area deprivation on COVID-19 incidences: a study of Chennai megacity, India*. *Public Health*, 185, 266-269.
- Dasgupta, S., Bowen, V., Leidner, A., Fletcher, K., Musial, T., Rose, C., et al. (2020). *Association Between Social Vulnerability and a County's Risk for Becoming a COVID-19 Hotspot — United States, June 1–July 25, 2020*. US Department of Health and Human Services/Centers for Disease Control and Prevention. *Morbidity and Mortality Weekly Report (MMWR)*. October 23, 2020 / Vol. 69 / No. 42.
- De Groot, Jiska y Charlotte Lemanski (2021). *COVID-19 Responses: Infrastructure Inequality and Privileged Capacity to Transform Everyday Life in South Africa*. *Environment and Urbanization* 33 (1): 255–72.
- Dell, M (2010). *The Persistent Effects of Peru's Mining Mita*. *Econometrica*, 78, no. 6, pp. 1863-1903.
- Desmet, K. and Wacziarg, R. (2021). *JUE insight: Understanding Spatial Variation in COVID-19 across the United States*. *Journal of Urban Economics*.
- Díaz-Cassou, Javier, Deza, María Cecilia y Kenji Moreno (2020). *Perú: Desafíos del Desarrollo en el post COVID-19*. Banco Interamericano de Desarrollo. Documento para Discusión 790.
- Elacqua, G., López, K., Méndez, C., Novella, R., Prada, M. F. y D. Vásquez. *El impacto de la crisis del COVID-19 en estudiantes vulnerables de educación superior: El caso de Perú*. Nota Técnica del BID 2449.
- Fortaleza, C., Guimarães, R., De Almeida, G., Pronunciate, M., & Ferreira, C. (2020). *Taking the inner route: spatial and demographic factors affecting vulnerability to COVID-19 among 604 cities from inner São Paulo State, Brazil*. *Epidemiology and Infection*, 148, 1-5.
- Freire de Souza, Carlos Dornels, Michael Ferreira Machado, and Rodrigo Feliciano Do Carmo. (2020). *Human Development, Social Vulnerability and COVID-19 in Brazil: A Study of the Social Determinants of Health*. *Infectious Diseases of Poverty* 9 (1): 1–11.
- Friedrichs, M., Wagstaff, S., Sage, K., LaCross, N., Bui, D., et al. (2020). *Disparities in COVID-19 Incidence, Hospitalizations, and Testing, by Area-Level Deprivation — Utah, March 3–July 9, 2020*. *Morbidity and Mortality Weekly Report*, 69(38), 1369-1373
- Gianella, C., Iguñiz-Romero, R., Romero, M.J., y J. Gideon (2020). *Good Health Indicators are Not Enough: Lessons from COVID-19 in Peru*. *Health and Human Rights Journal*, 2020. 22(2): p. 317-319
- Godini, H., Hoseinzadeh, E. y H. Hossini. *Water and wastewater as potential sources of SARS-CoV-2 transmission: a systematic review*. *Reviews on Environmental Health*, vol. 36, no. 3, 2021, pp. 309-317.
- Greene, W. (2006). *A General Approach to Incorporating Selectivity in a Model*. Department of Economics, Stern School of Business, New York University.
- Harris, Jeffrey (2020). *The subways seeded the massive Coronavirus epidemic in New York City*. NBER Working Paper Series.
- Hawkins, D. (2020). *Social Determinants of COVID-19 in Massachusetts, United States: An Ecological Study*. *Journal of Preventive Medicine & Public Health*, 53(4), 220-227.

Heckman, J. (1979). *Sample Selection Bias as a Specification Error*. *Econometrica*, 47, pp. 153-161.

Howard, Guy, Bartram, Jamie *et al.* (2020). *COVID-19: Urgent Actions, Critical Reflections and Future Relevance of 'WaSH': Lessons for the Current and Future Pandemics*. *Journal of Water and Health* 18 (5): 613–30.

Hyde, K. (2020). *Residential Water Quality and the Spread of COVID-19 in the United States*. Social Science Research Network Web site.

Instituto Nacional de Estadística e Informática - INEI (2021). *Perú: Línea Base de los Principales Indicadores Disponibles de los Objetivos de Desarrollo Sostenible ODS 2021*.

Jaramillo, Miguel y Kristian López (2021). *Políticas para combatir la pandemia de COVID-19*. Grupo de Análisis para el Desarrollo - GRADE. Documentos de investigación 112.

Jones, D.L., Baluja, M.Q., Graham, D.W., Corbishley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H. y K. Farkas. (2020). *Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19*. *Science of Total Environment* 749.

Kamis, Christina, Allison Stolte, Jessica S. West, Samuel H. Fishman, Taylor Brown, Tyson Brown, and Heather R. Farmer (2021). *Overcrowding and COVID-19 Mortality across U.S. Counties: Are Disparities Growing over Time?* *SSM - Population Health* 15: 100845.

Keele, Luke y Rocío Titiunik (2015). *Geographic Boundaries as Regression Discontinuities*. *Political Analysis* (2015) 23:127–155.

Kumar Thakur, A., Sathyamurthy, R., Velraj, R., Lynch, I., Saidur, R., Pandey, A.K., Sharshir, S.W., Elnaby Kabeel, A., Hwang, J., y P. GaneshKumar, (2021) *Secondary transmission of SARS-CoV-2 through wastewater: Concerns and tactics for treatment to effectively control the pandemic*. *Journal of Environmental Management*, Volume 290, 2021.

Lee, David S. y Thomas Lemieux. *Regression Discontinuity Designs in Economics*. *Journal of Economic Literature*. Vol. 48, No. 2, June 2010 (pp. 281-355)

Levin, A. T., Cochran, K. B., and Walsh, S. P. (2020). *Assessing the age specificity of infection fatality rates for COVID-19: Meta-analysis & public policy implications*. NBER Working Paper 27597.

Marí-Dell'Olmo, M., Gotsens, M., Pasarín, I., Rodríguez-Sanz, M., Artazcoz, L., García de Olalla, P. *et al.* (2021). *Socioeconomic Inequalities in COVID-19 in a European Urban Area: Two Waves, Two Patterns*. *International Journal of Environmental Research and Public Health*, 18(3), 1256.

Ministerio de Economía y Finanzas - MEF (2019). *Plan Nacional de Infraestructura para la Competitividad*.

Ministerio de Vivienda, Construcción y Saneamiento (2021). *Plan Nacional de Saneamiento 2022-2026*.

Moscoe, Ellen, Bor, Jacob y Till Bärnighausen. *Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice*. *Journal of Clinical Epidemiology* 68 (2015) 132 – 143.

- Nicodemo, C., Barzin, S., Cavalli, N., Lasserson, D., Moscone, F., Redding, S., & Shaikh, M. (2020). *Measuring geographical disparities in England at the time of COVID-19: results using a composite indicator of population vulnerability*. *BMJ Open*, 10(9), 1-8.
- Niedzwiedz, C., O'Donnell, C., Dinesh, B., Demou, E., Ho, F., Celis-Morales, C., *et al.* (2020). *Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank*. *BMC Medicine*, 18(160), 1-14.
- Organización Mundial de la Salud (2017). *Guías para la calidad del agua de consumo humano: Cuarta edición que incorpora la primera adenda*.
- Organización Mundial de la Salud (2020). *Handwashing an effective tool to prevent COVID-19, other diseases*. <https://www.who.int/southeastasia/news/detail/15-10-2020-handwashing-an-effective-tool-to-prevent-covid-19-other-diseases> [consultado el 3 de febrero de 2022]
- Organización Mundial de la Salud (2022). *Tracking SARS-CoV-2 variants*. <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> [consultado el 9 de febrero de 2022]
- Revollo-Fernández, D., Rodríguez-Tapia, L., Medina-Rivas, C. y J.A. Morales-Novelo (2022). *Socioeconomic determinants of COVID-19 in Mexico*. *Public Health*.
- Romero, P., Dávila-Barclay, A., Salvatierra, S., González, L., Cuicapuza, D., Solís, L., Marcos-Carbajal, P., Huancachoque, K., Maturrano, L., y P. Tsukayama (2021). *The Emergence of Sars-CoV-2 Variant Lambda (C.37) in South America*. *Microbiology Spectrum*. Volume 9 Issue 2.
- Rozenfeld, Y., Beam, J., Maier, H., Haggerson, W., Boudreaud, K., Carlson, J., y Medows, R. (2020). *A model of disparities: risk factors associated with COVID-19 infection*. *International Journal for Equity in Health*.
- Scarpone, C., Brinkmann, S., Große, T., Sonnenwald, D., Fuchs, M., y Byron, B. (2020). *A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany*. *International Journal of Health Geographics*, 19(32), 1-17.
- Schwalb, A. y C. Seas (2021). *The COVID-19 Pandemic in Peru: What Went Wrong? The American Journal of Tropical Medicine and Hygiene*. 2021 Apr; 104(4): 1176–1178
- Sheng, J., Malani, A., Goel, A., y Botla, P. (2021). *JUE Insight: Does Mobility Explain Why Slums Were Hit Harder by COVID-19 in Mumbai, India?* *Journal of Urban Economics*, page 103357.
- Silva, R., Ribeiro, C., Moura, T., Santos, M., Santos, A., Tavares, D., & Santos, P. (2021). *Basic sanitation: a new indicator for the spread of COVID-19?* *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 15(7), 832-840.
- Stampini, Marco, Ibararán, Pablo, Rivas, Carolina y Marcos Robles (2021). *Adaptive, but Not by Design: Cash Transfers in Latin America and the Caribbean Before, During and After the COVID-19 Pandemic*. Banco Interamericano de Desarrollo. Nota técnica 2346.
- Stokes, Andrew C., Dielle J. Lundberg, Irma T. Elo, Katherine Hempstead, Jacob Bor, and Samuel H. Preston (2021). *COVID-19 and Excess Mortality in the United States: A County-Level Analysis*. *PLoS Medicine* 18 (5): 1–19.

- Strully, K., Yang, T.-C., & Liu, H. (2021). *Regional variation in COVID-19 disparities: connections with immigrant and Latinx communities in U.S. counties*. *Annals of Epidemiology*, 53, 56-62.
- Sunass - Superintendencia Nacional de Servicios de Saneamiento (2015). *Comprar agua por camión cisterna les cuesta a las familias limeñas 72 soles mensuales*. Nota de Prensa 161 – 2015.
- Suttorp, M, Siegerink, B., Jager, K.J., Zoccali, C. y F. W. Dekker (2015). *Graphical presentation of confounding in directed acyclic graphs*. *Nephrol Dial Transplant* (2015) 30: 1418-1423.
- Taylor, Dawn L., Tanya M. Kahawita, Sandy Cairncross, and Jeroen H.J. Ensink. 2015. *The Impact of Water, Sanitation and Hygiene Interventions to Control Cholera: A Systematic Review*. *PLoS ONE* 10 (8): 1–19.
- The Economist (2022). *The pandemic's true death toll. Our daily estimate of excess deaths around the world*. <https://www-economist-com.access.idm.oclc.org/graphic-detail/coronavirus-excess-deaths-estimates> [consultado el 3 de febrero de 2022]
- Vargas-Herrera, N., Araujo-Castillo, R.V., Mestanza, O., Galarza, M., Rojas-Serrano, N. y L. Solari-Zerpa (2022). *SARS-CoV-2 Lambda and Gamma variants competition in Peru, a country with high seroprevalence*. *The Lancet Regional Health – Americas*, Volumen 6, February 2022.
- Vázquez-Rowe, I. y A. Gandolfi (2020). *Peruvian efforts to contain COVID-19 fail to protect vulnerable population groups*. *Public Health in Practice* 1 - 2020.
- Waddington, Hugh y Birte Snilstveit (2009). *Effectiveness and Sustainability of Water, Sanitation, and Hygiene Interventions in Combating Diarrhoea.* *Journal of Development Effectiveness* 1 (3): 295–335.
- Wang, C.; Prather, K.; Sznitman, J.; Jimenez, J.; Lakdawala, S.; Tufekci, Z. y L. Marr. (2021). *Airborne transmission of respiratory viruses*. *Science* 373, 981.
- Wolf, Jennyfer, Paul R. Hunter, Matthew C. Freeman, Oliver Cumming, Thomas Clasen, Jamie Bartram, Julian P.T. Higgins, *et al.* (2018) *“Impact of Drinking Water, Sanitation and Handwashing with Soap on Childhood Diarrhoeal Disease: Updated Meta-Analysis and Meta-Regression.”* *Tropical Medicine and International Health* 23 (5): 508–25.
- Xu, Jay J, Jarvis T Chen, Thomas R Belin, Ronald S Brookmeyer, Marc A Suchard, Christina M Ramirez, y Jay J Xu (2021). *National Population-Level Disparities in COVID-19 Mortality Across the Intersection of Race / Ethnicity and Sex in the United States*. *MedRxiv*, 1–18.
- You, H., Wu, X., & Guo, X. (2020). *Distribution of COVID-19 Morbidity Rate in Association with Social and Economic Factors in Wuhan, China: Implications for Urban Development*. *International Journal of Environmental Research and Public Health*, 17(10), 1-14.
- Zhang, H., Liu, Y., Chen, F., Baibing, M., Zeng, L., & Pei, L. (2021). *The effect of sociodemographic factors on COVID-19 incidence of 342 cities in China: a geographically weighted regression model analysis*. *BMC Infectious Diseases*, 21(428), 1-8

Anexo 1 – Ficha técnica de la encuesta

a) Población objetivo

La población de estudio está definida como el conjunto de todos los hogares con y sin acceso al servicio de agua potable ubicados en las zonas periurbanas de los distritos de Ancón, Carabaylo, Lurigancho, Cieneguilla, Pachacamac, Lurín, San Juan de Lurigancho y Ventanilla, ubicados a una distancia de hasta 2km de la discontinuidad geográfica en la prestación del servicio.

Los distritos fueron seleccionados utilizando los resultados de los Censos Nacionales XII de Población, VII de Vivienda y III de Comunidades Indígenas realizados en el año 2017. Se seleccionaron los ocho distritos con menor acceso al servicio de agua potable, exceptuando los distritos en los balnearios del sur de Lima Metropolitana bajo el ámbito del proyecto Provisur. Esto debido a que el nivel de cobertura puede haber sufrido variaciones importantes debido al proyecto y a que la estacionalidad en la ocupación de las viviendas en estos distritos podría dificultar el trabajo de campo.

b) Distritos de análisis y tamaño total de muestra

Provincias	Distritos	Acceso al agua mediante red pública		Total general
		Sin acceso	Con acceso	
Lima	Ancón	30	10	40
	Carabaylo	80	61	141
	Cieneguilla	20	10	30
	Lurigancho	136	80	216
	Lurín	57	30	87
	Pachacamac	86	60	146
	San Juan de Lurigancho	174	181	355
Callao	Ventanilla	51	55	106
Total		634	487	1.121

Nota: Incluye algunas observaciones cuya distancia se ha verificado mayor a los 2km de la discontinuidad geográfica en la prestación del servicio y que no son consideradas en el estudio.

c) Margen de error y confiabilidad:

	Sin acceso	Con acceso
Margen de error	+/- 4,2%	+/- 4.3%
Confiabilidad	95%	95%

d) **Fechas de trabajo de campo:** El trabajo de campo fue realizado por la empresa Impacto Directo / Directo Marketing del 12 de octubre al 5 de noviembre de 2021.

- e) **Muestreo:** El diseño de regresión discontinua del estudio requería del muestreo de hogares alrededor de una discontinuidad geográfica. Ésta ha sido definida como una frontera entre aquellos hogares que cuentan con acceso a agua mediante red y aquellos que no cuentan con este acceso. Es decir, no se requiere que la muestra sea representativa de la totalidad de los hogares en los distritos estudiados. Por el contrario, la muestra debe permitir la adecuada comparación entre hogares a ambos lados de la frontera.

El marco muestral se diseñó sobre la base de los resultados de los Censos Nacionales XII de Población, VII de Vivienda y III de Comunidades Indígenas realizados en el año 2017. Con ayuda de los planos, se identificaron agrupaciones de manzanas donde existían discontinuidades en la prestación del servicio de agua potable. Debido a la antigüedad de esta información, se consideró que la frontera podría haberse desplazado (la cobertura de agua potable podría haberse incrementado). Por tal motivo, mediante trabajo de campo se validó y actualizó el marco muestral para los usuarios no conectados a la red pública.

Para la muestra de hogares sin conexión, se realizó un muestreo probabilístico, multietápico, estratificado y de selección final sistemática. Para la muestra de hogares con conexión, se utilizó el criterio de proximidad a la muestra de hogares sin conexión, registrando en ambos casos las coordenadas de geolocalización del hogar.

Procedimiento de selección de hogares:

Etapas	Unidad de muestreo	Método de selección
1	Centro Poblado	Selección proporcional al tamaño
2	Manzanas (urbano) y segmentos compactos (rural)	Selección proporcional al tamaño
3	Viviendas particulares	Selección sistemática de arranque aleatorio

- La Unidad Primaria de Muestreo está conformada por los centros poblados de cada distrito denominado conglomerado, dentro de los cuales se encuentran las manzanas.
- La Unidad Secundaria de Muestreo (USM) es el área periurbana del centro poblado que está conformado por las manzanas.
- La Unidad Terciaria de Muestreo (UTM) corresponde a la vivienda particular.

Anexo 2 – Test de medias para la proporción de individuos con prueba COVID-19

Variable	0,5 km de la frontera			1 km de la frontera			1,5 km de la frontera		
	Control (1)	Trat. (2)	P value t-test (1) = (2)	Control (1)	Trat. (2)	P value t-test (1) = (2)	Control (1)	Trat. (2)	P value t-test (1) = (2)
Proporción de individuos que se realizaron una prueba diagnóstica	0,259 (0,017)	0,300 (0,017)	0,191	0,270 (0,015)	0,292 (0,019)	0,456	0,269 (0,014)	0,290 (0,018)	0,432

Nota: Entre paréntesis se muestran errores estándar clusterizados por distrito

Anexo 3 – Resultados del test de medias (testeados vs no testeados)

	0,5 km de la frontera			1 km de la frontera			1,5 km de la frontera			2 km de la frontera		
	SIN PRUEBA (1)	CON PRUEBA (2)	P val for t-test (1) = (2)	SIN PRUEBA (1)	CON PRUEBA (2)	P val for t-test (1) = (2)	SIN PRUEBA (1)	CON PRUEBA (2)	P val for t-test (1) = (2)	SIN PRUEBA (1)	CON PRUEBA (2)	P val for t-test (1) = (2)
Ratio de hacinamiento	2,107 (0,126)	2,311 (0,065)	0,221	2,088 (0,128)	2,311 (0,057)	0,214	2,082 (0,131)	2,314 (0,054)	0,191	2,072 (0,131)	2,341 (0,045)	0,132
Vivienda precaria	0,475 (0,037)	0,423 (0,038)	0,141	0,477 (0,039)	0,444 (0,040)	0,275	0,476 (0,038)	0,445 (0,038)	0,279	0,483 (0,033)	0,449 (0,033)	0,211
Bono Midis	0,228 (0,023)	0,227 (0,041)	0,992	0,224 (0,021)	0,227 (0,034)	0,956	0,224 (0,021)	0,230 (0,037)	0,909	0,220 (0,021)	0,235 (0,039)	0,809
Bono Universal	0,185 (0,036)	0,197 (0,016)	0,754	0,178 (0,038)	0,189 (0,017)	0,725	0,177 (0,038)	0,189 (0,017)	0,700	0,174 (0,037)	0,187 (0,020)	0,621
Total de bonos recibidos	0,497 (0,048)	0,510 (0,041)	0,816	0,487 (0,051)	0,496 (0,045)	0,886	0,485 (0,049)	0,501 (0,049)	0,782	0,475 (0,049)	0,500 (0,048)	0,655
Presencia de comorbilidades	0,262 (0,028)	0,357 (0,035)	0,141	0,248 (0,023)	0,347 (0,033)	0,081	0,243 (0,024)	0,343 (0,031)	0,068	0,245 (0,023)	0,339 (0,030)	0,071
Trabajo presencial con atención al público	0,690 (0,048)	0,613 (0,036)	0,094	0,688 (0,057)	0,602 (0,029)	0,090	0,692 (0,055)	0,606 (0,024)	0,098	0,693 (0,053)	0,612 (0,023)	0,101
Ingreso menor a 750	0,334 (0,059)	0,314 (0,033)	0,513	0,320 (0,050)	0,325 (0,031)	0,843	0,329 (0,051)	0,325 (0,029)	0,890	0,340 (0,051)	0,330 (0,028)	0,727
Ingreso entre 750 a 1000	0,315 (0,054)	0,317 (0,051)	0,974	0,317 (0,057)	0,304 (0,049)	0,777	0,311 (0,051)	0,311 (0,050)	0,999	0,307 (0,051)	0,311 (0,047)	0,941
Ingreso mayor a 1000	0,351 (0,090)	0,369 (0,077)	0,671	0,362 (0,083)	0,371 (0,069)	0,850	0,360 (0,083)	0,365 (0,068)	0,928	0,352 (0,079)	0,359 (0,065)	0,871
Miembros del hogar menores a 17 años	1,152 (0,257)	1,323 (0,080)	0,513	1,102 (0,240)	1,347 (0,081)	0,309	1,106 (0,230)	1,361 (0,081)	0,262	1,088 (0,230)	1,381 (0,082)	0,189
Miembros del hogar mayores a 65 años	0,145 (0,026)	0,246 (0,036)	0,085	0,139 (0,023)	0,244 (0,032)	0,058	0,157 (0,026)	0,241 (0,032)	0,122	0,152 (0,027)	0,242 (0,032)	0,118
Madre con estudios secundarios como mínimo	0,661 (0,056)	0,684 (0,032)	0,741	0,654 (0,054)	0,688 (0,031)	0,643	0,656 (0,055)	0,680 (0,033)	0,760	0,661 (0,053)	0,667 (0,032)	0,927
Acceso a los servicios de agua potable	0,486 (0,061)	0,498 (0,040)	0,781	0,466 (0,073)	0,463 (0,052)	0,943	0,463 (0,072)	0,459 (0,054)	0,918	0,449 (0,065)	0,443 (0,052)	0,857

Nota: Entre paréntesis se muestran errores estándar clusterizados por distrito

Anexo 4 – Modelo con corrección por posible sesgo de selección

Conforme a la metodología propuesta por Greene (2006) para ecuaciones principales no lineales, se estiman mediante modelos probit tanto la ecuación de selección (que en este caso estima la probabilidad de realizarse una prueba diagnóstica) y la ecuación principal (que en este caso estima la probabilidad de registrar antecedente de infección por COVID-19).

Dentro de la ecuación principal se incluyen, además del *tratamiento*, las siguientes variables: *hacinamiento*, *uso de transporte público por parte del jefe de hogar*, *ingreso mayor a 1000 soles* y la presencia de *comorbilidades*. Para la ecuación de selección, además de las variables contenidas en la ecuación principal, se incluyeron como instrumentos si el individuo presenta *síntomas*, la *edad* del individuo, si *trabaja de forma presencial y con atención al público* y si se trata de un *trabajador dependiente*. A continuación, se muestran los efectos marginales y los resultados de la estimación. Asimismo, debe mencionarse que la estimación incluye variables que pueden diferir de las incluidas en los resultados primarios de este estudio. Las variables incluidas fueron las aseguraban que las estimaciones convergieran, teniendo en cuenta que se realizan mediante la maximización de la función de verosimilitud por simulaciones.

Efectos Marginales de la variable de tratamiento según distancia

Distancia	AME	SE	z	p	lower	upper
0,5 km	-0,043	0,017	-2,490	0,013	-0,076	-0,009
1 km	-0,036	0,013	-2,800	0,005	-0,061	-0,011
1,5 km	-0,036	0,013	-2,790	0,005	-0,061	-0,011
2 km	-0,038	0,012	-3,140	0,002	-0,062	-0,014

Fuente: Elaboración propia.

Resultados de la estimación (Heckman Probit)

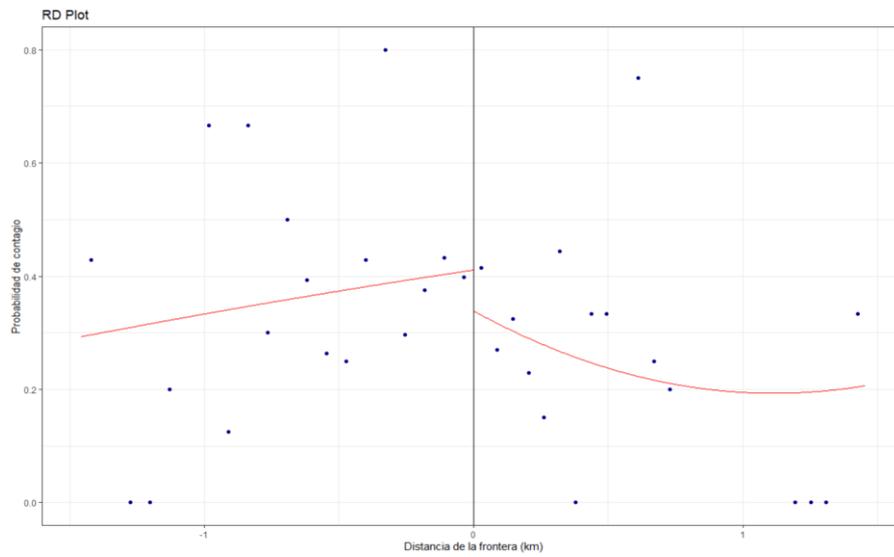
	Distancia de la frontera			
	<0,5 km	<1 km	<1,5 km	<2 km
Ecuación principal				
Variable dependiente: El resultado de la prueba fue positivo (sí=1)				
<i>Tratamiento</i>	-0,146*** (0,054)	-0,124*** (0,042)	-0,122*** (0,042)	-0,131*** (0,039)
<i>Hacinamiento</i>	0,0791 (0,051)	0,0781** (0,040)	0,0759** (0,038)	0,0674* (0,036)
<i>Uso de transporte público</i>	0,0873 (0,099)	0,0920 (0,086)	0,102 (0,082)	0,096 (0,081)
<i>Ingreso del hogar mayor a 1000</i>	-0,204* (0,122)	-0,117 (0,100)	-0,100 (0,096)	-0,098 (0,086)
<i>Comorbilidades</i>	0,0268 (0,094)	0,00923 (0,094)	0,030 (0,096)	0,012 (0,100)
<i>Constante</i>	0,727*** (0,130)	0,707*** (0,074)	0,688*** (0,072)	0,724*** (0,070)
Ecuación de Selección				
Variable dependiente: Prueba_covid				
<i>Individuo presenta síntomas</i>	1,196*** (0,107)	1,180*** (0,089)	1,181*** (0,084)	1,184*** (0,090)
<i>Edad del individuo</i>	0,00823** (0,003)	0,00767*** (0,003)	0,00798*** (0,002)	0,00841*** (0,003)
<i>Trabaja presencial con atención al público</i>	-0,0484*** (0,018)	-0,0745*** (0,028)	-0,0760*** (0,026)	-0,0748*** (0,026)
<i>Trabajo dependiente</i>	0,157** (0,063)	0,110*** (0,040)	0,111*** (0,037)	0,111** (0,045)
<i>Hacinamiento</i>	-0,104*** (0,027)	-0,103*** (0,032)	-0,101*** (0,031)	-0,0857*** (0,031)
<i>Utiliza el transporte público</i>	0,120 (0,090)	0,066 (0,079)	0,051 (0,074)	0,041 (0,073)
<i>Ingreso del hogar mayor a 1000</i>	0,179** (0,076)	0,0961* (0,054)	0,076 (0,056)	0,0903* (0,048)
<i>Comorbilidades</i>	0,062 (0,128)	0,062 (0,122)	0,040 (0,120)	0,022 (0,103)
<i>Constante</i>	-1,061*** (0,134)	-0,956*** (0,093)	-0,949*** (0,087)	-1,001*** (0,083)
<i>athrho</i>	-2,744*** (0,550)	-2,871*** (0,465)	-2,821*** (0,385)	-2,795*** (0,401)
Observaciones Seleccionadas	612	707	728	749
Observaciones No Seleccionadas	1991	1991	1991	1991
Total observaciones	2603	2698	2719	2740

Los niveles de significancia son: ***p<0,01, **p<0,05, *p<0,1

Nota: Entre paréntesis se muestran errores estándar robustos clusterizados a nivel distrital. Incluye variables de localización geográfica (altitud, latitud).

Fuente: Elaboración propia.

Anexo 5 – Probabilidad de infección por COVID-19 según distancia a la frontera



Nota: El panel de la izquierda corresponde a usuarios no conectados a la red de agua; mientras que el panel de la derecha corresponde a los usuarios conectados.

Fuente: Elaboración propia.