

# Seroprevalence of anti-SARS-CoV-2 antibodies in Colombia, 2020: A population-based study



Marcela Mercado-Reyes,<sup>a,\*</sup> Jeadran Malagón-Rojas,<sup>a</sup> Isabel Rodríguez-Barraquer,<sup>b</sup> Silvana Zapata-Bedoya,<sup>a</sup> Magdalena Wiesner,<sup>a</sup> Zulma Cucunubá,<sup>c</sup> Yesith Guillermo Toloza-Pérez,<sup>a</sup> Juan P. Hernández-Ortiz,<sup>d</sup> Jorge Acosta-Reyes,<sup>e</sup> Eliana Parra-Barrera,<sup>a</sup> Edgar Ibáñez-Beltrán,<sup>a</sup> Gianni G Quinche,<sup>a</sup> Lyda Muñoz-Galindo,<sup>a</sup> Vivian Rubio,<sup>a</sup> Marisol Galindo-Borda,<sup>a</sup> Erickson G Osorio-Velázquez,<sup>f</sup> Andrea Bermúdez-Forero,<sup>a</sup> Nelson Pinto-Chacón,<sup>a</sup> Gloria Puerto-Castro,<sup>a</sup> Carlos Franco-Muñoz,<sup>a</sup> María Isabel Estupiñán,<sup>g</sup> Luis Ángel Villar,<sup>g</sup> Nancy Gore-Saravia,<sup>h</sup> María Consuelo Miranda-Montoya,<sup>h</sup> Jaime Castellanos,<sup>i</sup> Edna Margarita Valle,<sup>f</sup> Edgar Navarro-Lechuga,<sup>e</sup> Juan Daniel Oviedo,<sup>f</sup> and Martha Ospina-Martínez<sup>a</sup>

<sup>a</sup>Instituto Nacional de Salud, Avenida Calle 26 51 -20, Bogotá, Colombia

<sup>b</sup>University of California, San Francisco, CA, USA

<sup>c</sup>Imperial College London, London, UK

<sup>d</sup>Universidad Nacional de Colombia, Medellín, Antioquia, Colombia

<sup>e</sup>School of Medicine, Universidad del Norte de Barranquilla, Barranquilla, Colombia

<sup>f</sup>Departamento Administrativo Nacional de Estadística – DANE, Bogotá, Colombia

<sup>g</sup>Centro de Atención y Diagnóstico de Enfermedades Infecciosas – CDI, Bucaramanga, Colombia

<sup>h</sup>Centro Internacional de Entrenamiento e Investigaciones Médicas – CIDEIM, Cali, Colombia

<sup>i</sup>Universidad El Bosque, Bogotá, Colombia

## Summary

**Background** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) continues to cause very high morbidity and mortality throughout Latin American countries. However, few population-based seroprevalence surveys have been conducted to quantify attack rates and characterize drivers of transmission.

**Methods** We conducted a population-based cross-sectional study to assess the seroprevalence of antibodies against SARS-CoV-2 in ten cities in Colombia between September and December 2020. The study involved multi-stage cluster sampling at each city. Participants provided a serum sample and answered a demographic and risk factor questionnaire. Prior infection by SARS-CoV-2 was ascertained using the "SARS-CoV-2 Total (COV2T) Advia Centaur - Siemens" chemiluminescence assay.

**Findings** A total of 17863 participants from 7320 households participated in the study. Seroprevalence varied substantially between cities, ranging from 26% (95%CI 23–29 %) in Medellín to 68% (95%CI 62–74 %) in Guapi. There were no differences in seroprevalence by sex, but seropositivity was higher in certain ethnic groups. There was substantial heterogeneity in seroprevalence within cities, driven to a large extent by a strong association between socioeconomic stratum and seropositivity.

**Interpretation** Colombia has been one of the Latin American countries most affected by the COVID-19 pandemic. This study documented very high attack rates in several Colombian cities by the end of 2020 and identified key drivers of heterogeneities including ethnicity and socioeconomic stratum. Few studies of seroprevalence of SARS-CoV-2 have been conducted in Latin America, and therefore this study contributes to the fundamental understanding of the pandemic in the region.

**Funding** The study was sponsored by, Ministerio de Ciencia y Tecnología e Innovación –CT361/2020, Ministerio de Salud y Protección Social, Fundación Universitaria del Norte, Imperial College of London, Universidad Nacional de Colombia (Sede Medellín), Universidad de Córdoba, California University, Unidad Nacional de Gestión del Riesgo, Centro de Atención y Diagnóstico de Enfermedades Infecciosas -CDI-, Centro Internacional de Entrenamiento e Investigaciones Médicas -CIDEIM-, Departamento Administrativo Nacional de Estadística - DANE, Fondo Nacional de Turismo -FONTUR-, Secretarías de Salud Departamentales, Distritales y Municipales and Instituto Nacional de Salud.

**Copyright** © 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The Lancet Regional Health - Americas 2022;9: 100195

Published online 5 February 2022

<https://doi.org/10.1016/j.lana.2022.100195>

\*Corresponding author.

E-mail address: [mmercado@ins.gov.co](mailto:mmercado@ins.gov.co) (M. Mercado-Reyes).

**Keywords:** SARS-CoV-2; Seroprevalence; COVID-19; Colombia; Public health; Population

### Research in context

#### *Evidence before this study*

Colombia is one of the most affected countries by COVID-19 in Latin America. Surveillance data suggests that the epidemic has been heterogeneous throughout the country, with early peaks in the North Coast and Amazonian border and later peaks elsewhere. Nevertheless, in the absence of seroprevalence studies, the true infection rate in Colombia remains unknown. We searched PubMed, Embase, and the Serotracker® Database for publications on the seroprevalence of SARS-CoV-2 antibodies, using the search terms “SARS-CoV-2”, “COVID-19”, “seroprevalence”, “Latin America” from September to December, 2020. Except for Brazil and some studies in Iquitos (Perú), Montería y Bucaramanga (Colombia) there is a lack of seroprevalence studies from Latin America region. Furthermore, in most studies, the overall prevalence estimates were not further stratified by geographical areas and did not bear in mind the changes of infection rate across various regions.

#### *Added value of this study*

In this population-based study we assessed the seroprevalence of SARS-CoV-2-specific antibodies in some of the most populated cities in Colombia. This is the first seroprevalence study in the Andean region to report the prevalence of anti-SARS-CoV-2 antibodies in the general population during the second semester of 2020. Our findings imply that almost 4.1 million of individuals from the included cities were infected between March and December 2020. Seroprevalence estimates of SARS-CoV-2-specific antibodies presented an elevated heterogeneity across the general populations in different cities, ranging from 26% to 68%. Compared with other seroprevalence studies from America, Asia and Europe, the seroprevalence estimates in the general population of Colombia in this study were high. In spite of schools having been closed since March 2020, we observed that children and adolescents showed seroprevalence levels similar to the observed in young adults. In contrast, we observed consistently lower seroprevalence levels in adults older than 65 years in cities with late epidemics, with respect to those with early epidemics.

#### *Implications of all the available evidence*

The heterogeneity in the prevalence of SARS-CoV-2-specific antibodies between cities documented by this study confirms that a large proportion of the population in Colombia was still susceptible to the virus by December 2020. Variations on seroprevalence estimates between cities have important public health implications, suggesting the presence of local conditions that may influence adherence to infection control measures, which requires further investigation.

### Introduction

Population-based serological surveys are the gold standard to estimate what proportion of a population has been exposed to a pathogen. For SARS-CoV-2, and due to a large number of asymptomatic and mildly symptomatic cases, surveillance systems that rely on symptomatic cases are insufficient to estimate the total number of infections at the community level.<sup>1</sup> The World Health Organization (WHO) has underscored the relevance of studies quantifying the extent of SARS-CoV-2 transmission at the country level.<sup>2</sup> More importantly, quantifying SARS-CoV-2 seroprevalence is crucial to better understand the transmission dynamics, guiding novel avenues to design and implement public health interventions.<sup>3</sup> Nevertheless, except for cities in Brazil, few studies have been conducted in Latin American cities, despite the very high disease burden.

Between March and December 2020 there were more than 1.5 million confirmed cases and more than 43 thousand COVID-19 related deaths in Colombia.<sup>4</sup> A national lockdown was in place between March 25th and April 15th, followed by a slow but steady increase in mobility until August 2020. Since September, COVID-19 mitigation strategies and implementation of non-pharmaceutical interventions (NPIs) have been largely enforced at a local level, often guided by intensive care unit (ICU) occupancy rates. Similar to what has been observed in other settings, the SARS-CoV-2 epidemic in Colombia has been highly spatially heterogeneous. While some municipalities experienced explosive early peaks followed by periods of very low transmission despite the near absence of NPIs, many others have experienced several moderate peaks interspersed with plateaus of sustained transmission.

In this study, we aimed to provide a comprehensive panorama of the pandemic in Colombia by measuring the seroprevalence of antibodies against SARS-CoV-2 in ten selected cities during the second semester of 2020.

### Methods

#### **Study design, population, and sampling**

We designed a population-based cross-sectional study to measure the seroprevalence of antibodies against SARS-CoV-2 in ten cities across Colombia. The selected cities included seven of the ten most populated cities in Colombia (Bogotá, Medellín, Cali, Barranquilla, Cucuta, Bucaramanga and Villavicencio) as well as cities in key border regions (Leticia, Ipiales and Guapi).

We conducted multi-stage cluster sampling using updated population projections and maps from the national statistical office (DANE). The source population included all the non-institutionalized civilian population

between the ages of five and 80 years old, residing in the urban areas of each city (suburbs were not included). In the first stage, blocks -the primary sampling unit (PSU) - were selected with probability proportional to their population size. In the second stage, segments were selected by randomly sampling clusters of five contiguous dwellings. Households within each segment were invited to participate and surveyed.

The sample size was estimated assuming a seroprevalence of 30%,<sup>5</sup> and marginal sampling errors of 3% for Bogota, Cali, and Medellin, 3.5% for Leticia, Barranquilla, Bucaramanga, Cúcuta, Villavicencio and Ipiales, and 5.0% for Guapi and regions with geographically restricted access. These were equivalent to relative errors of 5%, 6% and 8.5%, respectively. Based on results from prior household surveys conducted by DANE, we assumed a design effect of 2. Detailed information on sample size calculation is provided in *Supplementary Table 1*.

### Study procedures

The fieldwork took place between September 21st and December 11th, 2020. Specific dates for each city are listed in *Supplementary Table 1*. The locations were visited in random order by the study teams, and selected households were approached and invited to participate. Due to logistical constraints, non-responding households were visited three times before neighboring households were invited to participate.

Individuals in the selected households were invited to participate and asked to sign a consent form. All participants were asked to provide a venous blood sample (6 to 7 mL) and to respond to the study questionnaire. In addition, a nasopharyngeal swab was obtained from participants who reported having any symptoms associated with COVID-19 at the time of the visit. After collection, blood samples were refrigerated and transported to a local laboratory where they were centrifuged to separate the serum. They were then stored at -30 °C to -80 °C until processing.

The study questionnaire included 52 questions on sociodemographic characteristics, potential risk factors for COVID-19 transmission, household characteristics, and behavior related to a COVID-19 infection. Before the initiation of the survey, this questionnaire was developed and validated by five experts who assessed clarity, coherence, relevance, and sufficiency. The questionnaire was piloted among 100 volunteers from Bogotá. RedCap® V. 10.1.2 (free license) was used to record the geographic and survey information.

### Laboratory procedures

Prior infection by SARS-CoV2 was ascertained by measuring total antibodies (IgM+ IgG) using the SARS-CoV-2 Total (COV2T) Advia Centaur – Siemens

chemiluminescent immunoassay (CLIA).<sup>6</sup> The COV2T is an in-sandwich one-step automated antigen test for the qualitative detection of total antibodies (IgM + IgG), against the receptor-binding domain (RBD) of the spike protein of SARS-CoV-2 virus, in serum or plasma. This CLIA assay was selected based on the good performance reported by prior published studies and after performing an internal validation with samples from the Colombian population. Negative controls included 221 serum samples collected before 2010, including samples from patients with Dengue, Zika and other arboviruses. Sera from 149 patients with SARS-CoV-2 infection, confirmed by RT-PCR and obtained less than 14 days after the onset of symptoms, were used as positive controls. This validation resulted in a sensitivity of 86% (95% CI% 79–91%) and specificity of 99% (95% CI 96–100%).<sup>6</sup>

### SARS-CoV-2 virus identification and sequencing

Nasopharyngeal swab samples were refrigerated and transported to the local laboratory in each city. The RNA was extracted and amplified according to the Berlin protocol, which was validated by the Colombian National Institute of Health (Instituto Nacional de Salud, INS).<sup>7</sup> Viral RNA extraction was performed by using the QIAamp Viral RNA Mini kit (Qiagen Inc., Chatsworth, CA, USA) or the MagNA Pure LC nucleic acid extraction system (Roche Diagnostics GmbH, Mannheim, Germany).<sup>8</sup> All positive samples by RT-PCR were sequenced. The library preparation and sequencing were performed following the ARTIC network (real-time molecular epidemiology for outbreak response) protocol and using Oxford Nanopore Technologies (Oxford Nanopore Technologies, Oxford).<sup>8</sup> The SARS-CoV-2 genome sequences were deposited in GISAID. The sequences were compiled and combined with representative genome sequences from the South America-focused subsampling available from NextStrain.

### Statistical analysis

General descriptive statistics were used to explore and compare the characteristics of participants and households. For each city we report crude seropositivity estimates as well as adjusted estimates that are jointly (a) weighted to match the population structure of each city by age, sex and socio-economic stratum, (b) adjusted for test-performance characteristics and (c) took into account clustering at the block level. Weights (expansion factors) were calculated using previously established methodology<sup>9</sup> by the National Statistical Office (DANE), and also considered aspects of survey design and non-response as described in *Supplementary Table 2*. Corrections for test performance used the method proposed by Gelman and Carpenter<sup>10</sup> and data on the sensitivity and specificity of the assay measured by

internal validation. Details on the model used for this adjustment are provided in the *Supplementary Methods file*.

To characterize associations between sociodemographic factors and seropositivity, we fit mixed-effect Poisson models using the lmer package in R, to estimate prevalence ratios and 95% confidence intervals.<sup>11,12</sup> All models included a random effect for blocks to account for clustering at this level. The Akaike information criterion (AIC) was used to compare the fit of competing models. To ensure that all models included the same number of observations, multivariate models were fit using a complete dataset where observations that contained missing data in any of the included variables were deleted. Results from an alternative analysis where missing observations were coded as a “missing” category can be found in *Supplementary Table 14*. Analyses were performed in RStudio V.1.4.1103, and RStan V. 2.21.2. Findings are reported according to the STROBE Statement guidelines.<sup>13</sup>

#### Ethics statement

The study proposal and protocol were approved by the ethics committee of Instituto Nacional de Salud (CEMIN 010/2020). We obtained written informed consent from each adult participant as well as oral assent and written parental permission from all participants between five and 17 years of age.

#### Role of the funding source

Beside the funding, the Ministerio de Ciencia y Tecnología e Innovación had no role in the study design, data collection, data analysis, data interpretation, or writing the report.

#### Results

The study included samples from 7,320 households, accounting for 17,863 participants from the ten selected Colombian cities, i.e. Bogotá ( $n = 4,515$ ; 25.2%), Barranquilla ( $n = 1,431$ ; 8.0%), Bucaramanga ( $n = 1,427$ ; 7.9%), Cali ( $n = 1,983$ ; 11.1%), Cúcuta ( $n = 1,448$ ; 8.1%), Medellín ( $n = 1,987$ ; 11.1%), Villavicencio ( $n = 1,467$ ; 8.2%), Leticia ( $n = 1,417$ ; 7.9%), Ipiales ( $n = 1,465$ ; 8.2%), and Guapi ( $n = 712$ ; 4.0%) (*Supplementary Table 1*). Household response rates ranged between 82% and 90%. Among eligible individuals from participating households, the median response rate was 90% but ranged between 83% and 95% (*Supplementary Table 2*).

Characteristics of participating households, and participants are described in *Table 1*. The median size of participating households was two persons (range: 1–17 persons). The majority of households (4122/7320, 56.3%) belonged to the lowest socio-economic strata (strata 1

and 2). This is a 6-level classification system used in Colombia that serves as an indicator of a household's socio-economic conditions. The average age of participants was  $37.1 \pm 20.7$  years old and females accounted for 61.47% of the sample ( $n = 10,980$ ), due to a lower response rate in males (*Table 1*). Almost 99% ( $n = 15,421/15,647$  participants who answered this question) of participants asserted to wash their hands with soap regularly (once or more per day). 98% ( $n = 14,803/15,132$ ) declared that they used a face mask whenever they left home.

Only 764 (4.3%) participants reported having been diagnosed with COVID-19 in the prior months, but 2843 (15.91%) reported “having been in contact with a COVID-19 case”. Characteristics of the participants enrolled in each city are reported in *Supplementary Tables 3–12*.

A total of 5811/17,863 (32.25%) participants tested positive for antibodies against SARS-CoV-2. As expected, crude seroprevalence was heterogeneous between cities and ranged between 26% and 59% (*Table 2*). After correcting for the sensitivity and specificity of the test and weighting for age, sex and socio-economic stratum, our results suggest that more than 50% of the population had been infected in Barranquilla, Guapi and Leticia by the time of the survey (September to December 2020) (*Table 2*, *Supplementary Table 13*). Except for Bucaramanga and Cúcuta, higher seroprevalence levels were observed in cities with populations of 1.5 million or less (*Table 2*). Cities that experienced the first peak during the first semester of 2020 showed higher seroprevalence levels (*Fig. 1*). Seroprevalence was highest in Guapi (68%, 95% CI 62–74%), a small Afro-Colombian city located on the Pacific coast and was significantly lower in the three most populated cities of the country: Bogotá (30%, 95% CI 27–33%), Medellín (26%, 95% CI 24–30%) and Cali (30%, 95% CI 27–33%).

#### Risk factors for seropositivity

At the household level, the main risk factor for seropositivity was the socio-economic stratum (*Fig. 2*). Individuals living in strata 5,6 households had, on average, 60% (95% CI 53–65) lower risk of being seropositive as compared to individuals living in stratum 1 households. This association was particularly strong in Barranquilla, where seropositivity among individuals living in strata 4–6 was 83% (95% CI 68–91%) lower than among individuals living in stratum 1. We also found associations between seropositivity and number of persons in the household (Prevalence Ratio (PR) 1.27, 95% CI 1.13–1.42 comparing households with 7 or more people as compared to those with 1 person), and the number of rooms in the house (PR 0.819, 95% CI 0.670–0.948 comparing households with 5 or more rooms as compared to 1 room). There was no association between

Variable	Categories	Overall	Positive	Crude seroprevalence
95% CI				
Sex	Female	10980	3615	0.33 (0.32–0.34)
	Male	6880	2193	0.32 (0.31–0.33)
	Missing	3	3	1 (0.44–1)
Age class	[5,10]	1150	418	0.36 (0.34–0.39)
	(10,18]	2267	861	0.38 (0.36–0.4)
	(18,40]	6365	2200	0.35 (0.33–0.36)
	(40,60]	4982	1580	0.32 (0.3–0.33)
	(60,100]	2964	708	0.24 (0.22–0.25)
	Missing	135	44	0.33 (0.25–0.41)
Socioeconomic strata	1	3979	1914	0.48 (0.47–0.5)
	2	6571	2205	0.34 (0.32–0.35)
	3	5424	1323	0.24 (0.23–0.26)
	4	1173	203	0.17 (0.15–0.2)
	5	259	39	0.15 (0.11–0.2)
	6	156	15	0.1 (0.06–0.15)
	Missing	301	112	0.37 (0.32–0.43)
Comorbidities	None	12759	4240	0.33 (0.32–0.34)
	1–3	4022	1209	0.3 (0.29–0.31)
	4+	71	18	0.25 (0.17–0.37)
	Missing	1011	344	0.34 (0.31–0.37)
Race	White	3409	954	0.28 (0.27–0.3)
	Afro	1265	637	0.5 (0.48–0.53)
	Indigenous	779	398	0.51 (0.48–0.55)
	Mestizo	8855	2701	0.31 (0.3–0.31)
	Other	1264	372	0.29 (0.27–0.32)
	Missing	2291	749	0.33 (0.31–0.35)
	Civil Status	Single	7306	2296
Married		3693	991	0.27 (0.25–0.28)
Cohabiting		3630	1409	0.39 (0.37–0.4)
Widower		680	197	0.29 (0.26–0.32)
Divorced		600	173	0.29 (0.25–0.33)
Other		1721	667	0.39 (0.36–0.41)
Missing		233	78	0.33 (0.28–0.4)
Educational level	Elementary	4807	1763	0.37 (0.35–0.38)
	High school	6333	2209	0.35 (0.34–0.36)
	Technical degree	2689	788	0.29 (0.28–0.31)
	Bachelor degree +	2687	577	0.21 (0.2–0.23)
	None	1055	386	0.37 (0.34–0.4)
	Missing	292	88	0.3 (0.25–0.36)
Occupation	Student	3940	1379	0.35 (0.34–0.37)
	Employed	3985	1186	0.3 (0.28–0.31)
	Independent	3253	1079	0.33 (0.32–0.35)
	Housewife	3632	1226	0.34 (0.32–0.35)
	Informal worker	350	162	0.46 (0.41–0.52)
	Retired	809	133	0.16 (0.14–0.19)
	Unemployed	1436	499	0.35 (0.32–0.37)
	Missing	458	147	0.32 (0.28–0.37)
Health insurance affiliation	Private	9586	2575	0.27 (0.26–0.28)
	Public	6894	2727	0.4 (0.38–0.41)
	None	958	374	0.39 (0.36–0.42)
	Unknown	45	13	0.29 (0.18–0.43)
	Missing	380	122	0.32 (0.28–0.37)

Table 1 (Continued)

Variable	Categories	Overall	Positive	Crude seroprevalence
Access to sewage	Yes	16621	5123	0.31 (0.3–0.32)
	No	1103	639	0.58 (0.55–0.61)
	Missing	139	49	0.35 (0.28–0.43)
Access to electricity	Yes	17580	5713	0.32 (0.32–0.33)
	No	114	45	0.39 (0.31–0.49)
	Missing	169	53	0.31 (0.25–0.39)
Number of people per household	1,2	4957	1355	0.27 (0.26–0.29)
	3,4	7511	2283	0.3 (0.29–0.31)
	5,6	3269	1233	0.38 (0.36–0.39)
	7+	1779	827	0.46 (0.44–0.49)
	Missing	347	113	0.33 (0.28–0.38)
Number of rooms per household	1	1353	545	0.4 (0.38–0.43)
	2	4871	1685	0.35 (0.33–0.36)
	3	7127	2167	0.3 (0.29–0.31)
	4	2808	854	0.3 (0.29–0.32)
	5+	1565	524	0.33 (0.31–0.36)
	Missing	139	36	0.26 (0.19–0.34)
Contact with COVID–19 patient	Yes	2843	1268	0.45 (0.43–0.46)
	No	13531	4001	0.3 (0.29–0.3)
	Missing	1489	542	0.36 (0.34–0.39)
Usage of facemask	Yes	15076	4756	0.32 (0.31–0.32)
	No	562	271	0.48 (0.44–0.52)
	Missing	2225	784	0.35 (0.33–0.37)
Daily frequency of handwashing	0 times	226	100	0.44 (0.38–0.51)
	1–3 times	2381	885	0.37 (0.35–0.39)
	4,5 times	3042	947	0.31 (0.3–0.33)
	6+ times	9998	3090	0.31 (0.3–0.32)
	Missing	2216	789	0.36 (0.34–0.38)
Influenza vaccine during last year	Yes	2726	860	0.32 (0.3–0.33)
	No	14722	4824	0.33 (0.32–0.34)
	Missing	415	127	0.31 (0.26–0.35)
City	Barranquilla	1426	697	0.49 (0.46–0.51)
	Bogota	4515	1186	0.26 (0.25–0.28)
	Bucaramanga	1425	398	0.28 (0.26–0.3)
	Cali	1983	524	0.26 (0.25–0.28)
	Cucuta	1454	508	0.35 (0.33–0.37)
	Guapi	721	423	0.59 (0.55–0.62)
	Ipiales	1465	445	0.3 (0.28–0.33)
	Leticia	1417	726	0.51 (0.49–0.54)
	Medellin	1987	464	0.23 (0.22–0.25)
	Villavicencio	1470	440	0.3 (0.28–0.32)
Overall	-	17863	5811	0.33 (0.32–0.33)

**Table 1: Baseline characteristics of the study participants.**

access to public utilities (electricity, water, or sewage) and seropositivity (Table 3).

At the individual level, seropositivity was similar between males and females and between children and adults, but lower among participants 60 years or older (PR 0.74 (95%CI 0.65–0.83) compared to participants between 18 and 39 years of age (Fig. 2, Table 3). This was particularly true in cities with large populations (Bogotá, Cali, and Medellín). Consistent with this

finding, seroprevalence was lower among the group of participants who were retired compared with participants who were employed at the time of the study (PR 0.76 (0.61–0.94)). In addition, seroprevalence in certain cities was also higher among the group of participants who were informal or independent workers compared with participants who were employed at the time of the study. We also found a strong association between the educational level and seropositivity.



City	Population size 2020	n	Positive	Crude seroprevalence (95% CI)	Adjusted seroprevalence (95%CI)
Bogotá	7,220,002	4515	1,186	0.262 (0.237–0.287)	0.299 (0.273–0.328)
Barranquilla	1,183,865	1426	697	0.488 (0.451–0.525)	0.565 (0.52–0.618)
Bucaramanga	559,018	1425	398	0.279 (0.235–0.323)	0.319 (0.284–0.356)
Cali	2,050,433	1983	524	0.264 (0.226–0.301)	0.301 (0.270–0.335)
Cúcuta	689,415	1454	508	0.349 (0.307–0.390)	0.402 (0.363–0.445)
Guapi	12,622	721	423	0.586 (0.539–0.633)	0.681 (0.622–0.747)
Ipiales	71,491	1465	445	0.303 (0.261–0.346)	0.348 (0.313–0.386)
Leticia	31,480	1417	726	0.512 (0.475–0.548)	0.594 (0.547–0.649)
Medellín	2,342,939	1987	464	0.233 (0.195–0.272)	0.265 (0.235–0.296)
Villavicencio	465,347	1470	440	0.299 (0.256–0.342)	0.343 (0.307–0.382)
Total	14,626,612	17,863	5,811	..	..

**Table 2: Seroprevalence in the ten cities of Colombia, 2020.**

Participants with a bachelor or technical degree had a lower risk of being seropositive compared with participants without any schooling (PR 0.65 95%CI 0.53–0.79 and PR 0.77 95%CI 0.63–0.93), respectively (Table 3). Seroprevalence also varied by ethnicity and was greater in afro descendants (PR 1.27, 95%CI 1.07–1.51) participants compared to white participants, even after adjusting for socioeconomic stratum and city (Table 3). We did not observe an association between seroprevalence and health insurance type.

Not surprisingly, seroprevalence was higher among participants who reported contact with COVID-19 cases (PR 1.56, 95%CI 1.46–1.68), particularly those who reported caring for or sharing a room with COVID-19 cases (Table 3 and Supplementary Table 14).

#### SARS-CoV-2 isolated strains

Nasopharyngeal swabs were obtained from 272 participants who reported symptoms consistent with COVID-19. Of these, it was possible to obtain full-length sequences of 37 samples that had CTs values < 25. The most frequent sequenced isolated lineages were B.1, B.1.1 and B.1.5 (Supplementary Table 15).

## Discussion

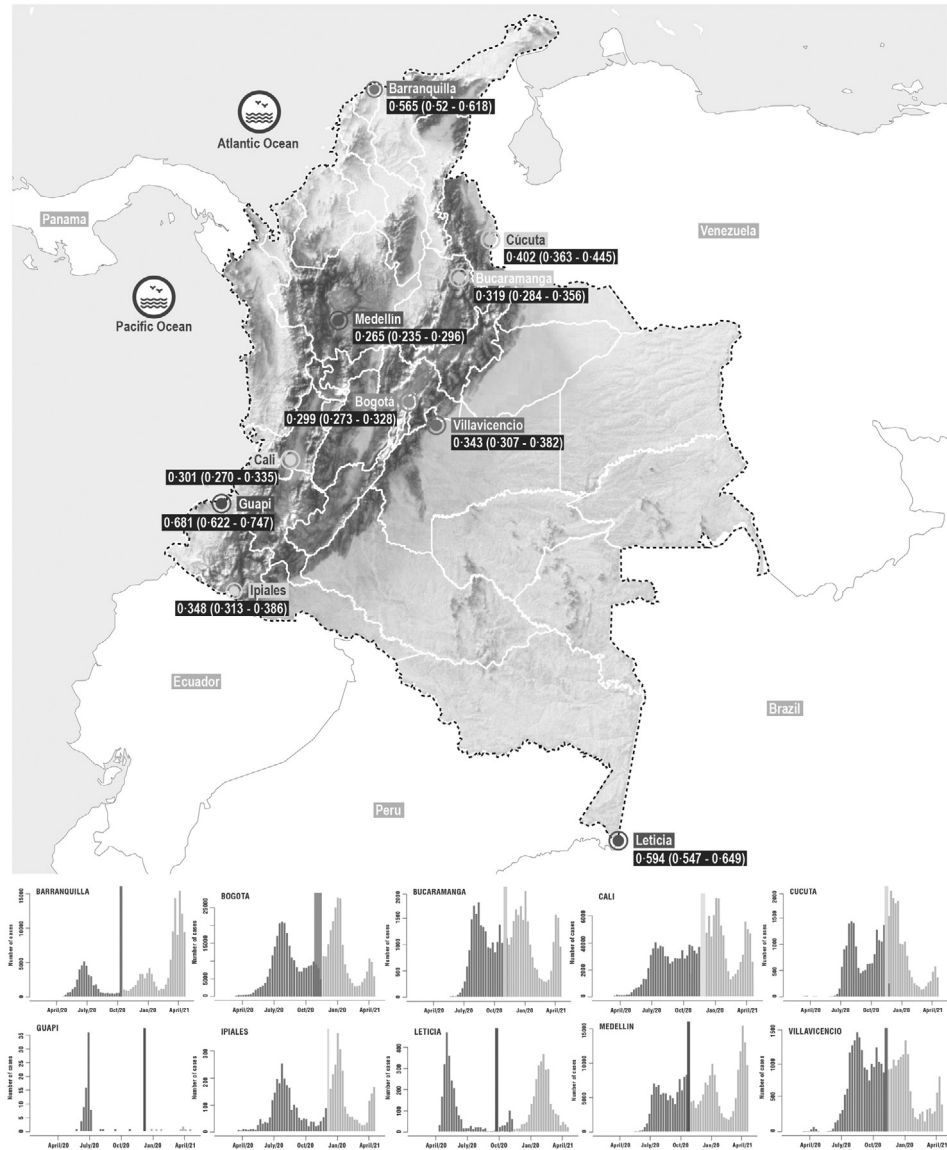
Colombia has been one of the Latin American countries most affected by the COVID-19 pandemic. In this population-based serosurvey, we documented high prevalence of total antibodies against SARS-CoV-2 in several Colombian cities by the end of 2020, consistent with high attack rates during the first pandemic year.

Our results are consistent with large heterogeneity in SARS-CoV-2 transmission experienced in Colombian cities during 2020, with estimated seroprevalences ranging from 26% to 68% in the 10 cities included in this study. Similar heterogeneities have been described in other settings, including a recent study conducted in 18 cities in Iran, where seropositivity ranged between 1.7% and 72.6%.<sup>14</sup>

Our results are also consistent with large variation in seropositivity within cities, largely driven by socio-economic status. Household socio-economic stratum was the strongest independent predictor of seropositivity, and there were significant differences in seropositivity by ethnicity, educational level, and family composition. We did not observe an association between the access to public utilities and greater seroprevalence. Afro-Colombian participants, from low socioeconomic strata, with low educational level and residing in crowded dwellings with fewer rooms have experienced a higher risk of infection by SARS-CoV-2. These results are consistent with prior efforts to characterize disparities in SARS-CoV-2 transmission.<sup>15,16</sup> The self-reported adherence to usage of face masks and hand washing was high (98% and 99%, respectively) in our study. Therefore, it was not possible to establish an association between the usage of face masks or hand hashing and seroprevalence.

Estimated seropositivity was highest in Guapi (78%) and Leticia (62%). Guapi is a small afro-Colombian city located on the Pacific coast. The population is predominantly black, and a large proportion of households live under extreme poverty. Leticia is located in the Amazon jungle in the tri-country border between Colombia-Brazil-Peru and experienced an early epidemic peak in May 2020. Prior studies have shown high seropositivity in other cities in the Amazon including Manaus, Brazil (66%)<sup>17</sup> and Iquitos, Peru (71%).<sup>18</sup> and it has been suggested that virus spread was highly correlated with the main river routes along the Amazon river.<sup>19</sup> Inconsistent policies to mitigate the pandemic in this tri-border area, may have contributed to the high attack rates experienced by these populations.

Other bordering cities included in the study were Cúcuta, next to Venezuela, and Ipiales, which borders Ecuador. The estimated seroprevalence in Cucuta was 38%, even though a clear epidemic peak was not observed by the time of the survey. According to the WHO, more than 40 thousand Venezuelans used to cross the border into Cucuta every day until March 14th,



**Figure 1.** Map of study locations and epidemiological curves.

Map of Colombia showing the locations of the 10 cities where the serosurveys were conducted. Panels surrounding the map show the epidemic curves at each city (weekly number of reported cases). The colored lines and polygons (red, orange, blue, yellow, green, cyan, purple, lilac and light yellow) indicate the period of sample collection for each city relative to the ongoing epidemic.

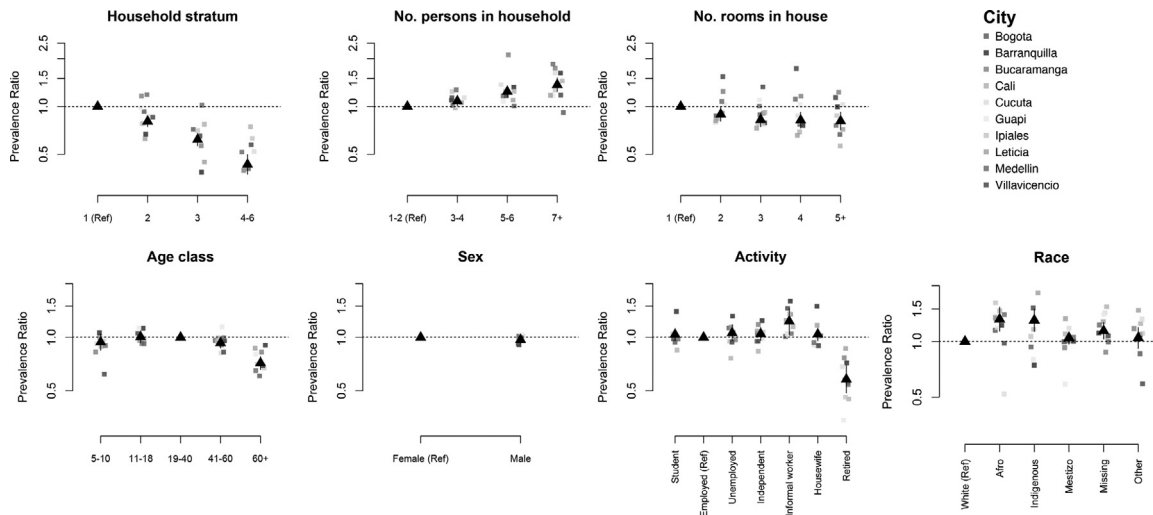
when Colombia closed its borders.<sup>20</sup> However, no estimates have been published from neighboring cities in Venezuela. In Ipiales, the seroprevalence was 31%, consistent with local studies from Ecuador that estimated seroprevalence levels of 44% in the neighboring cities.<sup>21</sup>

Interestingly, the most populated cities in the country (Bogotá, Medellín, and Cali) showed lower seroprevalence levels. These cities implemented strict social distancing measures starting in March 2020, which may have impacted transmission rates. Nevertheless,

despite the lower attack rates, these cities have experienced the largest number of cases and deaths by SARS-CoV-2.<sup>22</sup>

While several prior population-based serosurveys have shown differences in attack rates by sex, we did not find differences in seropositivity between males and females.<sup>23,24</sup> Similarly, we didn't find differences in seropositivity between children and adults, even though schools were still mostly closed by the time of the survey. Interestingly, we did find lower seroprevalence among adults over 60 years of age in several cities,





**Figure 2.** Risk factors for seropositivity.

Association between individual and household risk factors and seropositivity to SARS-CoV-2. Black symbols show the unadjusted mean associations (prevalence ratios) and 95% confidence intervals estimated using data from the 10 cities. Colored symbols show the mean association estimated for each city independently.

		Prevalence ratio (95% CI)	p
Sex	Male	1	
	Female	1.037 (0.963–1.116)	0.335
Age class	[5,10]	0.849 (0.595–1.212)	0.368
	(10,18]	0.996 (0.851–1.165)	0.961
	(18,40]	1	
	(40,60]	0.892 (0.824–0.967)	0.005
	(60,100]	0.735 (0.652–0.828)	0.000
Race	White	1	
	Afro	1.274 (1.073–1.512)	0.006
	Indigenous	1.092 (0.926–1.288)	0.298
	Mestizo	1.036 (0.95–1.13)	0.424
	Other	1.019 (0.881–1.178)	0.800
Educational level	Elementary	0.993 (0.839–1.176)	0.936
	High school	0.908 (0.764–1.08)	0.275
	Technical degree	0.773 (0.639–0.935)	0.008
	Bachelor degree +	0.652 (0.533–0.798)	0.000
	None	1	
Socioeconomic strata	1 (lower)	1	
	2	0.822 (0.752–0.898)	0.000
	3	0.648 (0.583–0.719)	0.000
	4	0.54 (0.448–0.652)	0.000
	5	0.49 (0.338–0.71)	0.000
	6	0.317 (0.181–0.555)	0.000
Occupation	Student	0.867 (0.747–1.006)	0.061
	Employed	1	
	Independent	1.02 (0.922–1.128)	0.701
	Housewife	0.96 (0.861–1.072)	0.470
	Informal worker	1.126 (0.915–1.385)	0.262
	Retired	0.758 (0.609–0.944)	0.013
	Unemployed	0.956 (0.841–1.086)	0.485
Health insurance affiliation	Private	0.988 (0.911–1.071)	0.763

**Table 3** (Continued)

		Prevalence ratio (95% CI)	p
	Public	1	
	None	1.074 (0.932–1.239)	0.324
	Unknown	0.789 (0.421–1.475)	0.457
Access to sewage	Yes	1	
	No	1.04 (0.893–1.21)	0.616
Number of people per household	1,2	1	
	3,4	1.057 (0.972–1.15)	0.194
	5,6	1.175 (1.06–1.302)	0.002
	7+	1.259 (1.109–1.428)	0.000
Number of rooms per household	1	1	
	2	0.895 (0.788–1.016)	0.087
	3	0.851 (0.747–0.969)	0.015
	4	0.85 (0.733–0.985)	0.031
	5+	0.798 (0.676–0.943)	0.008
Contact with COVID-19 patient	Yes	1	
	No	0.623 (0.576–0.674)	0.000
City	Barranquilla	1.396 (1.22–1.597)	0.000
	Bucaramanga	1.003 (0.867–1.161)	0.965
	Cali	0.837 (0.732–0.956)	0.009
	Cucuta	0.956 (0.826–1.107)	0.550
	Guapi	1.311 (1.031–1.666)	0.027
	Ipiales	0.914 (0.785–1.064)	0.244
	Leticia	1.334 (1.143–1.556)	0.000
	Medellin	0.776 (0.675–0.893)	0.000
	Villavicencio	1.052 (0.911–1.214)	0.490
	Bogota	1	

**Table 3: Results of multivariate mixed-effects Poisson regression quantifying the association between seropositivity to SARS-CoV-2 and household and individual factors.<sup>1</sup>**

<sup>1</sup> Model was fitted to a dataset (n = 12,083), where individuals with missing observations in any of the model variables were removed. An alternative analysis, where missing observations for each variable are modeled are included in the supplement (*Supplementary Table 14*).

which may have resulted from social distancing measures aiming to protect this highly vulnerable group.

The nasopharyngeal samples from symptomatic individuals in this study were sequenced to determine the lineages of the virus in the positive cases. The most frequent lineages during the time of sampling were B, B.1.1, B.1.5 (*Supplementary Table 14*). According to the GISAID, in Colombia, the SARS-CoV-2 diversity has been classified mainly into 12 sub-lineages, i.e. A.1.2, A.2, A.5, B, B.1, B.1.1, B.1.3, B.1.5, B.1.8, B.1.11, B.2, and B.2.5. These results serve to generate a baseline for the identification of imported or future local lineages.

Our study has several strengths including the probabilistic sampling of participants, the high participation rates, and the use of a serological assay that has shown excellent sensitivity and specificity in validation studies. Nevertheless, the study also has several limitations. First, even though the study was conducted in several of the most populated areas and border cities of Colombia, the results are not representative of the unsampled populations or of the country as a whole. Notably, our study did not include rural or institutionalized populations that might have experienced distinct epidemic patterns

from those captured here. Similarly, the study did not include pregnant women or people with certain pre-existing conditions. Second, since the study was conducted over a three-month period, some differences observed between cities might have been a consequence of timing. Some of the cities that were enrolled in the study were already experiencing the start of a second epidemic peak, and this may have resulted in higher seroprevalences as compared to cities that were sampled earlier. Additional serosurvey rounds would be required to ascertain the current seroprevalence levels in these populations, particularly since most of the included cities have experienced epidemic peaks after this study was conducted. Finally, our study quantified the presence of total Ig antibodies, which is a marker of prior infection by SARS-CoV-2 but does not necessarily reflect protective immunity.

Results from this study are important to understand the epidemiology of SARS-CoV-2 in Latin America and provide critical information on the burden of pandemic during 2020. These results also provide a baseline for future serological studies aiming to characterize transmission of the virus during 2021. The findings also

show the need to keep health preventive measures to limit the spread of the virus.

### Contributors

MM, MO, MW, IRB, JAR, JPHO, GO, MIC, MCM, MIE, JC LAV, JC and ZMC conceived the study design. EN, EMV, JDO, NP, MM, NS and MO designed the field operative. MM, JMR, GP, MW, MG, SZ, NP, MG, MIC coordinated the field operatives. VR and LM performed the CLIA tests at the lab. CFM performed the RT-PCR and sequencing of SARS-CoV-2 in nasal swabs. SZ, YGTP, EI, and JMR performed the data curation. SZ, IRB, ZMC, MM, JMR, YT, and EI performed the statistical analysis. MM, EP, JMR, IRB, EI, JPHO, and YGTP prepared the draft of the manuscript. All the authors have reviewed and approved the final manuscript.

### Data sharing

The study protocol and anonymized individual participant data that underlie the results reported in this manuscript, may be shared with investigators, whose proposed use of the data has been approved by the ethical and intellectual property committees of *Instituto Nacional de Salud*. Data can be provided for each individual participant, data meta-analysis, or other projects comparing the seroprevalence estimates in different regions. The proposals should be directed to the corresponding author at [mmmercado@ins.gov.co](mailto:mmmercado@ins.gov.co). To gain access, data requesters will need to sign a data access agreement, confirmed by MM as principal investigator.

### Funding

The study was sponsored by, Ministerio de Ciencia y Tecnología e Innovación –CT 361/2020-, Ministerio de Salud y Protección Social, Fundación Universitaria del Norte, Imperial College of London, Universidad Nacional de Colombia (Sede Medellín), Universidad de Córdoba, California University, Unidad Nacional de Gestión del Riesgo, Centro de Atención y Diagnóstico de Enfermedades Infecciosas -CDI-, Centro Internacional de Entrenamiento e Investigaciones Médicas -CIDEIM-, Departamento Administrativo Nacional de Estadística -DANE, Fondo Nacional de Turismo -FONTUR-, Secretarías de Salud Departamentales, Distritales y Municipales and Instituto Nacional de Salud. Beside the funding, the Ministerio de Ciencia y Tecnología e Innovación had no role in the study design, data collection, data analysis, data interpretation, or writing the report.

### Editor note

The *Lancet* Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

### Declaration of interests

All the authors declare no competing interests.

### Acknowledgments

The authors thank all professionals from the FETP epidemiologists, the COVID-19 research team at the Instituto Nacional de Salud, and all the Health Secretaries who supported the research in the field. A special thanks to Jency Catama, Zonia Alarcón, Stephany Botero, Maria Teresa Herrera, Jonathan Reales-González, Ligia Oviedo, Yolima Reyes-Pinto, Paula Díaz, Juan F Bedoya, Mónica Palma, Liliana Serrano, Sofía Duque, Ruth Palma, Ronald López and Norma Celly for their technical support, logistical management to carry out the project. Also, a special gratitude to the Armada Nacional de Colombia and Policía Nacional de Colombia and Unidad Nacional para la Gestión del Riesgo y Desastres. The Universidad Nacional de Colombia and the One-Health Genomic Laboratory are thankful to Grupo ISA for their financial support that leveraged sampling logistics in five of the 10 cities. Finally, we are grateful for the valuable collaboration of community leaders and residents of the municipalities for their altruistic participation.

### Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.lana.2022.100195](https://doi.org/10.1016/j.lana.2022.100195).

### References

- Stringhini S, Wisniak A, Piumatti G, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. *Lancet N Am Ed*. 2020;396:313–319.
- World Health Organization. *Population-Based age-Stratified Seroepidemiological Investigation Protocol for COVID-19 Virus Infection*, 17 March 2020. Geneva: World Health Organization; 2020. <https://apps.who.int/iris/handle/10665/331656>.
- Wolff F, Dahma H, Duterme C, et al. Monitoring antibody response following SARS-CoV-2 infection: diagnostic efficiency of 4 automated immunoassays. *Diagn Microbiol Infect Dis*. 2020;98: 115140.
- Instituto Nacional de Salud. Estimación del número reproductivo efectivo Rt para COVID-19 en Colombia. 2020. <https://www.ins.gov.co/Direcciones/ONS/modelos-de-estimacion> (accessed Jan 19, 2021).
- Instituto Nacional de Salud. Seroprevalencia de SARS-CoV-2 durante la epidemia en Colombia: estudio de país. 2021 <http://www.ins.gov.co/BibliotecaDigital/Informe-seroprevalencia-SARS-CoV-2-durante-la-epidemia-en-Colombia.pdf> (accessed Feb 17, 2021).
- Mercado-Reyes M., Zabaleta G. Validación secundaria y verificación del desempeño de la prueba serológica “SARS-CoV-2 Total (COV2T) Advia Centaur – Siemens”. 2020. [https://www.ins.gov.co/Direcciones/Investigacion/Informacion-sobre-pruebas/Pruebas%20serol%C3%B3gicas%20CLIA%20y%20ELISA/2-Validacion\\_prueba\\_serologica\\_SARS-CoV-2\\_Total\\_COV2T\\_Advia\\_Centaur\\_Siemens.pdf](https://www.ins.gov.co/Direcciones/Investigacion/Informacion-sobre-pruebas/Pruebas%20serol%C3%B3gicas%20CLIA%20y%20ELISA/2-Validacion_prueba_serologica_SARS-CoV-2_Total_COV2T_Advia_Centaur_Siemens.pdf) (accessed Jan 19, 2021).
- Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*. 2020;25. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.
- Laiton-Donato K, Villabona-Arenas CJ, Usme-Ciro JA, et al. Genomic epidemiology of severe acute respiratory syndrome coronavirus 2, Colombia. *Emerg Infect Dis*. 2020;26:2854–2862.

- 9 DANE. Ficha metodológica Gran Encuesta integrada de Hogares GEIH. 2016. [https://www.dane.gov.co/files/investigaciones/fichas/empleo/ficha\\_metodologica\\_GEIH-01\\_V10.pdf](https://www.dane.gov.co/files/investigaciones/fichas/empleo/ficha_metodologica_GEIH-01_V10.pdf) (accessed June 3, 2021).
- 10 Gelman A, Carpenter B. Bayesian analysis of tests with unknown specificity and sensitivity. *J R Stat Soc C*. 2020;69:1269–1283.
- 11 Deddens JA, Petersen MR. Approaches for estimating prevalence ratios. *Occup Environ Med*. 2008;65:501–506.
- 12 Santos CAS, Fiaccone RL, Oliveira NF, et al. Estimating adjusted prevalence ratio in clustered cross-sectional epidemiological data. *BMC Med Res Methodol*. 2008;8:80.
- 13 von EE, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandembroucke JP. The strengthening of reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet N Am Ed*. 2007;370:1453–1457.
- 14 Poustchi H, Darvishian M, Mohammadi Z, et al. SARS-CoV-2 antibody seroprevalence in the general population and high-risk occupational groups across 18 cities in Iran: a population-based cross-sectional study. *Lancet Infect Dis*. 2021;21:473–481.
- 15 Muñoz-Price LS, Nattinger AB, Rivera F, et al. Racial disparities in incidence and outcomes among patients With COVID-19. *JAMA Netw Open*. 2020;3:e2021892.
- 16 Millett GA, Jones AT, Benkeser D, et al. Assessing differential impacts of COVID-19 on black communities. *Ann Epidemiol*. 2020;47:37–44.
- 17 Buss LF, Prete CA, Abraham CM, et al. COVID-19 herd immunity in the Brazilian Amazon. *Infectious diseases (except HIV/AIDS)*, 2020 DOI:10.1101/2020.09.16.20194787.
- 18 Álvarez-Antonio C, Meza-Sánchez G, Calampa C, et al. Seroprevalence of anti-SARS-CoV-2 antibodies in Iquitos, Peru in July and August, 2020: a population-based study. *Lancet Glob Health*. 2021;0. [https://doi.org/10.1016/S2214-109X\(21\)00173-X](https://doi.org/10.1016/S2214-109X(21)00173-X).
- 19 Hallal PC, Hartwig FP, Horta BL, et al. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. *Lancet Glob Health*. 2020;8:e1390–e1398.
- 20 Daniels JP. Venezuelan migrants “struggling to survive” amid COVID-19. *Lancet North Am Ed*. 2020;395:1023.
- 21 Del Brutto OH, Mera RM, Recalde BY, Costa AF. Social determinants of health and risk of SARS-CoV-2 infection in community-dwelling older adults living in a rural Latin American setting. *J Commun Health*. 2021;46:292–297.
- 22 Malagón-Rojas JN, Mercado-Reyes M, Toloza-Pérez YG, et al. Seroprevalence of the SARS-CoV-2 antibody in healthcare workers: a multicentre cross-sectional study in 10 Colombian cities. *Occup Environ Med*. 2021. oemed-2021-107487.
- 23 Lai CC, Wang JH, Hsueh PR. Population-based seroprevalence surveys of anti-SARS-CoV-2 antibody: an up-to-date review. *Int J Infect Dis*. 2020;101:314–322.
- 24 Šmigelskas K, Petrikonis K, Kasiulevičius V, et al. SARS-CoV-2 seroprevalence in Lithuania: results of national population survey. *1* 2021; 28: 2–2.