Evaluation of Vaccine Effectiveness in Brazil against COVID-19 (VEBRA-COVID): A Test-Negative Case-Control Study Protocol

Version: 02 / 5 April 2021

I. 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.[1–3] Of grave concern is the resurgence of COVID-19 cases and hospitalizations observed in the city of Manaus, despite high estimated seroprevalence.[4,5] 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 this city in February 2021.[6–8] This lineage has accrued mutations associated with decreased neutralization,[9,10] and has since spread throughout Brazil.

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 was Sinovac's CoronaVac vaccine.[11,12] 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.[13] CoronaVac was approved for emergency use on 17 January in Brazil, and used to vaccinate healthcare workers and the general population, beginning with the oldest age groups, on 19 January 2021. AstraZeneca-Oxford's ChAdOx1 vaccine was approved on the same day and was administered beginning on 23 January 2021.

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. Several studies have demonstrated the effectiveness of COVID-19 vaccines against infection caused by the B.1.1.7 variant.[14] However, large-scale real-world investigations on vaccine effectiveness have not been conducted in regions where the P.1 variant is prevalent.

We propose a test-negative case-control study[15,16] of healthcare workers (HCWs) from the city of Manaus and the general population from the State of São Paulo to evaluate the effectiveness of vaccines in preventing COVID-19 in a setting of widespread P.1 VOC transmission.[7] Manaus was selected as the site for the HCW study since it was the first city to aggressively vaccinate HCWs in response to the P.1 epidemic. São Paulo was selected as the site to evaluate vaccine effectiveness in the general population because it is the largest state in Brazil (with a population of 46 million) and is currently experiencing a large outbreak. The HCW study will be limited to evaluating the effectiveness of CoronaVac since 97% of vaccinated HCWs in Manaus received this vaccine. The investigation in São Paulo will initially focus on the elderly since individuals >72 years of age were the first members of the general population to be eligible for vaccination. We will expand the study population as additional age groups become eligible for vaccination. Although the majority of vaccines administered in São Paulo is at present CoronaVac, there may be sufficient numbers to evaluate effectiveness of ChAdOx1. Furthermore, we expect that additional vaccines will be

approved and will evaluate their effectiveness. We will therefore continue to amend the protocol and its objectives accordingly to address these new questions.

II. Objectives

1. To estimate the effectiveness of CoronaVac against symptomatic SARS-CoV-2 infection amongst healthcare workers from the city of Manaus.
2. To estimate the effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection among the population of São Paulo State.

III. Methods

1. Study Design: 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 HCWs (Section 3) and the general population (Section 4) as of the day that the COVID-19 vaccination campaign was initiated at the study sites. The study will evaluate vaccine effectiveness on the primary outcome of symptomatic SARS-CoV-2 infection and secondary outcome of SARS-CoV-2 RT-PCR test positivity regardless of symptoms. We will identify cases and matched controls by extracting information from health surveillance records and ascertain the type and data 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. We will perform interim analyses aimed at evaluating the effectiveness of receiving at least one vaccine dose and a final analysis that will evaluate the effectiveness of completing the approved vaccine series. In a separate analysis, we will assess the association between vaccination and hospitalization and/or death among individuals who have tested positive for SARS-CoV-2.

2. 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 Barcelona Institute for Global Health was conducted to inform the public health response and was therefore covered under Public Health Response Authorization under the US Common Rule.

3. Objective 1 Study: Effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection amongst healthcare workers from Manaus

Study Site: 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 (cumulative incidence: 6,522 per 100,000 population) and 7,605 deaths (cumulative mortality: 342 per 100,000 population). Manaus has 40 Family Health (*Plano Saúde da Familia*) 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. The Municipal Secretary of Health of Amazonas initiated its COVID-19 vaccination campaign on 19 January 2021 and is administering two vaccines, CoronaVac and ChAdOx1. CoronaVac has been used in >97% of the vaccinations of HCWs.

Data Sources and Integration: The overall approach will be to: 1) Identify the cohort of all HCW from Manaus from *state HCW registries*; 2) Identify eligible cases and controls from the cohort who test positive and negative, respectively, from the *state laboratory testing registry* of public health laboratory network; 3) Determine vaccination status from *municipal vaccination registries*; and 4) Extract information from *national healthcare and surveillance databases* that will be used to define outcomes, match controls to cases, determine vaccination status, serve as covariates for post-

stratification and provide a source for cross-validation of information from databases. Data sources will include:

- SES-AM HCW registry
- National health plan registry of users (**CADSUS**)
- National laboratory testing registry (**GAL**) of the network of public health laboratories
- Municipal COVID-19 vaccination registry
- National surveillance database of severe acute respiratory illnesses (**SIVEP-Gripe**) created by Ministry of Health Brazil in 2009
- National surveillance system of suspected cases of COVID-19 (**e-SUS**) from mild to moderate "influenza like illness", created by the Ministry of Health Brazil in 2020
    - e-SUS includes information from healthcare telemonitoring, whereby teams of healthcare practioners make daily telephone calls, assess symptoms, identify signs of severity and triage patients to healthcare facilities.
- National mortality registry (**SIM**) from the Ministry of Health

We will build 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.

We will use CPF numbers (Brazilian citizens' unique identifier code) to integrate datasets. For those entries missing CPF in SEMSA, we will perform 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 will use the Reclink III software,[17] and consider 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 will compare 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 will be manually reviewed and revised.

Some variables were reported in multiple data sources. To define a single variable, we drew from each database with priority given to databases that were more complete, reliable, and up-to-date. We will choose age from the data sources in the following order: CADSUS, SEMSA, GAL, e-SUS, SIVEP-Gripe. We will define neighborhoods (*bairro*) by extracting information on CEP (Brazilian zipcode) and transforming them to neighborhoods or directly extracting information on neighborhoods from data sources in the following order: CADSUS (CEP), CADSUS (*bairro*), SEMSA (CEP), and e-SUS (CEP).

Study Population

Inclusion criteria:

- Healthcare worker as defined by the SEMSA registry,
- Has a residential address within the city of Manaus,
- Age ≥ 18 years before 19 January 2021,
- 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,
- Does not have a residential address within the city of Manaus,
- Aged < 18 years before 19 January 2021,

- 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: We will use information from integrated GAL/SIVEP-Gripe/e-SUS databases to identify eligible cases. 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
- Did not have a positive RT-PCR test in the preceding 90 day period,
- Have complete and consistent data on SARS-CoV-2 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 integrated GAL/SIVEP-Gripe/e-SUS databases to identify eligible controls. 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,
- Did not have a subsequent positive PCR test in the following 7-day period
- Have complete and consistent data on SARS-CoV-2 PCR test result,
- Did not receive a dose of ChAdOx1 vaccine before the date of respiratory sample collection.

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):

- Symptomatic illness status at time of testing, defined as the presence or absence of one or more reported COVID-19 related symptom with onset within 0-10 days before the date of their positive and negative PT-PCR test, respectively, for case and controls,
- Residential address (neighborhood [*bairro*], which is identified based on the first 5 digits of 8 digit CEP),
- Age (categorized as <30, ≥30 and <60, and ≥60 years; Figure 3 shows similar rates of testing positive for individuals aged 30-60),
- Window of ±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.

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.

Exposure definition: CoronaVac vaccination in 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

Statistical Analyses: We will evaluate the effectiveness of CoronaVac for the following SARS-CoV-2 infection outcomes:

- Primary: Symptomatic COVID-19, defined as one or more reported COVID-19 related symptom with onset within 0-10 days before the date of their positive RT-PCR test
- Secondary:
    - SARS-CoV-2 RT-PCR test positivity
    - COVID-19 associated hospitalization within 14 days of the symptom onset
    - COVID-19 associated death within 28 days of symptom onset

We will evaluate vaccine effectiveness for the primary outcome and the secondary outcome of test positivity in case control analyses according to the test-negative design. Table 2 shows a list of all planned analyses in the test-negative design. The test-negative design may introduce bias when evaluating outcomes of hospitalizations and deaths during an epidemic. We will therefore perform survival analyses of HCWs who test positive to evaluate the association of vaccination status and the risk for hospitalization and death after infection.

Case-control analysis: Analyses of the primary outcome will be restricted to case and control pairs who are matched based on the presence of a COVID-19 related symptom before or at the time of testing. Analyses of the secondary outcomes of test positivity will include additional case and control pairs who are matched based on the absence of a COVID-19 related symptom before or at the time of testing.

We will use conditional logistic regression to estimate the odds ratio (OR) of vaccination among cases and controls, accounting for the matched design, where 1-OR provides an estimate of vaccine effectiveness under the standard assumptions of a test-negative design. The reference group will be individuals who have not received a first dose of CoronaVac by the date of respiratory sample collection. Date of notification of the testing result will be used if the date of respiratory sample collection is missing. To evaluate potential biases and the timing of vaccine effectiveness after administration, we will evaluate the windows of vaccination status corresponding to 0-13 days and ≥14 days after the first dose and 0-13 days and ≥14 days after the 2nd dose.

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, this data may have different degrees of missingness between databases and between cases and control groups. Adjusting for comorbidities using complete case data will likely introduce bias. We will explore the feasibility of multiple imputation of comorbidity in a sensitivity analysis. Additional sensitivity

analyses will evaluate potential effect modification of the vaccine effectiveness by history of a positive RT-PCR, antigen or serological test result prior to the vaccination campaign.

***Survival analysis of h*ospitalization and death**: We will perform proportional survival analyses for hospitalization and death amongst HCWs who test positive and estimate the hazards according to vaccination status at the date of positive test, adjusting for covariates described in the case-control analyses. Sensitivity analyses will be conducted to evaluate the association of influence of a positive RT-PCR, antigen or serological test result prior to the vaccination campaign.

Sample size calculations and timing of analyses: 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. Moreover, the estimate of the odds ratio for one level of a categorical variable compared to baseline is determined by the distribution of all discordant pairs. As vaccine coverage and incidence are changing over time, the latter in ways we cannot predict, and there is no power formula for this analysis, we will simulate power and enroll individuals until we have reached a target power, which we can assess without analyzing the data.

Timing of interim and final analyses: Interim analyses will be performed 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. For the primary outcome (symptomatic SARS-CoV-2 infection), we will perform an 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 simulated power of 70% to detect vaccine effectiveness of 60% of at least one dose ≥14 days after the first dose. Once the interim analysis above has been triggered, we will perform one additional analysis when 80% power is achieved. To correct for multiple testing we will use the O'Brian Fleming alpha-spending method, meaning that for two interim analyses, the critical p-values at each analysis will be 0.0054 and 0.0492. We will perform final analyses of the primary outcome upon reaching simulated 80% power to detect vaccine effectiveness of 70% ≥14 days after the second dose.

4. Objective 2 Study: Effectiveness of COVID-19 vaccines against symptomatic SARS-CoV-2 infection among the population of São Paulo State

Study Site: The State of São Paulo (23°3'S, 46°4'W) is the most populous state in Brazil: an estimated population of 46,289,333 in 2020. São Paulo State reported 2,352,438 COVID-19 cases (cumulative incidence rate: 5,082 per 100,000 population) and 68,904 deaths (cumulative mortality: 149 per 100,000 population), by 24/03/2021. The State Secretary of Health of Sao Paulo (SES-SP) initiated its COVID-19 vaccination campaign on 17 January 2021 and is administering two vaccines, CoronaVac and ChAdOx1. As of 24 March 2021, 4.0 million and 927 thousand doses of CoronaVac and ChAdOx1 have been administered to HCWs and the general population, among which roll out has prioritized the oldest age groups.

Data Sources and Integration: We will follow the protocol described for Objective 1 with the following modifications. The approach will 1) Identify eligible cases and controls from the State of São Paulo who test positive and negative, respectively, from the *state laboratory testing registry* of public health laboratory network; 2) Determine vaccination status from *state vaccination registries*; and 3) Extract information from *national healthcare and surveillance databases* that will be used to define outcomes, match controls to cases, determine vaccination status, serve as covariates for post-stratification and provide a source for cross-validation of information from databases.

Registries are not available which enables constructing a cohort of people eligible for vaccination in the general population. Data sources for this study will include:

- State laboratory testing registry (**GAL**) of the network of public health laboratories
- State COVID-19 vaccination registry (**Vacina Já**)
- National surveillance database of severe acute respiratory illnesses (**SIVEP-Gripe**) created by Ministry of Health Brazil in 2009
- National surveillance system of suspected cases of COVID-19 (**e-SUS**) from mild to moderate "influenza like illness", created by the Ministry of Health Brazil in 2020

We will integrate data sources and build a parent database as described for Objective 1 with the following modifications. CPF will be extracted from GAL and used to integrate data sets. When discordant entries are observed between data sources, we will draw from each data base with the following priority order: GAL, e-SUS, SIVEP-Gripe. We will define a suitable geographical unit from e-SUS and/or SIVEP-Gripe.

Study Population

Inclusion criteria:

- Has a residential address in the State of São Paulo,
- Eligible to receive a COVID-19 vaccine based on age,
- With complete information, which is consistent between databases, on age, sex, race, and residential address
- With consistent vaccination status and dates for those who were vaccinated.

Exclusion criteria:

- Does not have a residential address in the State of São Paulo,
- Not eligible to receive a COVID-19 vaccine based on age,
- With missing or inconsistent information on age, sex, or residential address defined by CEP (zip code)
- With existing but inconsistent vaccination status or dates.

Case definition and eligibility: We will perform independent analyses of the effectiveness of Coronavac, ChAdOx1 and additional vaccines as they are approved and implemented. In analyses of a specific vaccine, we will exclude recipients of other vaccines in the selection of cases and their matched test negative controls. We will use information from integrated GAL/SIVEP-Gripe/e-SUS databases to identify cases that 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 17, 2021 and 7 days prior to database extraction of information
- Did not have a positive RT-PCR test in the preceding 90 day period,
- Have complete and consistent data on SARS-CoV-2 PCR test result

Control definition and eligibility: We will use integrated GAL/SIVEP-Gripe/e-SUS databases to identify eligible controls. In analyses of a specific vaccine, we will exclude recipients of other vaccines in the selection of controls, as done with cases. 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 17, 2021,
- Did not have a subsequent positive PCR test in the following 7-day period,
- Have complete and consistent data on SARS-CoV-2 PCR test result

Matching: As in Objective 1, test-negative controls will be matched 1:1 to the cases. Matching factors will include:

- Symptomatic illness status at time of testing, defined as the presence or absence of one or more reported COVID-19 related symptom with onset within 0-10 days before the date of their positive and negative PT-PCR test, respectively, for case and controls,
- Residential address, defined at a suitable geographical unit,
- Age according to 3-year brackets,
- Gender
- Window of ±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.

Standard algorithms will be used to conduct matching as described in Objective 1.

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

Statistical Analyses: As with Objective 1, we will evaluate the effectiveness of Coronavac for the following SARS-CoV-2 outcomes:

- Primary: Symptomatic COVID-19, defined as one or more reported COVID-19 related symptom with onset within 0-10 days before the date of their positive RT-PCR test
- Secondary:
    - SARS-CoV-2 RT-PCR test positivity
    - COVID-19 associated hospitalization within 14 days of symptom onset
    - COVID-19 associated death within 28 days of symptom onset

We will evaluate vaccine effectiveness for the primary outcome and the secondary outcome of test positivity in case control analyses according to the test-negative design. Table 2 shows a list of all planned analyses in the test-negative design. The test-negative design may introduce bias when evaluating outcomes of hospitalizations and deaths during an epidemic. We will therefore perform survival analyses of HCWs who test positive to evaluate the association of vaccination status and the risk for hospitalization and death after infection.

Case-control analysis: Analyses of the primary outcome will be restricted to case and control pairs who are matched based on the presence of a COVID-19 related symptom before or at the time of testing. Analyses of the secondary outcomes of test positivity will include additional case and control pairs who are matched based on the absence of a COVID-19 related symptom before or at the time of testing.

We will use conditional logistic regression to estimate the odds ratio (OR) of vaccination among cases and controls, accounting for the matched design. 1-OR provides an estimate of vaccine effectiveness under the standard assumptions of a test-negative design. The reference group will be individuals who have not received a first dose of CoronaVac by the date of respiratory sample collection. Date of notification of the testing result will be used if the date of respiratory sample

collection is missing. To evaluate potential biases and the timing of vaccine effectiveness after administration, we will evaluate the windows of vaccination status corresponding to 0-13 days and ≥14 days after the first dose and 0-13 days and ≥14 days after the 2nd dose.

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
- 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 have different mechanism of missingness between databases. Adjusting for comorbidities using complete case data will likely introduce bias. We will assess the level of missingness in the data, and if needed we will explore the feasibility of multiple imputation of comorbidity in a sensitivity analysis. Additional sensitivity analyses will evaluate possible effect modification of vaccine effectiveness by history of a positive RT-PCR or serological test result prior to the vaccination campaign.

Survival analysis of hospitalization and death: We will perform proportional survival analyses for hospitalization, ICU admission, invasive respiratory support and death amongst the general population who test positive by RT-PCR and estimate the hazards according to vaccination status at the date of positive test, adjusting for covariates described in the case-control analyses. Sensitivity analyses will be conducted to evaluate the association of influence of a positive RT-PCR, antigen or serological test result prior to the vaccination campaign.

Sample size calculations and timing of analyses: As in Objective 1, we will simulate power and enroll individuals until we have reached a target power (which we can assess without analyzing the data).

Interim analyses will be performed 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. For the primary outcome (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 simulated power of 70% to detect vaccine effectiveness of 50% of at least one dose ≥14 days after the first dose. Once the interim analysis above has been triggered, we will perform one additional analysis when 80% power is achieved. To correct for multiple testing we will use the O'Brian Fleming alpha-spending method, meaning that for two interim analyses, the critical p-values at each analysis will be 0.0054 and 0.0492. We will perform final analyses of the primary outcome upon reaching simulated 80% power to detect vaccine effectiveness of 70% ≥14 days after the second dose.

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, Wade Schulz

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, Abrahim 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 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.

7 Faria NR, Mellan TA, Whittaker C, *et al.* Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv* 2021; published online March 3. DOI:10.1101/2021.02.26.21252554.

8 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.

9 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.

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 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.

12 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.

13 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 (accessed Feb 27, 2021).

14 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.

15 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.

16 Verani JR, Baqui AH, Broome CV, *et al.* Case-control vaccine effectiveness studies: Data collection, analysis and reporting results. *Vaccine* 2017; **35**: 3303–8.

17 Camargo KR, Coeli CM. [Reclink: an application for database linkage implementing the probabilistic record linkage method]. *Cad Saude Publica* 2000; **16**: 439–47.

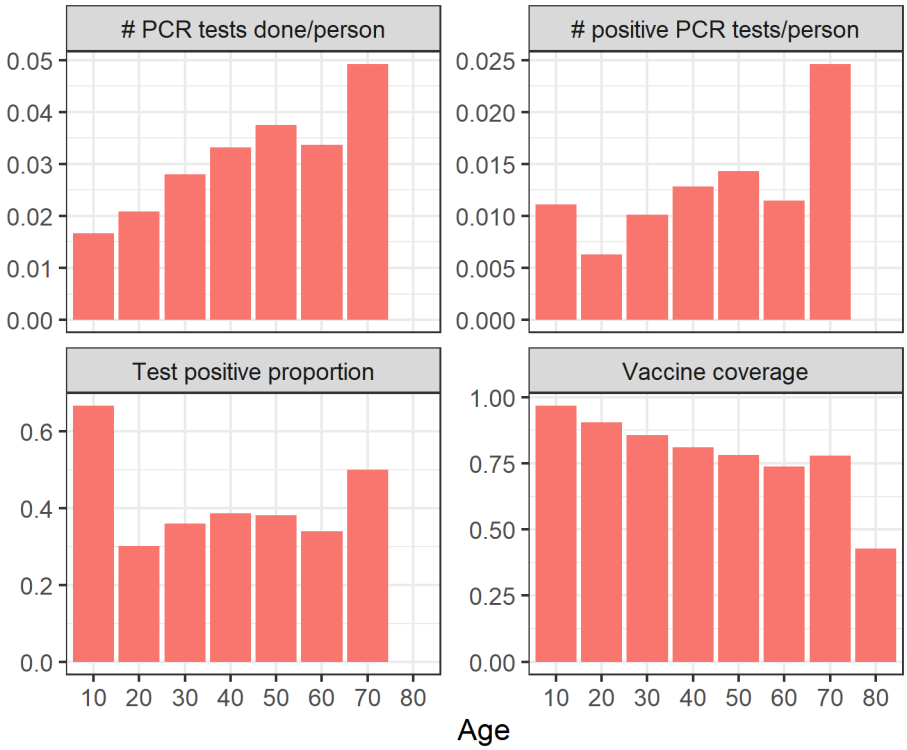

Figure 1: PCR testing rate, PCR positive testing rate, test-positive proportion, and vaccine coverage, by age (from data extracted on February 26, 2021)

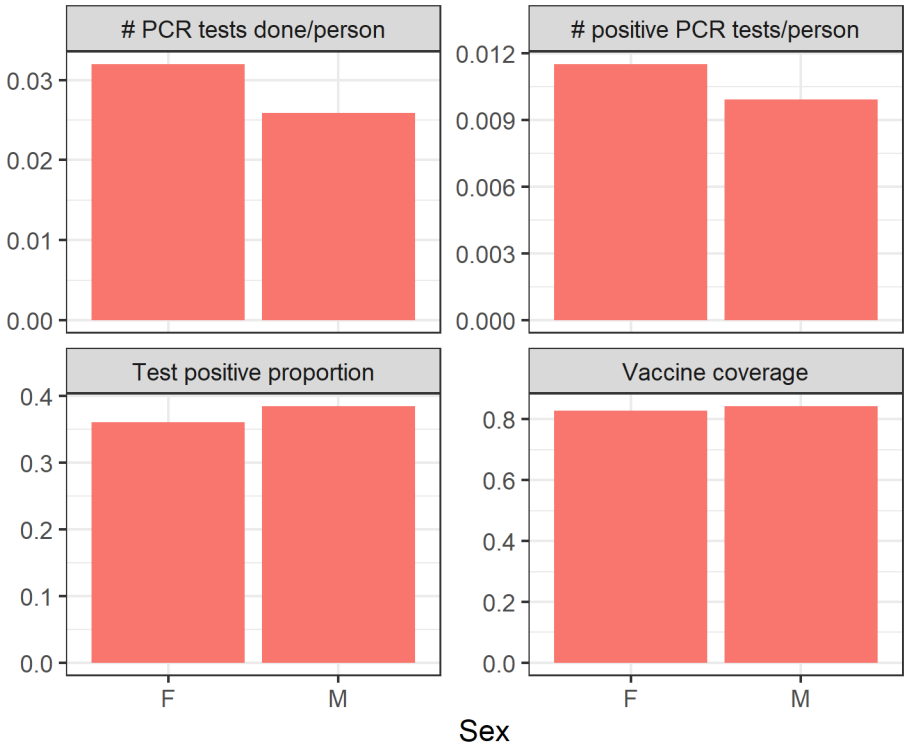

Figure 2: PCR testing rate, PCR positive testing rate, test-positive proportion, and vaccine coverage, by sex (from data extracted on February 26, 2021)

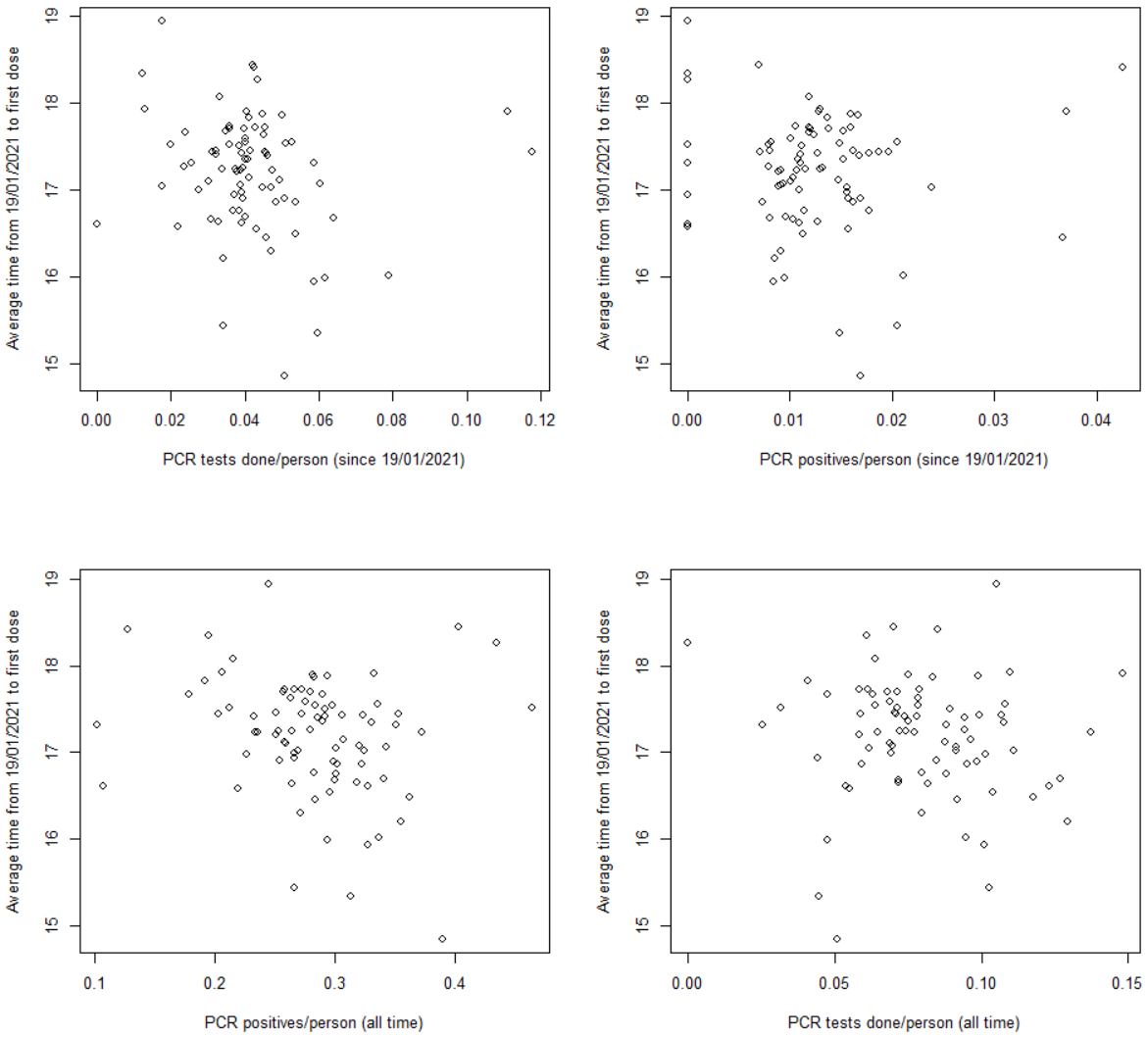

Figure 3: PCR testing rate, and PCR positive testing rate against average time from start of campaign to first dose administration, by *bairro*, (from data extracted on March 16, 2021)

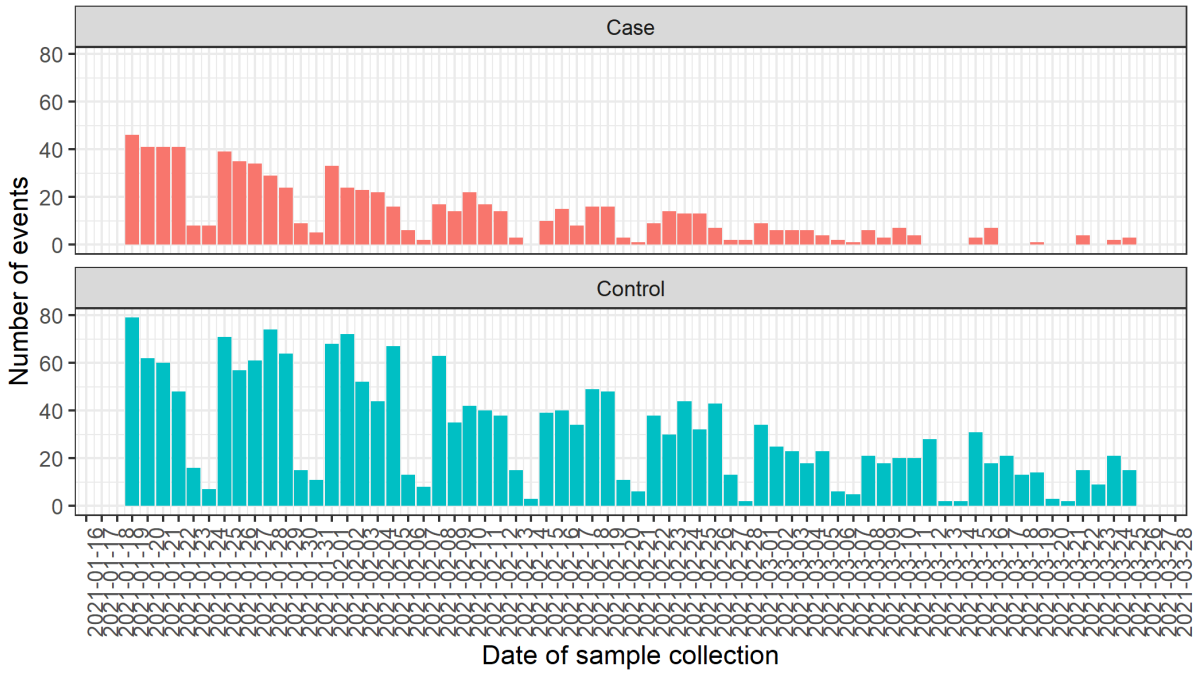

Figure 4: Number of positive and negative PCR tests over time in the study period, in the study population (from data extracted on April 4, 2021)

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_evolucaocaso	E-SUS Notifica	case evolution
notifica_profissionalsaude	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 - immunosupression

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

Table 2: Table of planned analyses

Analysis	Exposure	Outcome
Primary outcome, primary exposure	Two-dose regimen of CoronaVac in the period starting 14 days after administration of the 2nd dose	Positive test for SARS-CoV-2, with at least one COVID-19 symptom reported 0-10 days before sample collection date
Primary outcome, exposure for interim analysis	Receiving at least one dose of CoronaVac, in the period starting 14 days after administration of the 1st dose	
Primary outcome, secondary exposure (2-dose)	Two-dose regimen of CoronaVac in the period 0-13 days after administration of the 2nd dose	
Primary outcome, secondary exposure (1-dose)	One-dose regimen of CoronaVac, in the period starting 14 days after administration of the 1st dose	
Primary outcome, bias indicator	One-dose regimen of CoronaVac, in the period 0-13 days after administration of the 1st dose	
Secondary outcome, primary exposure	Two-dose regimen of CoronaVac in the period starting 14 days after administration of the 2nd dose	Positive test for SARS-CoV-2
Secondary outcome, exposure for interim analysis	Receiving at least one dose of CoronaVac, in the period starting 14 days after administration of the 1st dose	
Secondary outcome, secondary exposure (2-dose)	Two-dose regimen of CoronaVac in the period 0-13 days after administration of the 2nd dose	
Secondary outcome, secondary exposure (1-dose)	One-dose regimen of CoronaVac, in the period starting 14 days after administration of the 1st dose	
Secondary outcome, bias indicator	One-dose regimen of CoronaVac, in the period 0-13 days after administration of the 1st dose	