

# **Evaluation of COVID-19 vaccine effectiveness in Manaus, Brazil: Test-negative case-control study protocol**

Version: 01 / March 16<sup>th</sup> 2021

## **Background**

Since the emergence of severe acute respiratory virus coronavirus 2 (SARS-CoV-2), Brazil has experienced one of the world's highest incidence and mortality rates in the world, with over 10 million reported infections as of the end of February 2021.<sup>1-3</sup> Of grave concern is the resurgence of COVID-19 cases and hospitalizations observed in Manaus, despite high estimated seroprevalence.<sup>4,5</sup> One hypothesis for this resurgence is loss of immunity to the P.1 Variant of Concern (VOC), which was first detected in Manaus on Jan 12, 2021, and now consists the majority of new infections, including all genomes sequenced in February 2021.<sup>6-9</sup> This lineage has accrued mutations associated with decreased neutralization.<sup>10,11</sup>

The rapid development of novel vaccines against COVID-19 allowed countries to start vaccine distribution programs within a year of the identification of the novel virus. Among the first vaccines to be developed and evaluated in Brazil was Sinovac's CoronaVac vaccine.<sup>12,13</sup> Phase III trials were conducted in Turkey and Brazil, and results were announced on February 5, 2021, in which effectiveness after 14 days following vaccination with 2 doses of vaccine was reported to be 50.65% for all symptomatic cases of COVID-19, 84% for cases requiring medical attention, and 100% for hospitalized, severe, and fatal cases.<sup>14</sup> CoronaVac was approved for emergency use on 17 January in Brazil, and healthcare workers in Manaus received vaccine starting on 19 January.

As vaccine programs continue, there has been much interest in estimation of vaccine effectiveness through observational studies, and specifically in settings where VOC are circulating. Such studies have advantages over clinical trials, including increased size and follow-up time, and reduced cost. However, as vaccinated and unvaccinated individuals are likely different in their SARS-CoV-2 risk and healthcare access, these studies must address bias through design and analysis. Israel has led the world in vaccination rates, and several studies have been conducted demonstrating the effectiveness of Pfizer's vaccine, in a context in which the B.117 variant is prevalent.<sup>15</sup> However, studies have not been conducted in a context in which the P.1 variant is prevalent, nor have observational effectiveness studies of CoronaVac.

We propose a test-negative case-control study<sup>16,17</sup> in healthcare workers (HCWs) residing in Manaus to evaluate the effectiveness of CoronaVac in preventing COVID-19 in a real-world setting and a site where the P.1 VOC is responsible for a majority of transmission. We will expand the study population to include other groups that are eligible for vaccination, including individuals eligible based on age, once reasonable vaccine coverage has been achieved in those populations; to estimate the vaccine effectiveness of the ChAdOx1 vaccine and potentially other vaccines, if they achieve reasonable coverage in the population following the campaign vaccination; and to include additional sites in Brazil. We will amend the protocol and its objectives accordingly to address these new questions. .

## **Local setting and context**

Manaus (3°5'S, 60°W) is the capital of the state of Amazonas, the major urban metropolis in the middle of the Amazon jungle and a major river port for seafaring vessels. In 2020, Manaus, with an estimated population of 2,219,580 inhabitants, reported 144,767 COVID-19 cases (incidence rate: 6,522

/100,000) and 7,605 deaths (mortality rate: 342/100,000). Manaus has 40 Family Health teams, 53 primary health care centers and 15 other health units under the responsibility of the Secretariat of Health of Manaus, and 20 private or public hospitals.

## **Objectives**

To estimate the effectiveness of CoronaVac against COVID-19 among healthcare workers in Manaus.

## **Methods**

### Study design and study period

We will conduct a retrospective matched case-control study, enrolling cases who test positive for SARS-CoV-2 and controls who test negative for SARS-CoV-2 amongst residents of Manaus as of January 19, 2021, the day COVID-19 vaccination was initiated in the city. We will identify cases and matched controls by reviewing health surveillance records and ascertain the type and date of vaccination by reviewing the state COVID-19 vaccination registry. In this design, the odds ratio of vaccination comparing cases and controls estimates the direct effect of vaccination on the disease outcome.

Our initial analyses, described below, will evaluate effectiveness of CoronaVac among healthcare workers (HCWs), which was the initial population to receive the vaccine. The State Secretary of Health of Amazonas (SES-AM) is administering two vaccines, CoronaVac and ChAdOx1, in its mass vaccination campaign, but our initial analyses will focus on evaluating effectiveness of CoronaVac since >97% of the vaccinations of HCWs at present have been performed with this vaccine. We will expand the scope of the analyses and evaluate the effectiveness of ChAdOx1 in HCWs if significant coverage with this vaccine is achieved and amend the protocol accordingly.

We will perform interim analyses aimed at evaluating the effectiveness of receiving at least one dose of the vaccine and a final analysis that will evaluate the effectiveness of receiving two doses of the vaccine.

### Data Sources and Integration

The study base consists of HCWs who are residents of the municipality of Manaus. We identified eligible individuals by using the HCWs registry (SEMSA) of SES-AM.

The integration of the cohort datasets with other systems was made by using CPF numbers (Brazilian citizens' unique identifier code). For those entries missing CPF in SEMSA, we performed a probabilistic record linkage between SEMSA and CADSUS (registration database of users of the public universal health system [SUS] in Brazil). For the probabilistic method we used the Reclink III software,<sup>18</sup> and we considered sex, the phonetic code of the first and last name and the phonetic code of the first name of the mother as blocking variables. We compared the similarity of the name, mother's name with a threshold of >85% similarity, and date of birth a threshold of >65% similarity. All pairs identified by the probabilistic method were manually revised.

Using these approaches, we obtained information on CPF for 67,718 (99%) of the eligible 68,698 HCWs in SEMSA registry.

We will extract information on healthcare utilization, including those associated with testing and clinical care for COVID-19, deaths, and type, date and order of vaccination from five databases:

- GAL: the laboratory testing registry of the network of public health laboratories)

- SIVEP-Gripe: the national surveillance database of severe acute respiratory illnesses created by Ministry of Health Brazil in 2009
- e-SUS: the national surveillance system of suspected cases of COVID-19 from mild to moderate "influenza like illness", created by the Ministry of Health Brazil in 2020
- SIM: the national mortality registry from the Ministry of Health
- SES-AM COVID-19 vaccination registry

Datasets from e-SUS will include information from telemonitoring protocols established by SES-AM, whereby a team of doctors and nurses make daily telephone calls to assess symptoms, to identify signs of severity and to refer the patient to a health unit.

We built a parent database using the MySQL language and integrated individual datasets using an application programming interface (API), which was developed using ElasticSearch. Table 1 in the Supplementary material lists the variables extracted from datasets and incorporated in the parent database. The database will be updated on a weekly basis.

Some variables were reported in multiple data bases. To define a single variable, we drew from each data base with priority given to data bases that were more complete, reliable, and up-to-date. We chose age from the data bases in the following order: CADSUS, SEMSA, GAL, e-SUS, SIVEP-Gripe; and we chose neighborhood from the data bases in the following order: CADSUS (CEP variable, transformed to neighborhood), CADSUS (*bairro* variable), SEMSA, and e-SUS.

## **IRB and Ethics Statement**

The protocol has been submitted to the Ethical Committee for Research of Federal University of Mato Grosso do Sul (CAAE: 43289221.5.0000.0021). The work of investigators at the University of Florida, Yale University, Stanford University, and University of Barcelona was conducted to inform the public health response and was therefore covered under Public Health Response Authorization under the US Common Rule.

## **Study Population**

Inclusion criteria:

- Healthcare worker as defined by the SEMSA/SES-AM registry,
- Residing in the city of Manaus,
- Age  $\geq 18$  years at the index date,
- With complete information, which is consistent between databases, on age, sex, and residential address defined by CEP (zip code)
- With complete and consistent vaccination status and dates.

Exclusion criteria:

- Not a healthcare worker as defined by the SEMSA/SES-AM registry,
- Not residing in the city of Manaus,
- Aged  $< 18$  years at the index date,
- With missing or inconsistent information on age, sex, or residential address defined by CEP (zip code)
- With incomplete or inconsistent vaccination status or dates.

### Case definition and eligibility

For both cohorts, we will use information from GAL/SIVEP-Gripe/e-SUS VE databases to identify eligible cases. For the primary analysis, cases are defined as eligible members of the study population (as defined above, Study Population) who:

- Had a sample with a positive SARS-CoV-2 RT-PCR, which was collected between January 19, 2021 and 7 days prior to database extraction of information
- Have complete and consistent data on SARS-CoV-2 PCR test result,
- Did not have a positive RT-PCR test in the preceding 90 days' period,
- Reported at least one COVID-19 related symptom whose onset was within 0-10 days before the date of their positive PT-PCR test result,
- Did not receive a dose of ChAdOx1 vaccine before the date of respiratory sample collection.

### Control definition and eligibility

We will use GAL/SIVEP-Gripe/e-SUS VE databases to identify eligible controls. For the primary analysis in both studies, controls are defined as eligible members of the study population who:

- Had a sample with a negative SARS-CoV-2 RT-PCR result, which was collected after January 19, 2021 and  $\pm 3$  days from collection of the positive RT-PCR sample of its matched case,
- Have complete and consistent data on SARS-CoV-2 PCR test result,
- Reported at least one COVID-19 related symptom whose onset was within 0-10 days after the date of their negative PT-PCR test result,
- Did not have a subsequent positive PCR test in the following 14-day period
- Did not receive a dose of ChAdOx1 vaccine before the date of respiratory sample collection.

An individual who fulfils the control definition and eligibility and later has a sample tested that fulfils the case definition and eligibility can be included in the study as both a case and a control. An individual who fulfils the control definition for multiple different sample collection dates can be included in the study as a control for each collection date, up to a maximum of three times.

### Matching

Test-negative controls will be matched 1:1 to the cases.

Matching factors will include variables that are anticipated to be causes of the likelihood of receiving the vaccine, risk of infection and likelihood of receiving PCR testing for SARS-CoV-2 (i.e. healthcare access and utilization) (see Figures 1-3):

- Residential address (neighborhood [*bairro*], which is identified based on the first 5 digits of 8 digit CEP),
- Age (categorized as  $<30$ ,  $\geq 30$  and  $<60$ , and  $\geq 60$ ),
- Window of  $\pm 3$  days between collection of RT-PCR positive respiratory sample for cases and collection of RT-PCR negative respiratory sample for controls. If the date of respiratory sample collection is missing, the date of notification of testing result will be used.

We will use the standard algorithms to conduct matching which include: 1) setting a seed, 2) locking the database, 4) creating a unique identifier for matching after random ordering, 5) implementing exact matching based on matching variables, sampling controls at random if more than one available per case within strata.

## Exposure definition

COVID-19 vaccination by type of vaccine and the following stratifications:

- Received the first vaccine dose, and not having received a second dose, in the following time periods relative to sample collection for their PCR test:
  - 0-13 days
  - ≥14 days
- Received the second dose in the following time periods relative to sample collection for their PCR test:
  - 0-13 days
  - ≥14 days

The reference group for all analyses will be individuals who had not received a first dose of vaccine by the date of respiratory sample collection. This includes individuals who have not received any vaccine doses as of the date of data base extraction, individuals who received CoronaVac after the date of sample collection, and individuals who received ChAdOx1 after the date of sample collection. Date of notification of the testing result will be used if the date of respiratory sample collection is missing. We will restrict our initial analyses to the evaluation of CoronaVac, but plan to expand these analyses to include the ChAdOx1 vaccine in subsequent analyses.

## Statistical design

Primary analysis:

- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of symptomatic COVID-19 in the period starting 14 days after administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

Secondary analyses:

*Interim analyses of primary endpoint: symptomatic COVID-19*

- Estimate the effectiveness of receiving at least one dose of CoronaVac in reducing the risk of symptomatic COVID-19 in the period starting 14 days after administration of the 1<sup>st</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

*Analysis of secondary exposures of primary endpoint: symptomatic COVID-19*

- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of symptomatic COVID-19 in the period 0-13 days after administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of symptomatic COVID-19 in the period between 14 days after administration of the 1<sup>st</sup> dose and receiving the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Bias indicator: Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of symptomatic COVID-19 in the period from 0-13 days after administration of the 1<sup>st</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

*Analysis of secondary endpoint: test positive*

- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of testing positive for SARS-CoV-2 regardless of symptoms in the period starting 14 days after administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of testing positive for SARS-CoV-2 regardless of symptoms in the period from 0-13 days after

administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

- Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of testing positive for SARS-CoV-2 regardless of symptoms in the period between 14 days after administration of the 1<sup>st</sup> dose and receiving the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Bias indicator: Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of testing positive for SARS-CoV-2 regardless of symptoms in the period 0-13 days after administration of the 1<sup>st</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

*Analysis of secondary endpoint: hospitalizations*

- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of hospitalization with COVID-19 in the period starting 14 days after administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Estimate the effectiveness of a two-dose regimen of CoronaVac in reducing the risk of hospitalization with COVID-19 in the period from 0-13 days after administration of the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of hospitalization with COVID-19 in the period between 14 days after administration of the 1<sup>st</sup> dose and receiving the 2<sup>nd</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.
- Bias indicator: Estimate the effectiveness of a one-dose regimen of CoronaVac in reducing the risk of hospitalization with COVID-19 in the period from 0-13 days after administration of the 1<sup>st</sup> dose, compared to individuals who received no doses of a COVID-19 vaccine.

We will use conditional logistic regression to estimate the odds ratio (OR) of vaccination among cases and controls, accounting for the matched design, and 1-OR provides an estimate of vaccine efficacy under the standard assumptions of a test-negative design. We will include the following covariates in the adjusted model, which we hypothesize are predictive of vaccination, the risk of SARS-CoV-2 infection and COVID-19 severity and healthcare access and utilization:

- Age as continuous variable
- Sex
- Occupation category
- Self-reported race/skin color
- Number of previous entries in e-SUS or SIVEP-Gripe surveillance databases
- Evidence of prior SARS-CoV-2 infection (defined as positive PCR test, antigen test or rapid antibody test)

Although data on comorbidities is available through e-SUS and SIVEP-Gripe, we expect this data to be missing not at random. Adjusting for comorbidities will therefore introduce bias. We will explore the feasibility of multiple imputation of comorbidity in a sensitivity analysis.

To assess vaccine effectiveness against SARS-CoV-2 infection regardless of symptoms, we will enroll an independent case-control sample with eligibility criteria updated as follows: the eligibility criterion regarding appearance and onset of symptoms is removed for cases and controls. In addition, we will include an extra matching factor, being “Had symptom onset date that was between 0-10 days before sample collection date”. We will match individuals that were assumed to seek testing due to active symptoms vs. those that were assumed to seek testing for other reasons, to control for differences in healthcare access and utilization between vaccinated and unvaccinated individuals.

To assess vaccine effectiveness against hospitalization with COVID-19, we will enroll an independent case-control sample with eligibility criteria updated as follows: the eligibility criterion regarding appearance and onset of symptoms is replaced with “Was hospitalized”.

### Sample size

The power of a matched case-control study depends on the assumed odds ratio and the number of discordant pairs (i.e. pairs in which the case is exposed and the control is unexposed, or vice versa), which is a function of the assumed odds ratio and the expected prevalence of exposure among controls. As vaccine coverage and incidence are changing over time, the latter in ways we cannot predict, we will enroll individuals until we have reached a target number of discordant pairs (which we can assess without analyzing the data).

- For the primary analysis of the effectiveness against symptomatic COVID-19, 25 discordant pairs would give 80% power to detect vaccine effectiveness of 70%  $\geq 14$  days after the second dose.
- For the analysis of effectiveness against symptomatic COVID-19 following a single dose, 69 discordant pairs would give 80% power to detect vaccine effectiveness of 50%.

Therefore, we will complete enrollment and perform the primary analysis once 25 discordant pairs have been enrolled for the primary outcome, and once 69 discordant pairs have been enrolled for the outcome of effectiveness of a single dose.

### Interim analysis

Formal analysis of the data will be performed at a number of time points. The final analysis will only be performed and reported upon reaching the target sample size for each primary outcome (as above). However, to allow for early reporting of significant results for the benefit of public health, specifically the question whether receiving at least one dose of the vaccine is effective, we plan an interim analysis. For the primary endpoint (vaccine effectiveness against symptomatic COVID-19), we will perform interim analysis of the effectiveness following at least one dose of the vaccine, as we expect this analysis to be the first to reach desired power. This interim analysis will be triggered upon reaching 70% power. We will perform the initial analysis once 33 discordant pairs have been enrolled, giving 70% power to detect vaccine effectiveness of 60% of at least one dose  $\geq 14$  days after the first dose.

Once the interim analysis above has been triggered, we will perform at most one additional analysis when 75% power is achieved. The number of discordant pairs that triggers this analysis is 36. To correct for multiple testing we will use the O’Brien Fleming alpha-spending method, meaning that for three interim analyses the critical p-values at each analysis will be 0.0006, 0.0151, and 0.0471. If only two interim analyses are performed, the critical p-values at each analysis will be 0.0054 and 0.0492.

### **Privacy**

Only SEMSA, SES-AM and OPAS technicians had access to the identified dataset to linkage the datasets by name, date of birth, mother's name and CPF. After the linkage, the CPF was encrypted and the de-identified dataset was sent to the team for analysis.

Working group: Matt Hitchings, Otavio T. Ranzani, Julio Croda, Albert I. Ko, Derek Adam Cummings, Silvano Barbosa de Oliveira, Wildo Navegantes de Araujo, Jason R. Andrews, Ryan Borg, Roberto Dias de Oliveira, Daniel Barros, Andre Siqueira, Patricia Vieira da Silva, Bernardino Cláudio de

Albuquerque, Shadia Hussami Hauache Fraxe, Cristiano Fernandes da Costa, Mario Sergio  
Sacaramuzzini Torres, Felipe Gomes Naveca, Vanderson de Souza Sampaio



## References

- 1 Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* 2020; **20**: 533–4.
- 2 de Souza WM, Buss LF, Candido D da S, *et al.* Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav* 2020; **4**: 856–65.
- 3 Ranzani OT, Bastos LSL, Gelli JGM, *et al.* Characterisation of the first 250 000 hospital admissions for COVID-19 in Brazil: a retrospective analysis of nationwide data. *The Lancet Respiratory Medicine* 2021; : S2213260020305609.
- 4 Buss LF, Prete CA, Abraham CMM, *et al.* Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science* 2021; **371**: 288–92.
- 5 Sabino EC, Buss LF, Carvalho MPS, *et al.* Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *The Lancet* 2021; **397**: 452–5.
- 6 Fujino T, Nomoto H, Kutsuna S, *et al.* Novel SARS-CoV-2 Variant Identified in Travelers from Brazil to Japan. *Emerg Infect Dis* 2021; **27**. DOI:10.3201/eid2704.210138.
- 7 Naveca F, Nascimento V, Souza V, *et al.* COVID-19 epidemic in the Brazilian state of Amazonas was driven by long-term persistence of endemic SARS-CoV-2 lineages and the recent emergence of the new Variant of Concern P.1. In Review, 2021 DOI:10.21203/rs.3.rs-275494/v1.
- 8 Faria NR, *et al.* Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings. Virological. 2021; published online Jan 12. <https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586> (accessed Feb 27, 2021).
- 9 Faria NR, Mellan TA, Whittaker C, Claro IM, *et al.* Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. <https://github.com/CADDE-CENTRE/>. <https://github.com/CADDE-CENTRE/Novel-SARS-CoV-2-P1-Lineage-in-Brazil> (accessed Feb 27, 2021).
- 10 Jangra S, Ye C, Rathnasinghe R, *et al.* The E484K mutation in the SARS-CoV-2 spike protein reduces but does not abolish neutralizing activity of human convalescent and post-vaccination sera. *medRxiv* 2021; published online Jan 29. DOI:10.1101/2021.01.26.21250543.
- 11 Nonaka CKV, Franco MM, Gräf T, *et al.* Genomic Evidence of SARS-CoV-2 Reinfection Involving E484K Spike Mutation, Brazil. *Emerg Infect Dis* 2021; **27**. DOI:10.3201/eid2705.210191.
- 12 Wu Z, Hu Y, Xu M, *et al.* Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine (CoronaVac) in healthy adults aged 60 years and older: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. *The Lancet Infectious Diseases* 2021; : S1473309920309877.
- 13 Zhang Y, Zeng G, Pan H, *et al.* Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine in healthy adults aged 18–59 years: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial. *Lancet Infect Dis* 2021; **21**: 181–92.

- 14 Sinovac Announces Phase III Results of Its COVID-19 Vaccine-SINOVAC - Supply Vaccines to Eliminate Human Diseases. [http://www.sinovac.com/?optionid=754&auto\\_id=922](http://www.sinovac.com/?optionid=754&auto_id=922) (accessed Feb 27, 2021).
- 15 Dagan N, Barda N, Kepten E, *et al.* BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass Vaccination Setting. *N Engl J Med* 2021; : NEJMoa2101765.
- 16 Verani JR, Baqui AH, Broome CV, *et al.* Case-control vaccine effectiveness studies: Preparation, design, and enrollment of cases and controls. *Vaccine* 2017; **35**: 3295–302.
- 17 Verani JR, Baqui AH, Broome CV, *et al.* Case-control vaccine effectiveness studies: Data collection, analysis and reporting results. *Vaccine* 2017; **35**: 3303–8.
- 18 Camargo KR, Coeli CM. [Reclink: an application for database linkage implementing the probabilistic record linkage method]. *Cad Saude Publica* 2000; **16**: 439–47.

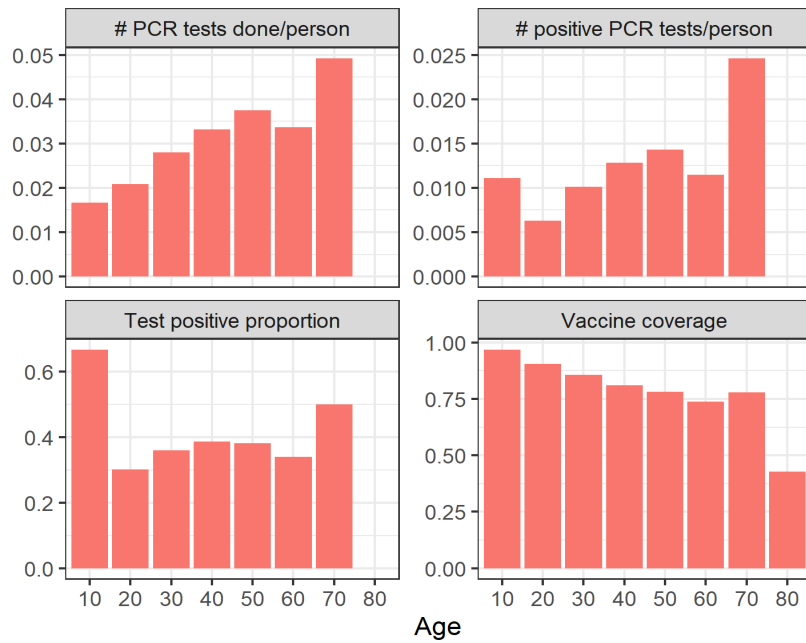


Figure 1: PCR testing rate, PCR positive testing rate, test-positive proportion, and vaccine coverage, by age

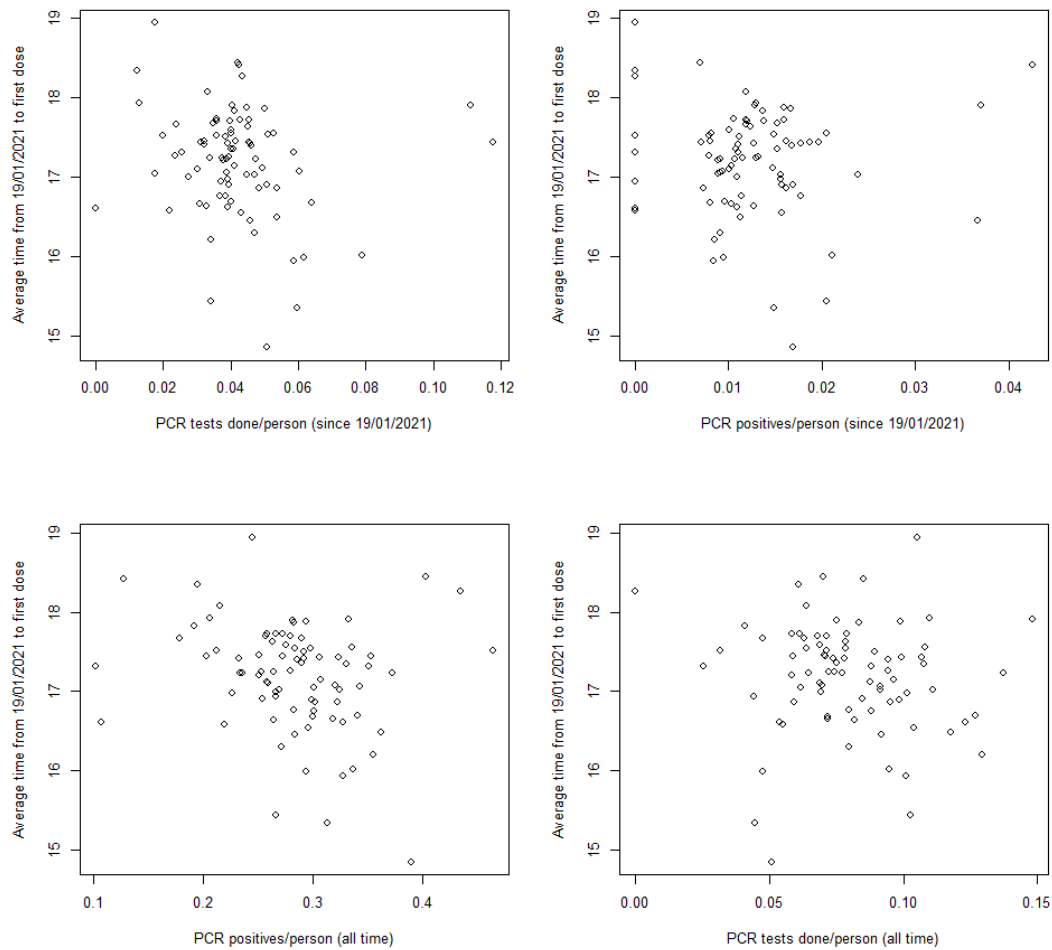


Figure 2: PCR testing rate, and PCR positive testing rate against time to first dose of vaccine, by neighborhood

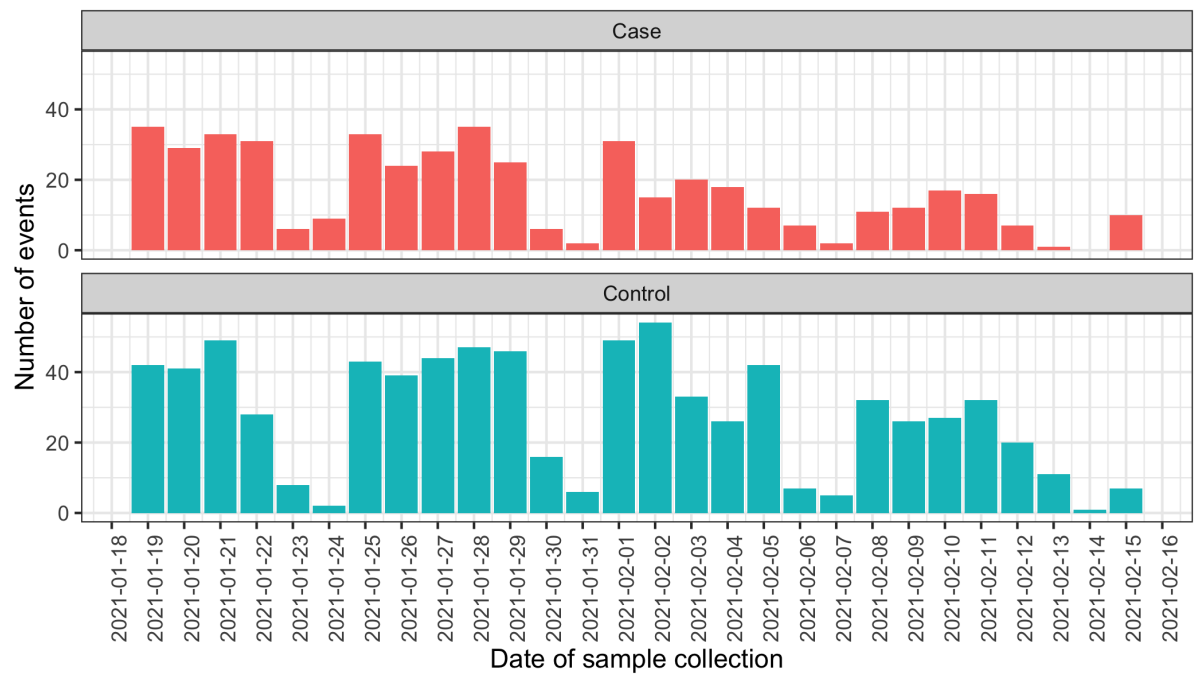


Figure 3: Number of positive and negative PCR tests over time in the study period, in the study population

## Supplementary material

Table 1. Variables included in study

Variable	Source	Label
id	SEMSA/SES-AM/CADSUS	citizen id encrypted
lot_id_1_dose	SMV	vaccine code - dose 1
lot_id_2_dose	SMV	vaccine code - dose 2
category_id	SMV	main priority group according to MoH
application_date_1_dose	SMV	date of vaccine application - dose 1
application_date_2_dose	SMV	date of vaccine application - dose 2
servicegroup_id	SMV	service group (subset) according to MoH
notifica_datainiciosintomas	E-SUS Notifica	date of symptom onset
notifica_datanotificacao	E-SUS Notifica	date of report
notifica_datanascimento	E-SUS Notifica	birthdate
notifica_sexo	E-SUS Notifica	sex
notifica_racacor	E-SUS Notifica	race/color
notifica_id_geo	E-SUS Notifica	zip code of residence
notifica_tipoteste	E-SUS Notifica	type of test
notifica_resultadoteste	E-SUS Notifica	result of test
notifica_classificacaofinal	E-SUS Notifica	final classification of case
notifica_evolucao	E-SUS Notifica	case evolution
notifica_profissionaisaude	E-SUS Notifica	health care worker
notifica_cbo	E-SUS Notifica	occupation Brazilian code
notifica_dispineia	E-SUS Notifica	symptom dyspnea
notifica_dorgarganta	E-SUS Notifica	symptom sore throat
notifica_febre	E-SUS Notifica	symptom fever
notifica_tosse	E-SUS Notifica	symptom cough
notifica_diabetes	E-SUS Notifica	comorbidity - diabetes
notifica_doencacardiacacronica	E-SUS Notifica	comorbidity–chronic cardiovascular disease
notifica_doencarenal	E-SUS Notifica	comorbidity - chronic kidney disease
notifica_doencarespcronica	E-SUS Notifica	comorbidity - pneumopathy
notifica_gestantealtorisco	E-SUS Notifica	comorbidity - high-risk pregnant woman
notifica_imunossupressao	E-SUS Notifica	comorbidity - immunosuppression

notifica_portadordoencacromo	E-SUS Notifica	comorbidity - chromosomal disease
semsa_cargo	SEMSA - local health care worker	profession
semsa_vinc	SEMSA - local health care worker	type of occupation contract
semsa_salario	SEMSA - local health care worker	wage (in BRL)
semsa_sex	SEMSA - local health care worker	sex
semsa_race	SEMSA - local health care worker	race/color
semsa_lotacao	SEMSA - local health care worker	place of work
semsa_dtnascto	SEMSA - local health care worker	birthdate
semsa_id_geo	SEMSA - local health care worker	zip code of residence
semsa_dtadmissao	SEMSA - local health care worker	admission date
sivep_nu_notific	SIVEP Gripe - SARI Cases	number of notification
sivep_uti	SIVEP Gripe - SARI Cases	hospitalized in ICU
sivep_cs_raca	SIVEP Gripe - SARI Cases	race/color
sivep_perd_olft	SIVEP Gripe - SARI Cases	loss of smell
sivep_dt_entuti	SIVEP Gripe - SARI Cases	date of ICU admission
sivep_pcr_para3	SIVEP Gripe - SARI Cases	result of PCR to Parainfluenza 3
sivep_dispneia	SIVEP Gripe - SARI Cases	symptom dyspnea
sivep_fadiga	SIVEP Gripe - SARI Cases	symptom fatigue
sivep_dt_pcr	SIVEP Gripe - SARI Cases	date of result of PCR test
sivep_an_para3	SIVEP Gripe - SARI Cases	result of antigen to Parainfluenza 3
sivep_pcr_fluasu	SIVEP Gripe - SARI Cases	result of PCR to Influenza
sivep_dt_interna	SIVEP Gripe - SARI Cases	date of hospitalization
sivep_cs_escol_n	SIVEP Gripe - SARI Cases	education degree
sivep_dt_coleta	SIVEP Gripe - SARI Cases	date of test collection
sivep_pcr_boca	SIVEP Gripe - SARI Cases	result of PCR to Bocavirus
sivep_dt_saiduti	SIVEP Gripe - SARI Cases	date of leaving ICU
sivep_pac_cocbo	SIVEP Gripe - SARI Cases	occupation Brazilian code
sivep_tp_idade	SIVEP Gripe - SARI Cases	age type
sivep_res_an	SIVEP Gripe - SARI Cases	antigen result
sivep_an_para2	SIVEP Gripe - SARI Cases	result of antigen to Parainfluenza 2
sivep_pcr_metap	SIVEP Gripe - SARI Cases	result of PCR to Metapneumovirus
sivep_evolucao	SIVEP Gripe - SARI Cases	case evolution
sivep_criterio	SIVEP Gripe - SARI Cases	confirmation criteria

sivep_cardiopati	SIVEP Gripe - SARI Cases	comorbidity – chronic cardiovascular disease
sivep_pcr_para1	SIVEP Gripe - SARI Cases	result of PCR to Parainfluenza 1
sivep_nu_cep	SIVEP Gripe - SARI Cases	zip code of residence
sivep_obesidade	SIVEP Gripe - SARI Cases	comorbidity - obesity
sivep_sind_down	SIVEP Gripe - SARI Cases	comorbidity - Down Syndrome
sivep_pcr_vsr	SIVEP Gripe - SARI Cases	result of PCR to respiratory syncytial virus
sivep_res_iga	SIVEP Gripe - SARI Cases	serological test IgA to COVID-19
sivep_res_igm	SIVEP Gripe - SARI Cases	serological test IgM to COVID-19
sivep_dor_abd	SIVEP Gripe - SARI Cases	symptom - abdominal pain
sivep_asma	SIVEP Gripe - SARI Cases	comorbidity - asthma
sivep_cs_gestant	SIVEP Gripe - SARI Cases	gestational age
sivep_pcr_para4	SIVEP Gripe - SARI Cases	result of PCR to Parainfluenza 4
sivep_an_para1	SIVEP Gripe - SARI Cases	result of antigen to Parainfluenza 1
sivep_pneumopati	SIVEP Gripe - SARI Cases	comorbidity – chronic respiratory disorder
sivep_diarreia	SIVEP Gripe - SARI Cases	symptom diarrhea
sivep_hematologi	SIVEP Gripe - SARI Cases	comorbidity - hematological disease
sivep_imunodepre	SIVEP Gripe - SARI Cases	comorbidity - immunodepression
sivep_pcr_para2	SIVEP Gripe - SARI Cases	result of PCR to Parainfluenza 2
sivep_pcr_sars2	SIVEP Gripe - SARI Cases	result of PCR to SARS-nCOV2
sivep_tosse	SIVEP Gripe - SARI Cases	symptom - cough
sivep_an_vsr	SIVEP Gripe - SARI Cases	result of antigen to respiratory syncytial virus
sivep_dt_evoluca	SIVEP Gripe - SARI Cases	evolution date
sivep_garganta	SIVEP Gripe - SARI Cases	symptom sore throat
sivep_suport_ven	SIVEP Gripe - SARI Cases	Use of ventilatory support type
sivep_neurologic	SIVEP Gripe - SARI Cases	comorbidity - neurological disease
sivep_pcr_rino	SIVEP Gripe - SARI Cases	result of PCR to Rinovirus
sivep_pac_dscbo	SIVEP Gripe - SARI Cases	occupation
sivep_dt_notific	SIVEP Gripe - SARI Cases	date of report
sivep_dt_nasc	SIVEP Gripe - SARI Cases	birthdate
sivep_hospital	SIVEP Gripe - SARI Cases	hospitalized
sivep_raiox_res	SIVEP Gripe - SARI Cases	result of chest-X-ray
sivep_perd_pala	SIVEP Gripe - SARI Cases	symptom - loss of taste
sivep_classi_fin	SIVEP Gripe - SARI Cases	final classification

sivep_dt_sin_pri	SIVEP Gripe - SARI Cases	date of symptom onset
sivep_fator_risc	SIVEP Gripe - SARI Cases	risk factor
sivep_cs_sexo	SIVEP Gripe - SARI Cases	sex
sivep_dt_encerra	SIVEP Gripe - SARI Cases	date of conclusion
sivep_an_sars2	SIVEP Gripe - SARI Cases	result of antigen to SARS-nCoV2
sivep_renal	SIVEP Gripe - SARI Cases	comorbidity - chronic kidney disease
sivep_an_adeno	SIVEP Gripe - SARI Cases	result of antigen to Adenovirus
sivep_febre	SIVEP Gripe - SARI Cases	symptom - fever
sivep_hepatica	SIVEP Gripe - SARI Cases	comorbidity - chronic liver disease
sivep_nu_idade_n	SIVEP Gripe - SARI Cases	age
sivep_dt_res	SIVEP Gripe - SARI Cases	date of result of serological test
sivep_vomito	SIVEP Gripe - SARI Cases	symptom - vomit
sivep_diabetes	SIVEP Gripe - SARI Cases	comorbidity - diabetes
sivep_pcr_resul	SIVEP Gripe - SARI Cases	date of result of PCR test
sivep_res_igg	SIVEP Gripe - SARI Cases	serological test IgG to COVID-19
sivep_co_mun_res	SIVEP Gripe - SARI Cases	IBGE code of residence municipality
sivep_pcr_adeno	SIVEP Gripe - SARI Cases	result of PCR to Adenovirus
sivep_puerpera	SIVEP Gripe - SARI Cases	puerperal
sivep_dt_res_an	SIVEP Gripe - SARI Cases	date of result of antigen test
sivep_dtnasc	SIVEP Gripe - SARI Cases	birthdate
sim_nu_do	SIM (Mortality System)	number of death certificate
sim_dt_obito	SIM (Mortality System)	date of death
sim_dt_nascimento	SIM (Mortality System)	birthdate
sim_nu_idade	SIM (Mortality System)	age
sim_sg_sexo	SIM (Mortality System)	sex
sim_tp_raca_cor	SIM (Mortality System)	race/color
sim_tp_escolaridade	SIM (Mortality System)	education degree
sim_co_ocupacao_falecido	SIM (Mortality System)	occupation
sim_nu_cep_residencia	SIM (Mortality System)	zip code of residence
sim_co_cid_causa_morte	SIM (Mortality System)	ICD10 cause of death
sim_ds_causa_morte_a	SIM (Mortality System)	ICD10informed on line A
sim_ds_causa_morte_b	SIM (Mortality System)	ICD10informed on line B
sim_ds_causa_morte_c	SIM (Mortality System)	ICD10informed on line C



sim_ds_causa_morte_d	SIM (Mortality System)	ICD10informed on line D
sim_ds_causa_morte_2	SIM (Mortality System)	ICD10informed on line 2
sim_co_cid_causa_basica_bath	SIM (Mortality System)	ICD10 underlaying/basic cause of death
cadsus_CEP	CADSUS	Zipcode
cadsus_Bairro	CADSUS	neighborhood
cadsus_PaisNacionalidade	CADSUS	nationality
cadsus_NomeMunicipio	CADSUS	Municipality of residence
cadsus_RacaCor	CADSUS	race/color
cadsus_Sexo	CADSUS	sex
cadsus_UF	CADSUS	state of residence
cadsus_DataNascimento	CADSUS	birthdate
ses_am_DTNASCTO	SES-AM - state health care worker	birthdate
ses_am_GANHO	SES-AM - state health care worker	wage (in BRL)
ses_am_LOTACAO_1	SES-AM - state health care worker	place work 1
ses_am_LOTACAO_2	SES-AM - state health care worker	place work 2
ses_am_LOTACAO_3	SES-AM - state health care worker	place work 3
ses_am_LOCAL_1	SES-AM - state health care worker	unit health 1
ses_am_LOCAL_2	SES-AM - state health care worker	unit health 2
ses_am_LOCAL_3	SES-AM - state health care worker	unit health 3
ses_am_CARGO_1	SES-AM - state health care worker	profession 1
ses_am_CARGO_2	SES-AM - state health care worker	profession 2
ses_am_CARGO_3	SES-AM - state health care worker	profession 3
gal_requisicao	GAL	number of report
gal_data_requisicao	GAL	date of report
gal_paciente_sexo	GAL	sex
gal_paciente_data_nascimento	GAL	birthdate
gal_paciente_uf	GAL	state of residence
gal_municipio	GAL	Municipality of residence
gal_data_inicio_sintomas	GAL	date of symptom onset
gal_data_coleta	GAL	date of collection
gal_tempo_transporte_amostra	GAL	time to transportation the biological sample
gal_exame	GAL	type of exam
gal_metodo	GAL	method of exam

gal_resultado	GAL	result
gal_virus	GAL	virus detected
gal_subtipagem	GAL	subtype of Influenza
gal_linhagem	GAL	lineage of Influenza
gal_data_liberacao	GAL	date of exam result
gal_status	GAL	status of exam
gal_classificacao	GAL	final classification