

Investigating SARS-CoV-2 Incidence and Morbidity in Ponce, Puerto Rico: Protocol and Baseline Results from a Community Cohort Study

Chelsea G. Major, Dania M. Rodríguez, Liliana Sánchez-González, Vanessa Rodríguez-Estrada, Tatiana Morales-Ortíz, Carolina Torres, Nicole M. Pérez-Rodríguez, Nicole Anai Medina-López, Neal Alexander, David Mabey, Kyle Ryff, Rafael Tosado-Acevedo, Jorge Muñoz-Jordán, Laura E. Adams, Vanessa Rivera-Amill, Melissa Rolfes, Gabriela Paz-Bailey

Submitted to: JMIR Research Protocols
on: October 20, 2023

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 39

 Figures 40

 Figure 1..... 41

 Figure 2..... 42

Investigating SARS-CoV-2 Incidence and Morbidity in Ponce, Puerto Rico: Protocol and Baseline Results from a Community Cohort Study

Chelsea G. Major¹ MPH; Dania M. Rodríguez¹ MS; Liliana Sánchez-González¹ MD, MPH; Vanessa Rodríguez-Estrada² BS; Tatiana Morales-Ortiz² BS; Carolina Torres^{1,3} MPH; Nicole M. Pérez-Rodríguez¹ MPH, DrPH; Nicole Anai Medina-López¹ MPH; Neal Alexander⁴ PhD; David Mabey⁵ MD; Kyle Ryff¹ MPH; Rafael Tosado-Acevedo¹ PhD; Jorge Muñoz-Jordán¹; Laura E. Adams¹ DVM, MPH; Vanessa Rivera-Amill² PhD; Melissa Rolfes⁶ PhD; Gabriela Paz-Bailey¹ MD, PhD, MPH

¹Division of Vector Borne Diseases Centers for Disease Control and Prevention San Juan PR

²Ponce Research Institute Ponce Health Sciences University Ponce PR

³K?pili Services, LLC Alaka`ina Foundation Family of Companies Orlando US

⁴Faculty of Epidemiology and Population Health London School of Hygiene and Tropical Medicine London GB

⁵Faculty of Infectious and Tropical Diseases London School of Hygiene and Tropical Medicine London GB

⁶Influenza Division Centers for Disease Control and Prevention Atlanta US

Corresponding Author:

Chelsea G. Major MPH

Division of Vector Borne Diseases

Centers for Disease Control and Prevention

1324 Calle Cañada

San Juan

PR

Abstract

Background: A better understanding of SARS-CoV-2 infection risk among Hispanic and Latino populations and in low-resource settings in the U.S. is needed to inform control efforts and strategies to improve health equity. Puerto Rico has a high poverty rate and other population characteristics associated with increased vulnerability to COVID-19 and limited data to determine community incidence.

Objective: This report describes the protocol and baseline seroprevalence of SARS-CoV-2 in a prospective community-based cohort study (COCOVID) to investigate SARS-CoV-2 infection incidence and morbidity in Ponce, Puerto Rico.

Methods: In June 2020, we implemented the COCOVID study within the Communities Organized to Prevent Arboviruses (COPA) project platform among residents of 15 communities in Ponce, Puerto Rico age 1 year and older. Weekly, participants answered questionnaires on acute symptoms and preventive behaviors and provided anterior nasal swab samples for SARS-CoV-2 PCR testing; additional anterior nasal swabs were collected for expedited PCR testing from participants that reported ?1 COVID-like symptoms. At enrollment and every 6 months during follow-up, participants answered more comprehensive questionnaires and provided venous blood samples for SARS-CoV-2 IgG antibody testing (indicative of seroprevalence). Weekly follow-up activities concluded in April 2022 and 6-month follow-up in August 2022.

Results: We enrolled a total of 1,030 participants from 388 households. Relative to the general populations of Ponce and Puerto Rico, our cohort overrepresented middle-income households, employed and middle-aged adults, and older children ($P < .001$). Almost all participants (99.6%) identified as Latino/a, 17% had annual household incomes less than \$10,000, and 45% reported one or more chronic medical conditions. Baseline SARS-CoV-2 seroprevalence was low (1%) overall and increased with later study enrollment time ($P = .003$).

Conclusions: The COCOVID study will provide a valuable opportunity to better estimate the burden of SARS-CoV-2 in the Ponce region, assess the limitations of surveillance, and inform mitigation measures in Puerto Rico and other similar populations.

(JMIR Preprints 20/10/2023:53837)

DOI: <https://doi.org/10.2196/preprints.53837>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org>

Original Manuscript

Investigating SARS-CoV-2 Incidence and Morbidity in Ponce, Puerto Rico: Protocol and Baseline Results from a Community Cohort Study

Preprint
JMIR Publications

Abstract

Background: A better understanding of SARS-CoV-2 infection risk among Hispanic and Latino populations and in low-resource settings in the U.S. is needed to inform control efforts and strategies to improve health equity. Puerto Rico has a high poverty rate and other population characteristics associated with increased vulnerability to COVID-19 and limited data to date to determine community incidence. This report describes the protocol and baseline seroprevalence of SARS-CoV-2 in a prospective community-based cohort study (COCOVID) to investigate SARS-CoV-2 infection incidence and morbidity in Ponce, Puerto Rico.

Methods: In June 2020, we implemented the COCOVID study within the Communities Organized to Prevent Arboviruses (COPA) project platform among residents of 15 communities in Ponce, Puerto Rico age 1 year and older. Weekly, participants answered questionnaires on acute symptoms and preventive behaviors and provided anterior nasal swab samples for SARS-CoV-2 PCR testing; additional anterior nasal swabs were collected for expedited PCR testing from participants that reported ≥ 1 COVID-like symptoms. At enrollment and every 6 months during follow-up, participants answered more comprehensive questionnaires and provided venous blood samples for multi-antigen SARS-CoV-2 IgG antibody testing (indicative of seroprevalence). Weekly follow-up activities concluded in April 2022 and 6-month follow-up in August 2022. Primary study outcome measures include SARS-CoV-2 infection incidence and seroprevalence, relative risk of SARS-CoV-2 infection by participant characteristics, SARS-CoV-2 household attack rate, and COVID-19 illness characteristics and outcomes. In this report, we describe characteristics of COCOVID participants overall and by SARS-CoV-2 seroprevalence status at baseline.

Results: We enrolled a total of 1,030 participants from 388 households. Relative to the general populations of Ponce and Puerto Rico, our cohort overrepresented middle-income households, employed and middle-aged adults, and older children ($P < .001$). Almost all participants (99.6%) identified as Latino/a, 17% had annual household incomes less than \$10,000, and 45% reported one

or more chronic medical conditions. Baseline SARS-CoV-2 seroprevalence was low (1%) overall and increased with later study enrollment time ($P=.003$).

Conclusions: The COCOVID study will provide a valuable opportunity to better estimate the burden of SARS-CoV-2 and associated risk factors in a primarily Hispanic/Latino population, assess the limitations of surveillance, and inform mitigation measures in Puerto Rico and other similar populations.

Keywords: SARS-CoV-2; COVID-19; research methodology; prospective studies; cohort studies; epidemiologic studies; Hispanic or Latino; seroprevalence; incidence

Introduction

After the first detection of SARS-CoV-2 in late 2019, the virus has become a significant threat to health and economies globally, particularly in populations with high economic inequality and low access to health resources and infrastructure.[1, 2] In the U.S., racial and ethnic groups with well-documented socioeconomic disparities have also been disproportionately affected by SARS-CoV-2, with the highest COVID-19 morbidity and mortality rates among American Indian or Alaska Native, Non-Hispanic Black, and Hispanic or Latino populations.[3, 4] Despite tremendous progress in our understanding of SARS-CoV-2 and development of COVID-19 treatment and vaccines, control efforts have been hindered by the emergence of new, more contagious variants and vaccine hesitancy. [5-9] A better understanding of risk and risk factors for SARS-CoV-2 infection and transmission dynamics in minority groups and low-resource settings is needed to inform ongoing control efforts and develop strategies to improve health equity.

Puerto Rico, a Caribbean Island and U.S. Commonwealth, has a high prevalence of population characteristics associated with increased vulnerability to COVID-19. Among approximately 3.3 million Puerto Rico residents, almost all (99%) identify as Hispanic or Latino and 43% live below the federal poverty level.[10] Additionally, risk factors for severe COVID-19, including older age and chronic disease, are higher among Puerto Ricans relative to those of the continental U.S., and economic crises and natural disasters have weakened the local health care infrastructure over the past decade.[11-13] As in many regions, facility-based public health surveillance and seroprevalence data from blood banks have been and continue to be the primary sources of SARS-CoV-2 incidence and prevalence data in Puerto Rico.[14-16] While valuable, these data are limited in depth and scope due to reliance on patient health care seeking behavior and passive reporting by providers and diagnostic laboratories. Systematic diagnostic testing and follow-up of a defined population can better capture the spectrum of COVID-19 disease, determine true infection rates, and help identify risk factors for infection and morbidity.

As part of the COVID-19 emergency response efforts, the U.S. Centers for Disease Control and Prevention (CDC) supported multiple community-based cohort studies to better understand SARS-CoV-2 transmission and risk in different settings.[17-19] The ongoing Communities Organized to Prevent Arboviruses (COPA) cohort study in Ponce, Puerto Rico — implemented in 2018 through the collaboration of the Ponce Health Sciences University, Puerto Rico Vector Control Unit, and the CDC Dengue Branch (DB) — was identified as an existing platform that could be expanded to estimate SARS-CoV-2 community infection rates and other COVID-19 research objectives in the region. In June 2020, we implemented the COPA COVID-19 (COCOVID) study, a prospective, community-based cohort to investigate the incidence of and risk factors for SARS-CoV-2 infection and morbidity, among COPA participants and other residents of 15 community areas.

In this report, we describe the protocol and baseline population demographics and seroprevalence of SARS-CoV-2 from the COCOVID cohort. The study included weekly molecular testing and multi-antigen serology testing at baseline and every 6 months for SARS-CoV-2. We also collected sociodemographic data, weekly mobility and personal preventive behaviors employed, COVID-19 vaccination status, and symptoms and long-term outcomes among individuals with and without incident SARS-CoV-2 infection during the study period. To date, few longitudinal COVID-19 studies with regular SARS-CoV-2 molecular testing have included U.S. or Hispanic/Latino populations,[20, 21] and none have been conducted in Puerto Rico. Thus, data from the COCOVID study has the potential to enhance general knowledge of SARS-CoV-2 epidemiology in addition to providing unique insights into the COVID-19 epidemic in Puerto Rico, which may have implications for prevention efforts and future epidemics in geographically or demographically similar populations.

Methods

Study overview

In June 2020, we implemented the COCOVID (COPA COVID-19) sub-study among residents of 15 community areas in Ponce. We aimed to enroll between 900 and 1,000 participants for regular surveys and specimen collection, including weekly anterior nasal swab collection for SARS-CoV-2 PCR testing and venous blood draws for multi-antigen IgG antibody testing at baseline and every 6 months. Participants were also included in activities for the larger COPA cohort for at least the same period that they were active in the COCOVID study, including weekly reporting of acute COVID-like symptoms via text message and additional SARS-CoV-2 PCR testing among symptomatic participants.

Ethical considerations

Local ethics review and approval for COPA project and COCOVID sub-study protocols were obtained through the Ponce Medical School Foundation, Inc. Institutional Review Board (IRB) (Study #171110-VR). CDC also reviewed COPA and COCOVID protocols and determined the activities were conducted consistent with applicable federal law and CDC policy (see 45 C.F.R. part 46; 21 C.F.R. part 56).

Adult consent, minor assent, and parental/guardian permission were obtained at enrollment and subsequent 6-month follow-up visits and documented using paper forms. Receipt of consent and assent for all participants was also confirmed daily and tracked electronically by study coordinators. All participants under 21 years, the age of majority in Puerto Rico, were considered minors except for those emancipated by marriage or judicial decree. Before interview or specimen collection, staff explained project activities and important points of the informed consent to participants, answered any questions raised, and provided a paper copy of the consent form to the household. The household representative (an adult or an emancipated minor that was present at the time of visit) consented to provide household level information and allow study staff to invite other household members to

participate in the study.

All adults and emancipated minors (including household representatives) were asked to provide written consent for each of the following:

- Providing contact information;
- Responding to the COPA annual interview questions and providing blood and nasal swab specimens (where applicable);
- Being contacted weekly for acute illness surveillance (AIS) and annually for cohort follow-up;
- Responding to the COCOVID questionnaire and providing blood specimens every 6 months and nasal swabs weekly;
- Storage of specimens remaining after testing for future studies; and
- Permission for children for which they are the parent or legal guardian to participate in the same project activities.

The adult consent document requested separate enumeration of each child for which they have parental or guardianship rights, to account for any differences in which activities each child may participate. For minors aged 14–20 years, written assent was obtained using a designated form simplified from the adult version. Verbal assent was documented in a designated form for minors ages 7–13 years, and only parent or guardian consent was documented for minors ages 1–6 years. All consent and assent forms and questionnaires were available in both Spanish and English and written at an 8th grade or lower reading level.

This project involved collection of personal identifying information (PII) and included safeguard measures to ensure privacy and confidentiality of participants. Households and individual participants were assigned project-specific identifiers to track their location of residence and study status. Barcode stickers were used to link specimen tubes and paper forms to each participant visit during the project. All specimens were securely transported to and stored at the study office site,

CDC Dengue Branch, and other laboratory sites. All personnel involved in this study were required to complete ethics training requirements set by the IRB and adhere to an unwavering code of conduct regarding the confidentiality of patients' information.

Only study personnel have access to participant responses or laboratory results. Study questionnaires, laboratory test results, and other collected data are electronically stored in locations approved by CDC for storage of personal data, including a limited access encrypted server and shared drive within the CDC internal network, and, in the case of paper forms, in locked cabinets. Paper records will be kept for at least 3 years and then destroyed. Where possible, de-identified datasets will be created and used for analysis, and individual participants will not be identified any presentations or publications based on project results.

For compensation of the time dedicated to the COCOVID study activities, participants received \$5 (USD) each week upon staff receipt of the completed weekly questionnaire and self-collected respiratory specimen, which was given in \$20 installments for every four successful transfers of weekly questionnaires and specimens. Participants also received \$20 at each study appointment (enrollment, 6-month, 12-month, 18-month, and 24-month) when they completed serum specimen collection and the questionnaire. If these appointments overlapped with COPA annual follow-up visits (enrollment, 12-month, and 24-month), an additional \$20 (June 2021–February 2022) or \$30 (March–August 2022) in compensation was given for both activities for a total of \$40 or \$50.

COPA platform

The COPA study established an ongoing cohort of approximately 3,800 participants in 38 community areas in Ponce, Puerto Rico (Figure 1) to measure incidence and prevalence of arboviral infections and evaluate vector control interventions via a cluster randomized controlled trial design.[22, 23] Recruitment activities are conducted primarily through house-to-house visits, and random selections of residences in each community area are made until sample size goals are reached or all residences

have been approached for recruitment. Eligible individuals are 1–50 years of age, spend four or more nights a week in the selected residence, and have no definite plans to move in the next 12 months. Follow-up is conducted annually and includes serum collection for antibody testing for dengue and administration of a questionnaire (Figure 2). Cohort enrollment began in March 2018.

Figure 1. Map of the 38 COPA and 15 COCOVID community cohort study clusters areas in Ponce, Puerto Rico.

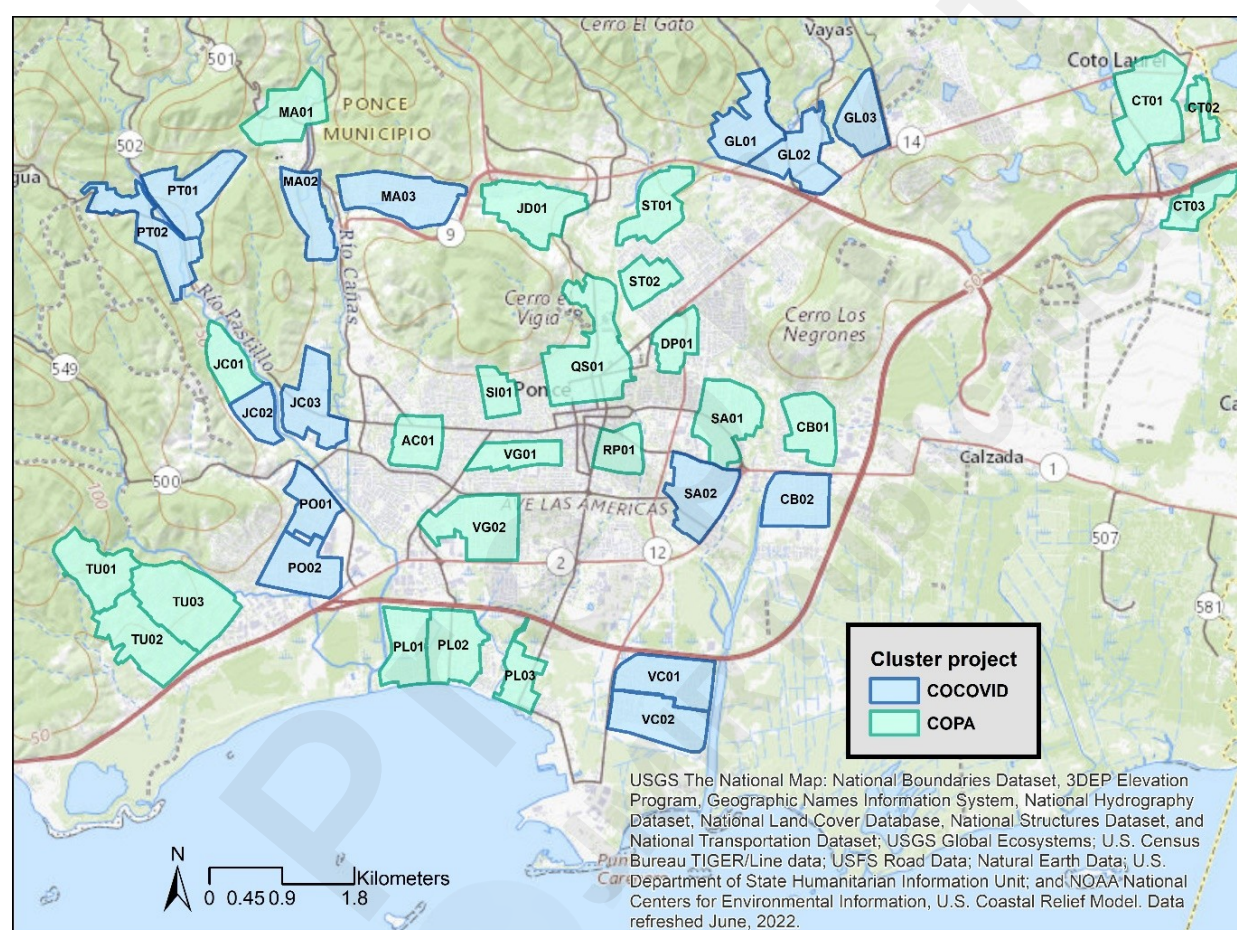
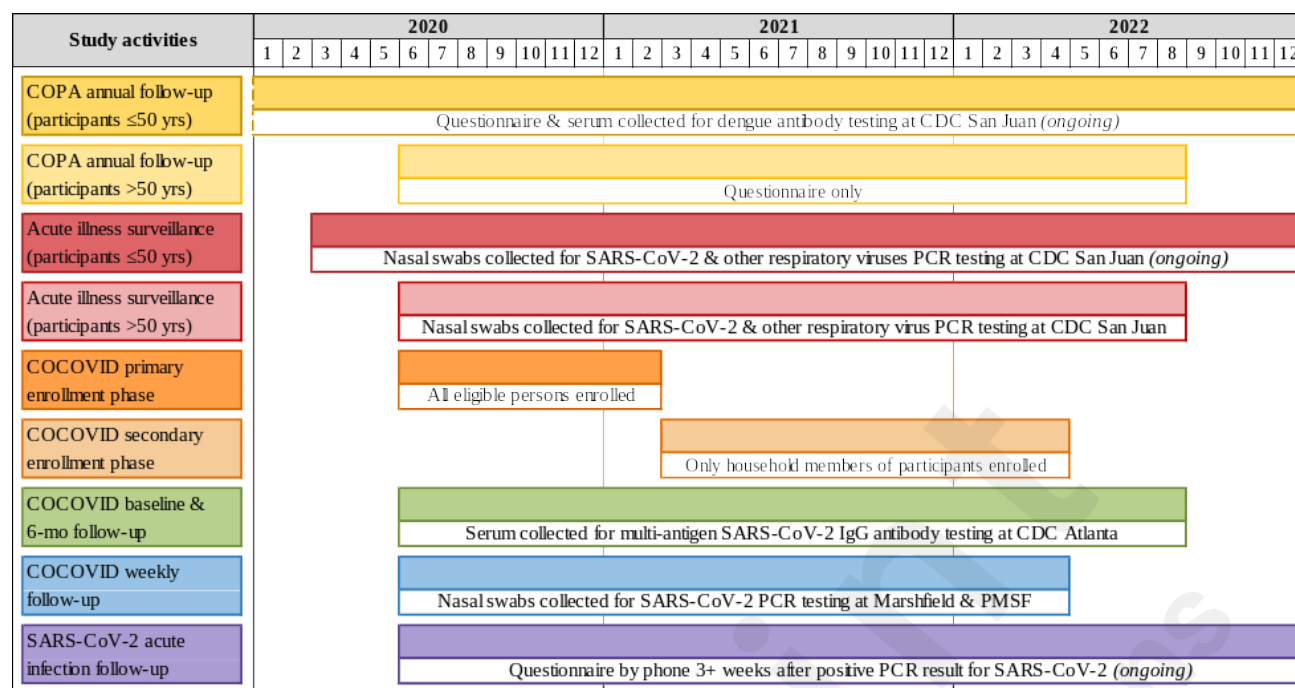


Figure 2. Timeline of key COPA and COCOVID cohort study activities in Ponce, Puerto Rico, January 2020–December 2022.



In April 2020, questions related to risk and perceptions of COVID-19 were added to the annual questionnaire and an acute illness surveillance (AIS) component was implemented (Figure 3). The AIS component is ongoing and uses an automated text messaging system to contact adult participants weekly to inquire if any household members have experienced fever or other COVID-like symptoms (i.e., cough, shortness of breath, sore throat, body pain, diarrhea, and loss of taste and/or smell) in the past 7 days (Figure 2). Participants that report fever are offered testing for dengue virus by reverse transcriptase polymerase chain reaction (RT-PCR) testing in serum, and those that report any qualifying symptom are offered an anterior nasal swab for SARS-CoV-2 and other respiratory virus RT-PCR testing. Beginning in December 2021, participants were also asked to provide an additional anterior nasal swab for SARS-CoV-2 rapid antigen testing. To ensure consistency of sample collection method for both tests, all nasal swabs are collected by a nurse or medical technologist. At the time of specimen collection, project staff also administer an AIS questionnaire on recent

exposures, symptoms, and health care seeking behaviors.

Study population and enrollment

Enrollment activities for the COCOVID sub-study for COVID-19 began in late June of 2020. We initially selected 12 of the 38 COPA study communities for COCOVID activities, and an additional three community areas were added during the enrollment process to meet sample size goals (Figure 1). We prioritized selection of community areas with high numbers of individuals aged under 18 and over 50 years in COPA-participating households to increase participation of these age groups. We also considered logistical priorities, including proximity to other selected clusters and ease of access to residential areas, in the selection process. Eligibility for COCOVID participation was expanded beyond that of COPA to include residents older than 50 years of age, in addition to spending four or more nights per week at a residence in a selected community area and having no definite plans to move in the next 12 months. COCOVID participants agree to participate in COPA annual follow-up questionnaires and, for participants ages 1–50 years, collection of blood specimens for arbovirus antibody testing for at least the sub-study duration; participation in the AIS activities is not required for COCOVID enrollment.

Using a standardized script, study staff contacted households with one or more COPA participants by phone to explain the study, offer enrollment to eligible household members, and schedule an initial study visit to complete consent processes and data and specimen collection activities. Enrollment of all eligible household members was encouraged but not required, and participants were also asked to refer family and neighbors residing in the 15 community areas. The primary enrollment phase occurred during June 2020–February 2021. Subsequently, only enrollment of members of households with one or more active participants was permitted to avoid major changes to study logistics and increases in resource use (secondary enrollment phase, Figure 3). Following the completion of the primary enrollment phase, we planned to continue participant weekly follow-up activities for a minimum of 12 months (through February 2022) and 6-month activities for a minimum of 18 months

(through August 2022).

Sample size calculation

Sample size for the COCOVID sub-study was based on the primary study objective of estimating community incidence of SARS-CoV-2 infection, which included allowing for comparisons by age group, accounting for potential similarity among participants within cluster areas, and assuming a cumulative incidence of infection of 10–30% over the study period. Age groups were defined to allow infection risk comparisons of pediatric (1–17 years) and older adult (≥ 50 years) populations with the general population (18–49 years); gender-based sample size estimates were not made as we assumed our household-based recruitment methods would result in representative samples by gender. Our calculations assumed enrollment of approximately equal numbers of participants across the three age groups and from the initially selected 12 community areas. Additionally, it was assumed that participants who were infected with SARS-CoV-2 prior to enrollment would not be at lower risk of further infection during the study period.

We expected a high coefficient of variation (COV) among the clusters for SARS-CoV-2 transmission, and calculations were made using a range of values from 0.5–0.8. We also expected high intra-cluster correlation among the responses within a cluster due to SARS-CoV-2 rapid spread, and calculations were made using values of 0.2 and 0.25. Precision of incidence is estimated computing 95% confidence intervals (CI) for one proportion in a cluster sampling design. Scenarios were run considering different cumulative incidence proportions, with different intra-cluster correlation, COV, and sample sizes for each cluster. Calculations showed that to reach a CI half width of 0.19 or less, a sample size of 215 participants per age group or a total of 645 participants was needed. To account for up to 20–30% loss to follow-up (high estimate to also allow for probable unequal recruitment across age groups) and an estimated 5% prevalence of prior SARS-CoV-2 infection at enrollment, we estimated that we needed to enroll a total of 900–1,000 participants.

Baseline and 6-month follow-up

Baseline and subsequent 6-month visits took place at the COPA central office facilities, a designated outdoor drive-thru site, or the participant home depending on local social distancing regulations and participant preferences. After completing informed consent processes, study staff administered a questionnaire on chronic conditions, recent COVID-19 exposures, and preventive behaviors, and, for participants that reported an acute COVID-19-like illness in the past 7 days, symptoms and health care seeking behaviors. Questions on COVID-19 vaccination status were added in January 2021, and questions on employment status and category, long-term COVID-19 symptoms (persisting or appearing more than one month after illness onset), and diagnosis of new chronic conditions after an acute SARS-CoV-2 infection were added in September 2021. A trained phlebotomist collected blood for multi-antigen SARS-CoV-2 IgG antibody testing by venipuncture using aseptic techniques. Approximately 7.5 milliliters of blood were collected from participants 1–4 years and ~15 milliliters from participants 5 years and older. Blood was centrifuged to serum within 4 hours of collection. To reduce study burden, COPA annual activities were conducted at the same time as COCOVID baseline and 6-month follow-up visits where possible (Figure 2). In these cases, an additional ~7.5 milliliters of blood for dengue antibody testing were collected during the COCOVID visit for participants ages 1–50 years, and the COPA annual questionnaire was administered to all participants by phone. This questionnaire included questions about household composition and income, demographics, employment status, and COVID-19 risk perception. The 6-month follow-up activities were conducted through August 2022 (Figure 3).

Weekly follow-up

Each week, participants self-collected (or had parent or guardian assist them in collecting) a flocked swab of the anterior nares (anterior nasal swab) and completed a paper questionnaire on new or worsening symptoms, illness among household member(s), and behaviors such as leaving their home for activities with people outside of their household, places visited, mask use, and social distancing.

Prior to distribution, a unique barcode label was applied to a specimen collection tube and corresponding weekly questionnaire, and questionnaires were pre-filled with the participant's name, date of birth, and study identifier. At the enrollment appointment, the first weekly nasal swab specimen and accompanying questionnaire were completed under study staff instruction and supervision, and an initial 4-week supply of paper questionnaires and nasal swab specimen collection supplies were provided. Participants were asked to collect nasal swab specimens and fill out the accompanying questionnaire on a consistent day of the week within 0–12 hours of a scheduled transfer to study staff.

Study staff drove to the participating houses at an agreed upon day and time for specimen and questionnaire pick-up each week. Households that were unavailable for scheduled pick-up (incidentally or always) were asked to drop-off collected specimens and questionnaires at an agreed upon day/time at the COPA office site. An automated text-messaging system via Twilio software was used to send reminders to households on the day of their scheduled pick-up or drop-off for respiratory specimen collection, and study staff contacted participants by phone within 0–2 days if specimen and questionnaire receipt was delayed without notice and to resolve other issues related to received specimens and questionnaires, such as specimen leaks and missing responses. If scheduled receipt of a weekly respiratory specimen was delayed by 4 or more days, no additional opportunities for specimen transfer were offered, and the specimen was considered missing for that week. Additional 4-week supplies of respiratory specimen collection kits and questionnaires were distributed to participants as needed during pick-up and drop-off activities.

Weekly follow-up activities were initially planned to conclude in February 2022, 12 months after the primary enrollment period ended. However, investigators agreed to extend the weekly follow-up period due to an increase in SARS-CoV-2 transmission following the introduction of the Omicron variant in Puerto Rico. Study materials and funds were identified to permit a 2-month extension of weekly follow-up to a total of 98 weeks from June 2020 through April 2022 (Figure 3).

Follow-up for participants with symptoms or SARS-CoV-2 infection

Participants that reported fever, cough, shortness of breath, sore throat, body pain, diarrhea, or loss of taste or smell in the weekly questionnaires (or through the AIS text message system) were asked to participate in a COPA AIS visit to collect additional samples for expedited testing for dengue, SARS-CoV-2, and other respiratory viruses and answer the AIS questionnaire.

Participants with a positive SARS-CoV-2 molecular test from nasal swab specimens collected during weekly COCOVID follow-up or as part of AIS were contacted by phone at least four weeks after specimen collection (depending on delays in receipt of positive test results) to complete a follow-up questionnaire on symptoms and illness duration, health care seeking behavior, travel and other exposures prior to infection, and taken measures to prevent household transmission while sick (Figure 2).

Specimen storage, transport, and testing

Participants were asked to store weekly anterior nasal swab specimens in their home refrigerators until they were transferred to study staff. Upon receipt of nasal swab specimens, staff stored them in the COPA laboratory refrigerator at 4°C or according to current CDC specimen storage guidelines. Staff centrifuged whole blood specimens within 4 hours of collection to separate out serum and stored them in the COPA laboratory refrigerator until transport. Specimens were transported by a courier three times weekly from the COPA office site to the CDC Dengue Branch (DB) laboratory in San Juan, Puerto Rico, for accessioning, processing, and storage at -70°C. Frozen nasal swab specimens from COCOVID weekly collection activities were shipped every 1–3 months and serum aliquots from COCOVID 6-month collection were shipped every 3–6 months to designated laboratories for testing.

Weekly nasal swab specimens were tested by multiplex real-time reverse transcriptase polymerase chain reaction (rRT-PCR) assays to detect SARS-CoV-2 RNA, including the Quidel Lyra® SARS-CoV-2 Assay, ThermoFisher Scientific TaqPath™ COVID-19 Combo Kit, ThermoFisher Scientific

TaqPath™ COVID-19, FluA, FluB Combo Kit, and Roche cobas® SARS-CoV-2 assay.[24-27] Cycle threshold (Ct) values for assay targets were provided for positive specimens and overall test result interpretations were provided for all tested specimens. Specimens collected during June 2020–February 2022 were tested by Marshfield Labs in Marshfield, Wisconsin; those collected during March–April 2022 were tested by the Ponce Medical Sciences Foundation (PMSF) Immunology Reference Laboratory in Ponce, Puerto Rico. Whole genome sequencing is ongoing and being performed at CDC headquarters (Atlanta, Georgia) or DB laboratories for weekly nasal swabs that tested positive by rRT-PCR and have sufficiently low Ct values for the N protein target (≤ 30 for headquarters and ≤ 27 for DB labs). Information on the detected variant lineage (via Pangolin COVID-19 Lineage Assigner software version 3.1.7) is provided in accordance with CDC guidance. [28]

Serum specimens collected for baseline and 6-month follow-up visits were tested by the Luminex xMAP SARS-CoV-2 Multi-Antigen Immunoglobulin G (IgG) Assay to detect past infection with SARS-CoV-2.[29] This assay has three antigen targets: the nucleocapsid (N) structural protein and the S1 subunit and the Receptor Binding domain (RBD), which are part of the spike (S) protein. Interim guidelines for SARS-CoV-2 testing from the U.S. CDC recommend using anti-N IgG antibody testing to evaluate prior SARS-CoV-2 infection in vaccinated individuals as COVID-19 vaccines are designed to encode the S protein or a portion of it. Thus, to reduce potential misclassification due to vaccination vs. past SARS-CoV-2 infection, specimens were only determined to be positive if IgG antibodies against the N protein in addition to those against at least one of the other two targets (S1 and RBD) were detected. The CDC Microbial Pathogenesis and Immune Response (MPIR) laboratory in Atlanta Georgia conducted testing for all serum specimens collected during the study period.

Serum and nasal swab specimens collected for AIS activities were tested at the CDC DB laboratory. Serum samples were tested by the CDC dengue virus (DENV)-1-4 rRT-PCR Multiplex Assay to detect RNA for dengue viruses 1, 2, 3, and 4 and by DENV Detect IgM Capture enzyme linked immunosorbent assay (ELISA) kit to detect IgM antibodies for DENV antigens.[30, 31] Anterior nasal swabs were tested by an in-house respiratory virus rRT-PCR multiplex assay to detect RNA for SARS-CoV-2, influenza A and B viruses, parainfluenza viruses, adenovirus, respiratory syncytial virus, and human metapneumovirus. The DENV and respiratory virus rRT-PCR assays for AIS specimens were run weekly or more often, and participants with positive results were notified by phone and given recommendations on preventing transmission and following-up with their healthcare providers as needed.

Data collection and management

Baseline and 6-month questionnaire

During June 2020–February 2022, the EpiInfo mobile application was used on Samsung tablets to administer the baseline and 6-month questionnaires and record responses, including participant identifiers and specimen collection data.[32] Beginning in March 2022, the Research Electronic Data Capture (REDCap) mobile application was used to administer these questionnaires.[33, 34] Paper questionnaire forms were available for use as a backup in case of electronic data collection failure. Paper forms were entered into the mobile app either daily or every 6 months as part of data review and cleaning processes. Questionnaire data were transferred daily from tablets to a secure, limited access CDC server via SSH File Transfer Protocol.

Weekly questionnaire

A paper version of the weekly questionnaire was completed by participants and received by study staff at scheduled specimen pick-up and drop-off activities. Staff reviewed the form for completeness and filled out the laboratory section to accompany the corresponding weekly nasal swab specimen. After the specimens were transported to and accessioned by the CDC DB laboratory, the forms were

scanned into the CDC laboratory samples database where they underwent automated data extraction using Microsoft Azure Form Recognizer software or manual data entry in cases of software extraction failure.[35]

Preliminary data cleaning was performed in EpiInfo or REDCap daily by study field staff and data quality team. Additional data cleaning processes were performed weekly using the R environment for statistical computing and involved data validation and standardization, including identification of missing forms, data inconsistencies, and variable completeness.[36] Final data checks and standardizations were also performed using R statistical software at 6-month intervals to generate analytic datasets.

Data analysis

Outcome measures

The primary outcome measures of interest for this study are (1) the incidence of SARS-CoV-2 infection as defined by detection of SARS-CoV-2 RNA by PCR testing of participant specimens collected as part of weekly follow-up or acute illness surveillance visits and (2) relative risks of incident SARS-CoV-2 infection by participant demographics, health history, and behaviors. Secondary outcomes included the (1) seroprevalence of SARS-CoV-2 defined by detection of N and either S1 or RBD IgG antibodies by ELISA; (2) household attack rate among households with at least two participants and one participant with incident SARS-CoV-2 infection; (3) frequency of illness characteristics and outcomes (i.e., symptom status and severity, health care seeking, death, and post-COVID conditions) among participants with incident SARS-CoV-2 infections.

Planned statistical analyses

SARS-CoV-2 infection incidence and risk factors: Using the number of incident SARS-CoV-2 infection events detected through PCR testing divided by the total person-time at risk contributed, we calculated SARS-CoV-2 infection incidence rates overall and approximately every 6 months during the study period (June 2020–August 2022). We will also estimate the incidence rates separately for

primary and secondary infections, during circulation of major SARS-CoV-2 variants, and by COVID-19 vaccination status. We will use Poisson regression to estimate adjusted infection risks by age group, primary SARS-CoV-2 variant circulating at time of infection, prior SARS-CoV-2 infection status, and COVID-19 vaccination status and identify other potential risk factors for infection in our cohort. Household will be accounted for as a clustering effect in the regression, using the sandwich estimator framework to calculate a clustered standard error.[37] We will extrapolate cumulative incidence of SARS-CoV-2 infections in the cohort to the general population in the Ponce region during the study, using population age and sex estimates from U.S. census data numbers, and compare this extrapolated estimate to surveillance data from the same time period.

SARS-CoV-2 infection seroprevalence: Using overall interpretation for the SARS-CoV-2 IgG antibody testing (positive only when N and either S1 or RBD antibodies detected), we will estimate the seroprevalence of SARS-CoV-2 infection among participants of the COCOVID cohort approximately every 6 months during the study period, using specimens collected during those periods. The frequency of seroreversion, defined as a positive SARS-CoV-2 IgG antibody test result following a negative result, will be assessed over the study period at the individual and household level. For the period during which both PCR and antibody testing were occurring (June 2020–April 2022), we will compare the proportion of participants with seroconversion detected with SARS-CoV-2 incidence estimates from weekly PCR testing. This comparison will be used to evaluate the sensitivity and specificity of repeated serology testing at 6-month intervals in estimating infection incidence. We will also use results of this comparison to inform estimates of SARS-CoV-2 infection incidence in the cohort that occur during May–August 2022 when only SARS-CoV-2 IgG test results are available.

Household attack rate: We will assess potential household transmission of SARS-CoV-2 among households with at least two participating members and one participating member with an incident SARS-CoV-2 infection (index case) during study follow-up. Additional household cases will be

defined based on detection of SARS-CoV-2 RNA by PCR testing in a household member within 14 days of the index case. We will calculate a crude household attack rate from the proportion of additional cases among all household members deemed susceptible to acquiring SARS-COV-2 infection from an index case. To account for household transmissions from non-index cases and identify factors associated with household transmission, we will consider use of a chain binomial transmission model to estimate the household attack rate and identify factors associated with increased and decreased likelihood of intrahousehold transmission.[38, 39]

Frequency of illness characteristics and outcomes: Among COCOVID participants with incident infections, we will assess the frequency of symptomatic vs. asymptomatic infection, self-reported symptoms, illness duration, health care seeking behaviors and outcomes, including uncomplicated recovery, hospitalization, death, and development of post-COVID conditions. If sample size permits, we propose using robust Poisson regression models to identify participant demographics, symptoms, and comorbidities associated with seeking care and being diagnosed with COVID-19, severe disease, and development of self-reported post-COVID conditions.[40] We will compare acute illness events reported in COCOVID weekly questionnaires associated with an incident SARS-CoV-2 infection to those not associated with SARS-CoV-2 infection to identify symptoms and symptom combinations that are highly indicative of infection.[41]

Statistical analyses of baseline data

In this report, we describe the baseline characteristics of the COCOVID cohort, using data collected from questionnaires applied at COCOVID enrollment or within the previous year for participants previously enrolled in COPA. Frequency statistics are presented overall for the COCOVID cohort and compared to the populations of Ponce and Puerto Rico based on data from the U.S. Census American Community Survey 2021 1-Year estimates.[10, 42] Pearson chi-square tests were used for comparisons of the cohort population with those of Ponce and Puerto Rico at a household and individual level; multiple comparisons were conducted for variables with more than two categories

and Bonferroni adjusted P-values are presented. For the cohort population, individual frequency statistics were also presented by serological evidence of previous SARS-CoV-2 infection at baseline as indicated by IgG antibody testing by timing of study enrollment, age group, sex, and annual household income. Pearson chi-square or Fisher exact tests, in the case of one or more expected values <5 , were used to compare categorical variables by baseline serostatus. A significance level of 0.05 was assumed for all comparisons, including following Bonferroni adjustment where applicable. All analyses were conducted using R (version 4.0.4).[36]

Results

During the first phase of enrollment (June 2020–February 2021), 1,012 individuals from 388 households were enrolled in the COCOVID cohort; an additional 18 members of participating households were enrolled in the secondary enrollment phase with the last participant enrolled in November 2021. Among 388 households with at least one COCOVID participants, 83% of all household members were enrolled with a mean of 3.2 participants per household; 40.3% of households had an annual income under \$20,000 (Table 1). Of the 1,030 total participants, 56.2% had previously participated in the COPA study, mean age was 35.9 years (range 1–97 years), and 53.4% were female. Most participants identified as Latino (99.6%) and Puerto Rican (99.1%), and around two-thirds (67.1%) identified as white. There were 719 adult participants (21 years of age or older) (69.8%) enrolled, of which approximately two-thirds had college or technical school degrees (58.8%) or post-graduate study (7.5%) and half were employed (53.0%). More than half (54%) of adult participants reported having been diagnosed with at least one chronic health condition.

Table 1. Household and individual characteristics of COCOVID cohort study population at baseline, June 2020–November 2021, compared to those Ponce municipality and Puerto Rico populations according to 2020 and 2021 U.S. Census American Community Survey estimates.

	Households	Households in	P-	Households	P-
--	------------	---------------	----	------------	----

	with ≥ 1 COCOVID participant (N = 388)	Ponce municipality (N = 50,007)	<i>value^b</i>	in Puerto Rico (N = 1,165,982)	<i>value^b</i>
Annual income for household, n (%)^a					
<\$10,000	61 (16.3)	16,659 (33.3)	<.001	286,499 (24.6)	<.001
\$10,000–\$19,999	90 (24)	9,783 (19.6)	.155	250,325 (21.5)	1
\$20,000–\$29,999	86 (22.9)	6,792 (13.6)	<.001	176,614 (15.1)	<.001
\$30,000–\$49,999	90 (24)	7,719 (15.4)	<.001	208,722 (17.9)	.01
>\$50,000	48 (12.8)	9,054 (18.1)	.04	243,822 (20.9)	<.001
	COCOVID study participants (N = 1,030)	Ponce municipality residents (N = 135,084)	<i>P- value^b</i>	Puerto Rico residents (N = 3,263,584)	<i>P- value^b</i>
Age group, n (%)					
1–9 years	75 (7.3)	11,105 (8.2)	1	254,038 (7.8)	1
10–17 years	174 (16.9)	12,603 (9.3)	<.001	291,750 (8.9)	<.001
18–24 years	119 (11.6)	12,823 (9.5)	.20	316,387 (9.7)	.35
25–34 years	129 (12.5)	17,660 (13.1)	1	402,204 (12.3)	1
35–44 years	181 (17.6)	14,827 (11)	<.001	400,607 (12.3)	<.001
45–54 years	198 (19.2)	15,391 (11.4)	<.001	415,974 (12.7)	<.001
55–64 years	52 (5.1)	17,600 (13)	<.001	442,138 (13.5)	<.001
≥ 65 years	102 (9.9)	330,75 (24.5)	<.001	740,486 (22.7)	<.001
Female, n (%)	550 (53.4)	70,555 (52.2)	.47	1,719,593 (52.7)	.67
Identify as Latino/a, n (%)^a	1021 (99.6)	134,560 (99.6)	.80	3,239,060 (99.2)	.27
Race, n (%)^a					
White	682 (68.9)	45,980 (34)	<.001	914,840 (28)	<.001
Black	102 (10.3)	5,171 (3.8)	<.001	200,391 (6.1)	<.001
Other	47 (4.7)	13,477 (10)	<.001	981,183 (30.1)	<.001
Multiple races	159 (16.1)	70,456 (52.2)	<.001	1167,170 (35.8)	<.001

	COCOVID study participants aged ≥25 years (N = 662)	Ponce municipality residents aged ≥25 years (N = 98,553)	<i>P- value^b</i>	Puerto Rico residents aged ≥25 years (N = 2,401,409)	<i>P- value^b</i>
Educational attainment, n (%)^a					
Less than high school graduate	38 (5.7)	19,235 (19.5)	<.001	488,780 (20.4)	<.001
High school graduate or equivalent	167 (25.3)	26,707 (27.1)	1	671,940 (28)	.13
Some college or associate degree	186 (28.2)	21,209 (21.5)	<.001	556,795 (23.2)	.01
Bachelor's degree or higher	268 (40.7)	31,402 (31.9)	<.001	683,894 (28.5)	<.001
Employment status, n (%)^a					
Employed	356 (55.3)	35,976 (36.5)	<.001	977,860 (40.7)	<.001
Unemployed	74 (11.5)	6,037 (6.1)	<.001	119,812 (5)	<.001

^a Participants or households who declined to respond or were missing a response were excluded from the denominator in percentage calculations (<4%).

^b P-values for characteristics with more than two categories were adjusted for multiple comparisons by Bonferroni correction.

Compared to the overall populations of both Ponce and Puerto Rico, the COCOVID cohort population differed significantly by annual household income, age group, race, education level, and employment status (Table 1). Households with at least one COCOVID participant more frequently belonged to middle income groups of \$20,000–29,999 (22.9% in cohort vs. 13.6% in Ponce and 15.1% in Puerto Rico) and \$30,000–49,000 (24% vs. 15.4 % and 17.9%) and less frequently highest, >\$50,000 (12.8% vs. 18.1% and 20.9%), and lowest, <\$10,000 (16.3% vs. 33.3% and 24.6%), income groups ($P<.05$). The cohort had relatively high proportions of participants in 10–17 (16.9% in cohort vs. 9.3% in Ponce and 8.9% in Puerto Rico), 35–44 (17.6% vs. 11% and 12.3%), and 45–54 (19.2% vs. 11.4% and 12.7%) years age groups and low proportions of participants in the oldest age groups, 55–64 (5.1% vs. 13% and 13.5%) and ≥65 years (9.9% vs. 24.5% and 22.7%) ($P<.001$). Cohort participants more frequently identified as white (68.9% in cohort vs. 34% in Ponce and 28% in Puerto Rico) or black (10.3% vs. 3.8% and 6.1%) and less frequently as another race (4.7% vs.

10.0% and 30.1%) or mixed race (16.1% vs. 52.2% and 35.8%) ($P<.001$). For populations 25 years and older, cohort participants were less likely to have a less than high school education (5.7% in cohort vs. 19.5% in Ponce and 20.4% in Puerto Rico) and more likely to have some college or associate degree (28.2% vs. 21.5% and 23.1%) and a bachelor's degree or higher (40.7% vs. 31.9% and 28.5%) ($P<.001$). Additionally, cohort participants 25 years or older were more likely to be in the labor force relative to Ponce and Puerto Rico residents in the same age group, including in employed (55.3% vs. 36.5 and 40.7%) and unemployed groups (11.5% vs. 6.1% and 5%) ($P<.001$).

At enrollment, 16 (1%) of COCOVID participants were seropositive for SARS-CoV-2, of which median age was 46 years (IQR 18.8), 56.3% were male, and half reported chronic health conditions (Table 2). There was a significant difference in the timing of study enrollment among participants with and without baseline seropositivity for SARS-CoV-2, with a higher proportion of seropositive participants recruited in later months of the study ($P=0.003$). No significant differences by age group, sex, or household income were observed among participants that were seropositive and those that were seronegative at baseline.

Table 2. Timing of study enrollment and demographics and of participants by serological evidence of previous SARS-CoV-2 infection at baseline, COCOVID study, Ponce, Puerto Rico, June 2020–November 2021 (N = 1,030).

	Participants seropositive for SARS-CoV-2 at baseline (N = 16)	Participants seronegative for SARS-CoV-2 at baseline (N = 1,014)	Prevalence of seropositivity for SARS-CoV-2 at baseline	<i>P-value</i>
Timing of study enrollment, n (%)				
Jun–Aug 2020	0 (0)	324 (32)	0	.003
Sep–Nov 2020	5 (31.3)	405 (39.9)	1.2	
Dec 2020–Feb 2021	10 (62.5)	268 (26.4)	3.6	
Mar 2021 or later	1 (6.3)	17 (1.7)	5.6	
Age group, n (%)				
1–17 years	1 (6.3)	248 (24.5)	0.4	.24
18 to 34 years	4 (25)	244 (24.1)	1.6	
35 to 49 years	5 (31.3)	301 (29.7)	1.6	

50 to 64 years	4 (25)	121 (11.9)	3.2	
≥65 years	2 (12.5)	100 (9.9)	1.9	
Sex, n (%)				
Female	7 (43.8)	543 (53.6)	1.3	.46
Male	9 (56.3)	471 (46.5)	1.9	
Annual income for household, n (%) ^a				
<\$10,000	3 (18.8)	172 (17)	1.7	.12
\$10,000–\$19,999	5 (31.3)	208 (20.6)	2.3	
\$20,000–\$29,999	0 (0)	234 (23.2)	0	
\$30,000–\$49,999	5 (31.3)	241 (23.9)	2	
>\$50,000	3 (18.8)	133 (13.2)	2.2	

^a Participants that declined to respond were excluded from the denominator in percentage calculations (<3%).

Discussion

Few studies with prospective assessment of incident SARS-CoV-2 infections at the community level have been described to-date, and none have been implemented in Puerto Rico or other primarily Latino or Hispanic U.S. population.[17-19, 43-45] The COCOVID community-based cohort study provides a unique opportunity to assess SARS-CoV-2 infection incidence, risk factors, and outcomes in Puerto Rico and also help further characterize asymptomatic infections, household transmission, and serological and molecular testing dynamics. We were able to build the COCOVID study into an existing cohort study platform, allowing for rapid enrollment and implementation of extensive data collection activities. Prospective follow-up via weekly PCR testing of participants and regular serology testing enabled comprehensive detection of SARS-CoV-2 infections within the cohort population. Additionally, our large sample size and extended follow-up period strengthened our assessment of SARS-CoV-2 incidence and outcomes by time-dependent factors, including lockdown measures, circulating SARS-CoV-2 variants, and COVID-19 vaccination status.

COCOVID surpassed its enrollment goal with 1,030 participants with minimum sample size goals (n=215) reached in the three age groups of interest (1–17 years old, 18–49 years old, and ≥50 years). The cohort included participants with characteristics associated with increased COVID-19 morbidity risk,[3, 46, 47] although sometimes with lower frequency than the general populations for Ponce and

Puerto Rico. Consistent with these populations, almost all COCOVID participants identified as Latino/a. Chronic health condition prevalence was high among adults in the cohort and similar to estimates from the Behavioral Risk Factor Surveillance System.[48] The cohort had relatively low proportions of participants in the oldest (65+ years) age group and households in the lowest (<\$10,000) annual income group compared to Ponce and Puerto Rico populations. However, these groups still accounted for at least 10% of cohort participants and household, respectively, which should facilitate adjusted analyses and extrapolation of results to the general population. SARS-CoV-2 seroprevalence at enrollment was low and increased among cohort participants enrolled later in the study period. Low but increasing seroprevalence of SARS-CoV-2 observed among study participants at baseline over time is consistent with surveillance data from the region.[14] As a result, there was a limited sample size to assess differences in SARS-CoV-2 seroprevalence by age group and other characteristics, and factors associated with seroprevalence of SARS-CoV-2 infection will be explored more fully using follow-up data from the study.

We did not use a formal sampling frame but rather our sampling strategy was determined by logistic constraints to include a limited selection of community areas with primarily single residence homes and no public housing or large apartment buildings. Additionally, most of our participants were previously enrolled in the larger COPA cohort, whose eligibility was restricted to participants under 50 years of age. This likely led to the observed oversampling of larger, middle-income households and employed, middle-aged adults and their children relative to the overall populations of Ponce and Puerto Rico. Additionally, individuals who agreed to enroll in COCOVID (and COPA) may be more concerned about disease risks and more likely to take measures to prevent SARS-CoV-2 infection than the general populations of Ponce and Puerto Rico. However, there was diversity across sociodemographic indicators in our study population, and the detailed data collected on household and individual characteristics and preventive behaviors can be used to adjust for these potential

impacts on SARS-CoV-2 related outcomes and extrapolate our findings to the general population, where possible.

Despite tremendous progress in our understanding of SARS-CoV-2 transmission and control, including the development of effective vaccines, the persistence of health inequities and other challenges underpin the need for more comprehensive data to inform mitigation measures and prevent future outbreaks. The COCOVID study provides a valuable opportunity to investigate community and household SARS-CoV-2 infections in a primarily Hispanic/Latino population with diversity across several COVID-19 related risk factors. Study methodology supports thorough detection of asymptomatic and symptomatic infections to help us better estimate the burden of SARS-CoV-2 in the region and assess the limitations of surveillance, which currently remains the primary resource for SARS-CoV-2 data for the island.

Acknowledgements

The authors would like to acknowledge the hard work of our study field team, including Emma Ramírez, Suheil Albizu, Lismarie Santiago Vargas, Aimara Medina, Evelynette Pérez, Juan Santiago, Kasey Kinzel, Nicole León, Wilmarie Rivera, Jared Abraham, Aurora Castillo, and José Molina, and the support of CDC staff and deployers in conducting participant recruitment and data collection and cleaning activities. We would like to thank Ronald Cancel Santiago, Cindia Díaz, and Oscar Padro for designing and implementing critical data management systems, including those for weekly visit tracking and automated data entry of paper questionnaires.

Authors affiliated with Ponce Health Sciences University/Ponce Research Institute wish to provide acknowledgement to grant number U01CK000580.

Data availability

Due to data security and confidentiality guidelines determined by the study governance board, restricted use datasets must be requested by submitting a concept proposal to the corresponding author or to dengue@cdc.gov.

Conflicts of interest

None declared.

Abbreviations

AIS: Acute Illness Surveillance

CDC: United States Centers for Disease Control and Prevention

COCOVID: COPA COVID-19 study

COPA: Communities Organized to Prevent Arboviruses study

COV: coefficient of variation

DB: Dengue Branch (CDC)

DENV: dengue virus

ELISA: enzyme-linked immunosorbent assay

IgG: immunoglobulin G

IgM: immunoglobulin M

MPIR: Microbial Pathogenesis and Immune Response (CDC)

PCR: polymerase chain reaction

PMRI: Ponce Medical Research Institute

REDCap: Research Electronic Data Capture

RNA: ribonucleic acid

rRT-PCR: real-time reverse transcription polymerase chain reaction

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2



References

1. Bottan N, Hoffmann B, Vera-Cossio D. The unequal impact of the coronavirus pandemic: Evidence from seventeen developing countries. *PLoS One*. 2020;15(10):e0239797. PMID: 33027272. doi: 10.1371/journal.pone.0239797.
2. Sepulveda ER, Brooker AS. Income inequality and COVID-19 mortality: Age-stratified analysis of 22 OECD countries. *SSM Popul Health*. 2021 Dec;16:100904. PMID: 34584934. doi: 10.1016/j.ssmph.2021.100904.
3. Mackey K, Ayers CK, Kondo KK, Saha S, Advani SM, Young S, et al. Racial and Ethnic Disparities in COVID-19-Related Infections, Hospitalizations, and Deaths : A Systematic Review. *Ann Intern Med*. 2021 Mar;174(3):362-73. PMID: 33253040. doi: 10.7326/m20-6306.
4. Centers for Disease Control and Prevention. COVID Data Tracker: Cases, Deaths, & Testing. [cited 2022 April 20]; Available from: https://covid.cdc.gov/covid-data-tracker/#trends_totalcases.
5. Mistry P, Barmania F, Mellet J, Peta K, Strydom A, Viljoen IM, et al. SARS-CoV-2 Variants, Vaccines, and Host Immunity. *Front Immunol*. 2022;12:809244-. PMID: 35046961. doi: 10.3389/fimmu.2021.809244.
6. Tian D, Sun Y, Xu H, Ye Q. The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. *J Med Virol*. 2022 Jun;94(6):2376-83. PMID: 35118687. doi: 10.1002/jmv.27643.
7. Olney AM, Smith J, Sen S, Thomas F, Unwin HJT. Estimating the Effect of Social Distancing Interventions on COVID-19 in the United States. *American journal of epidemiology*. 2021;190(8):1504-9. PMID: 33406533. doi: 10.1093/aje/kwaa293.
8. Fischer CB, Adrien N, Silguero JJ, Hopper JJ, Chowdhury AI, Werler MM. Mask adherence and rate of COVID-19 across the United States. *PLoS One*. 2021;16(4):e0249891. PMID: 33852626. doi: 10.1371/journal.pone.0249891.
9. Robinson E, Jones A, Lesser I, Daly M. International estimates of intended uptake and refusal of COVID-19 vaccines: A rapid systematic review and meta-analysis of large nationally representative samples. *Vaccine*. 2021 Apr 8;39(15):2024-34. PMID: 33722411. doi: 10.1016/j.vaccine.2021.02.005.
10. U.S. Census Bureau. 2020 Decennial Census: Puerto Rico. 2020 [cited 2021 September 24]; Available from: <https://data.census.gov/cedsci/table?q=Puerto%20Rico&tid=DECENNIALPL2020.P1>.
11. Kaiser Family Foundation. Puerto Rico: Fast Facts. 2017 [cited 2021 September 1]; Available from: <https://www.kff.org/racial-equity-and-health-policy/fact-sheet/puerto-rico-fast-facts/>.
12. Perreira K, Lallemand N, Napoles A, Zuckerman S. Environmental scan of Puerto Rico's health care infrastructure. Urban Institute, 2017.
13. García C, Rivera FI, García MA, Burgos G, Aranda MP. Contextualizing the COVID-19 Era in Puerto Rico: Compounding Disasters and Parallel Pandemics. *J Gerontol B Psychol Sci Soc Sci*. 2021;76(7):e263-e7. PMID: 33112945. doi: 10.1093/geronb/gbaa186.
14. Departamento de Salud de Puerto Rico. COVID-19 en CIFRAS en Puerto Rico. 2023 [cited 2022 April 20]; Available from: <https://covid19datos.salud.gov.pr/>.
15. Volkman HR, Pérez-Padilla J, Wong JM, Sánchez-González L, Acevedo-Molina L, Lugo-Menéndez M, et al. Characteristics and clinical outcomes of patients hospitalized with laboratory-confirmed COVID-19-Puerto Rico, March-August 2020. *PLoS One*. 2021;16(12):e0260599. PMID: 34855817. doi: 10.1371/journal.pone.0260599.
16. Jones JM, Stone M, Sulaeman H, Fink RV, Dave H, Levy ME, et al. Estimated US Infection- and Vaccine-Induced SARS-CoV-2 Seroprevalence Based on Blood Donations, July 2020-May 2021. *JAMA*. 2021;326(14):1400-9. PMID: 34473201. doi: 10.1001/jama.2021.15161.
17. Dawood FS, Porucznik CA, Veguilla V, Stanford JB, Duque J, Rolfes MA, et al. Incidence Rates,

Household Infection Risk, and Clinical Characteristics of SARS-CoV-2 Infection Among Children and Adults in Utah and New York City, New York. *JAMA Pediatr.* 2022 Jan 1;176(1):59-67. PMID: 34623377. doi: 10.1001/jamapediatrics.2021.4217.

18. Karron RA, Hetrich MK, Na YB, Knoll MD, Schappell E, Meece J, et al. Assessment of Clinical and Virological Characteristics of SARS-CoV-2 Infection Among Children Aged 0 to 4 Years and Their Household Members. *JAMA Netw Open.* 2022 Aug 1;5(8):e2227348. PMID: 36044218. doi: 10.1001/jamanetworkopen.2022.27348.

19. McLean HQ, McClure DL, King JP, Meece JK, Pattinson D, Neumann G, et al. mRNA COVID-19 vaccine effectiveness against SARS-CoV-2 infection in a prospective community cohort, rural Wisconsin, November 2020 to December 2021. *Influenza Other Respir Viruses.* 2022 Jul;16(4):607-12. PMID: 35178857. doi: 10.1111/irv.12970.

20. Oran DP, Topol EJ. The Proportion of SARS-CoV-2 Infections That Are Asymptomatic : A Systematic Review. *Annals of internal medicine.* 2021;174(5):655-62. PMID: 33481642. doi: 10.7326/M20-6976.

21. Buitrago-Garcia D, Ipekci AM, Heron L, Imeri H, Araujo-Chaveron L, Arevalo-Rodriguez I, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: Update of a living systematic review and meta-analysis. *PLoS Med.* 2022 May;19(5):e1003987. PMID: 35617363. doi: 10.1371/journal.pmed.1003987.

22. Adams LE, Sánchez-González L, Rodriguez DM, Ryff K, Major C, Lorenzi O, et al. Risk factors for infection with chikungunya and Zika viruses in southern Puerto Rico: A community-based cross-sectional seroprevalence survey. *PLoS Negl Trop Dis.* 2022 Jun;16(6):e0010416. PMID: 35696355. doi: 10.1371/journal.pntd.0010416.

23. Sánchez-González L, Adams LE, Saavedra R, Little EM, Medina NA, Major CG, et al. Assessment of community support for Wolbachia-mediated population suppression as a control method for *Aedes aegypti* mosquitoes in a community cohort in Puerto Rico. *PLoS Negl Trop Dis.* 2021 Dec;15(12):e0009966. PMID: 34871301. doi: 10.1371/journal.pntd.0009966.

24. Quidel. Lyra® SARS-CoV-2 Assay (EUA). 2020; Available from: https://www.quidel.com/sites/default/files/product/documents/Lyra_BRM120002EN00.pdf.

25. ThermoFisher Scientific. TaqPath™ COVID-19 Combo Kit and TaqPath™ COVID-19 Combo Kit Advanced (EUA). 2020; Available from: <https://www.fda.gov/media/136112/download>.

26. ThermoFisher Scientific. TaqPath™ COVID-19, FluA, FluB Combo Kit. 2021; Available from: <https://www.fda.gov/media/145931/download>.

27. Roche. cobas® SARS-CoV-2. 2022; Available from: <https://www.fda.gov/media/136049/download>.

28. U.S. Centers for Disease Control and Prevention. Guidance for Reporting SARS-CoV-2 Sequencing Results. 2021 [updated June 23, 2021; cited 2021 November]; 3]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/reporting-sequencing-guidance.html>.

29. Luminex. xMAP® SARS-CoV-2 Multi-Antigen IgG Assay (EUA). 2020.

30. Centers for Disease Control and Prevention. CDC DENV-1-4 Real-Time RT-PCR Multiplex Assay. 2019 [updated 10/17/2019 07/22/2022]; Available from: <https://www.cdc.gov/dengue/healthcare-providers/testing/molecular-tests/realtime.html>.

31. Namekar M, Ellis EM, O'Connell M, Elm J, Gurary A, Park SY, et al. Evaluation of a new commercially available immunoglobulin M capture enzyme-linked immunosorbent assay for diagnosis of dengue virus infection. *J Clin Microbiol.* 2013 Sep;51(9):3102-6. PMID: 23824771. doi: 10.1128/jcm.00351-13.

32. Dean AA, TG; Sunki, GG; Friedman, R; Lantinga, M; Sangam, S; Zubieta, JC; Sullivan, KM; Brendel, KA; Gao, Z; Fontaine, N; Shu, M; Fuller, G; Smith, DC; Nitschke, DA; Fagan, RF. Epi Info™, a

database and statistics program for public health professionals. Atlanta, GA, USA: Centers for Disease Control and Prevention; 2011.

33. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*. 2009 2009/04/01/;42(2):377-81. doi: <https://doi.org/10.1016/j.jbi.2008.08.010>.

34. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*. 2019 2019/07/01/;95:103208. doi: <https://doi.org/10.1016/j.jbi.2019.103208>.

35. Microsoft Azure Form Recognizer. [September 1, 2022]; Available from: <https://azure.microsoft.com/en-us/services/form-recognizer/#overview>.

36. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.

37. Zeileis A, Köll S, Graham N. Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R. *Journal of Statistical Software*. 2020 10/07;95(1):1 - 36. doi: 10.18637/jss.v095.i01.

38. Sharker Y, Kenah E. Estimating and interpreting secondary attack risk: Binomial considered biased. *PLoS Comput Biol*. 2021 Jan;17(1):e1008601. PMID: 33471806. doi: 10.1371/journal.pcbi.1008601.

39. Li F, Li YY, Liu MJ, Fang LQ, Dean NE, Wong GWK, et al. Household transmission of SARS-CoV-2 and risk factors for susceptibility and infectivity in Wuhan: a retrospective observational study. *Lancet Infect Dis*. 2021 May;21(5):617-28. PMID: 33476567. doi: 10.1016/s1473-3099(20)30981-6.

40. Chen W, Qian L, Shi J, Franklin M. Comparing performance between log-binomial and robust Poisson regression models for estimating risk ratios under model misspecification. *BMC Medical Research Methodology*. 2018 2018/06/22;18(1):63. doi: 10.1186/s12874-018-0519-5.

41. Antonelli M, Capdevila J, Chaudhari A, Granerod J, Canas LS, Graham MS, et al. Optimal symptom combinations to aid COVID-19 case identification: analysis from a community-based, prospective, observational cohort. *medRxiv*. 2021:2020.11.23.20237313. PMID: 33269364. doi: 10.1101/2020.11.23.20237313.

42. U.S. Census Bureau. 2021: ACS 1-Year Estimates Puerto Rico Public Use Microdata Sample. 2021.

43. Lanata CF, Gil AI, Ecker L, Cornejo R, Rios S, Ochoa M, et al. SARS-CoV-2 infections in households in a peri-urban community of Lima, Peru: A prospective cohort study. *Influenza Other Respir Viruses*. 2022 May;16(3):386-94. PMID: 34962079. doi: 10.1111/irv.12952.

44. Miller EM, Law EA, Ajeen R, Karasik J, Mendoza C, Abernathy H, et al. SARS-CoV-2 infection in central North Carolina: Protocol for a population-based longitudinal cohort study and preliminary participant results. *PLoS One*. 2021;16(10):e0259070. PMID: 34695156. doi: 10.1371/journal.pone.0259070.

45. Brasil P, Damasceno L, Fuller T, Bastos LS, Cruz OG, Medeiros F, et al. Cohort-profile: Household transmission of SARS-CoV-2 in a low-resource community in Rio de Janeiro, Brazil. *BMJ Open*. 2022 Dec 7;12(12):e067212. PMID: 36600372. doi: 10.1136/bmjopen-2022-067212.

46. Dessie ZG, Zewotir T. Mortality-related risk factors of COVID-19: a systematic review and meta-analysis of 42 studies and 423,117 patients. *BMC Infect Dis*. 2021 Aug 21;21(1):855. PMID: 34418980. doi: 10.1186/s12879-021-06536-3.

47. Khanijahani A, Iezadi S, Gholipour K, Azami-Aghdash S, Naghibi D. A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *Int J Equity Health*. 2021 Nov 24;20(1):248. PMID: 34819081. doi: 10.1186/s12939-021-01582-4.

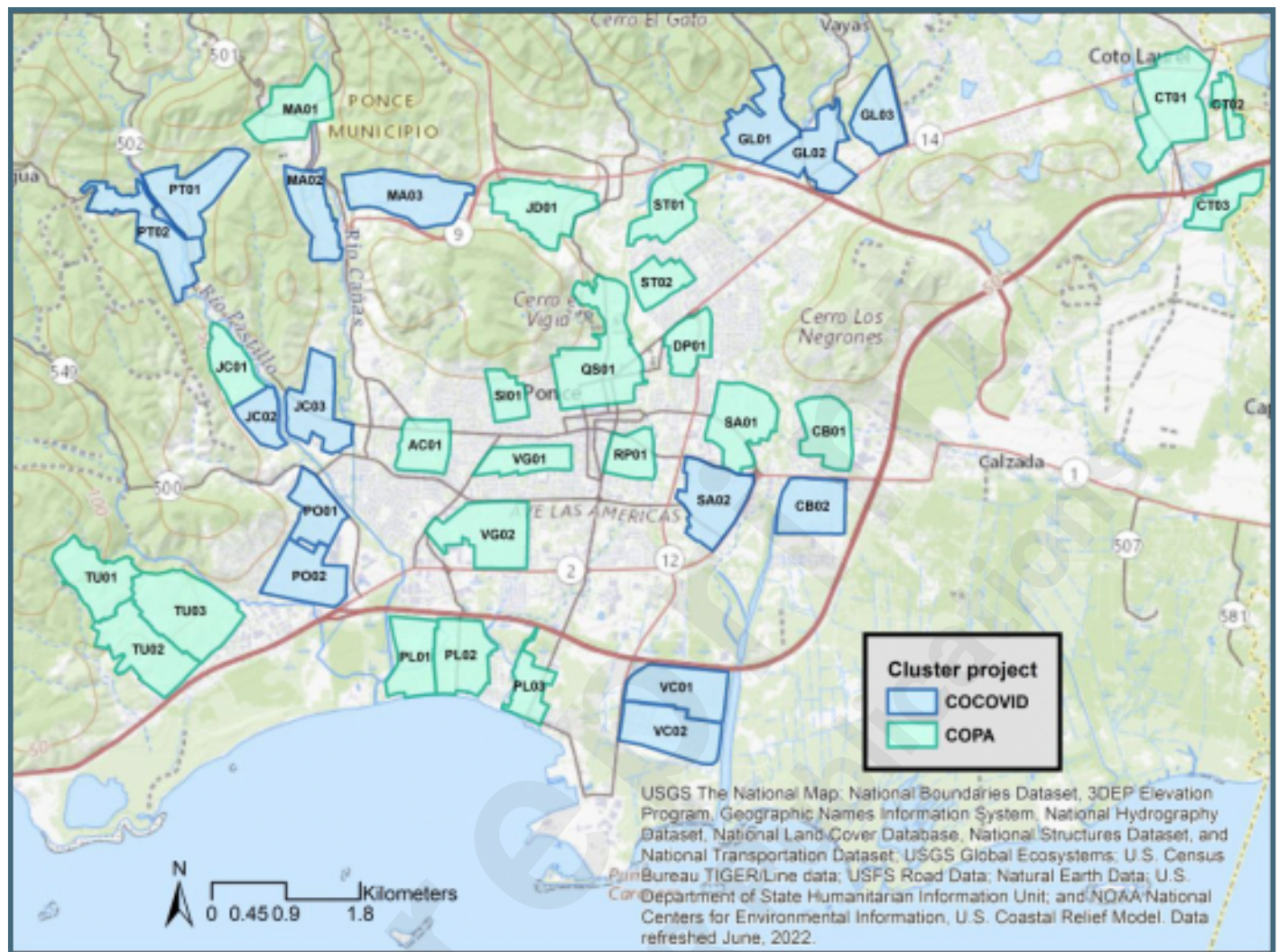
48. Centers for Disease Control and Prevention NCfCDPaHP, Division of Population Health,. BRFSS Prevalence & Trends Data [online]. 2020 [cited 2023 August 24]; Available from: <https://www.cdc.gov/brfss/brfssprevalence/>.



Supplementary Files

Figures

Map of the 38 COPA and 15 COCOVID community cohort study clusters areas in Ponce, Puerto Rico.



Timeline of key COPA and COCOVID cohort study activities in Ponce, Puerto Rico, January 2020–December 2022.

