Statistical analysis plan (SAP) of the study:

The waning of CoronaVac effectiveness and the effect of a BNT162b2 booster: A nationwide test-negative design case-control study in Brazil

Version History			
Version	**Date**	**Author**	**Notes**
V3	06/10/2021	RFO/TCS	Circulated to AS and SVK for comments
V4	02/12/2021	RFO/TCS/VSB/ESP	Circulated to SVK, NP, and CR for comments

Contents

1. Introduction

1.1 Background

CoronaVac, an inactivated-virus vaccine for COVID-19, has been used in all Brazilian states from January to the present (November 2021). Evaluating the effectiveness of vaccines used in COVID-19 immunisation programmes is necessary to guide the national public health response (1). A previous study assessed the effectiveness of CoronaVac among adults in Brazil (2). However, the waning of protection after completion of CoronaVac's two-dose has not previously been investigated.

2. Objectives

2.1 General objective

Using Brazilian national data, we aim to evaluate the vaccine effectiveness (VE) of the two-dose schedule of CoronaVac over time, and the effectiveness of the BNT162b2 vaccine as a booster dose, against confirmed symptomatic SARS-CoV-2 infection and severe COVID-19 (defined as COVID-19 related hospitalisation or death).

2.2 Specific objectives

a. To estimate the VE of CoronaVac over time against confirmed symptomatic SARS-CoV-2 infection, following receipt of the second CoronaVac dose, overall and by age group.
b. To estimate the VE of CoronaVac over time against severe COVID-19 (defined as COVID-19 related hospitalisation or death), following receipt of the second CoronaVac dose, overall and by age group.
c. To estimate the VE of the vaccination scheme: two doses of CoronaVac + BNT162b2 booster, against confirmed symptomatic SARS-CoV-2 infection, among adults who completed the two-dose schedule of CoronaVac, overall and by age group.
d. To estimate the VE of the vaccination scheme: two doses of CoronaVac + BNT162b2 booster, against severe COVID-19 (defined as COVID-19 related hospitalisation or death), among adults who completed the two-dose schedule of CoronaVac, overall and by age group.

3. Study design

3.1 Study design

The test negative design is a type of case-control study, which uses population test results, with the positive tests being the cases and the negative test being the controls (3). It is ideally suited to situations where not everyone in a population is being tested, since the factors that influence being tested (health seeking behaviour, access to health care, availability of testing, etc) will apply to both those who test positive and those who test negative (3,4). A test-negative case-control design

(TND) will be used to estimate the VE of the two-dose schedule of CoronaVac over time, and the VE of the BNT162b2 vaccine as a booster dose, against confirmed symptomatic SARS-CoV-2 infection and severe COVID-19 (defined as COVID-19 related hospitalisation or death), among individuals in Brazil who had COVID-19-like symptoms and were tested with RT-PCR for SARS-CoV-2 between January 18, 2021 and November 11, 2021.

3.2 Setting

Brazil.

3.3 Population

Individuals aged 18 years or older in Brazil who have COVID-19-like symptoms and were tested with RT-PCR for SARS-CoV-2 between January 18, 2021 and November 11, 2021.

3.4 Inclusion/exclusion criteria

The eligible study population will be individuals aged 18 years or older in Brazil who have COVID-19-like symptoms and were tested with RT-PCR for SARS-CoV-2.

We will exclude (i) individuals younger than 18 years, (ii) individuals who received first and second doses of different vaccines; (iii) individuals whose time interval between the first and second doses was less than 14 days; (iv) tests with missing information of age, sex, city of residence or sample collection date; (v) consecutive negative tests less than 14 days apart; (vi) negative tests followed by a positive test up to 7 days; (vii) any test after a positive test up to 90 days; (viii) tests with a symptom onset date greater than notification date.

3.5 Cases and controls for analysing confirmed symptomatic SARS-CoV-2 infection

Cases: individuals who have COVID-19-like symptoms and have a positive RT-PCR test result for SARS-CoV-2 from a sample collected within 10 days of symptoms onset.

Controls: individuals who have COVID-19-like symptoms and have a negative RT-PCR test result for SARS-CoV-2 from a sample collected within 10 days of symptoms onset.

3.6 Cases and controls for analysing severe COVID-19

Cases: individuals who have COVID-19-like symptoms, have a positive RT-PCR test result for SARS-CoV-2 within 10 days of symptoms onset, and have experienced COVID-19 related hospitalisation within 28 days of sample collection, or death up to 28 days of sample collection.

Controls: individuals who have COVID-19-like symptoms and have a negative RT-PCR test result for SARS-CoV-2 from a sample collected within 10 days of symptoms onset.

3.7 Valid RT-PCR test result

An RT-PCR test result will be considered valid if:

1. There is no positive test result in the previous 90 days to the current positive test result.
2. There is no positive test result in the 14 days after a negative test result.

3.8 Variables

3.8.1 Outcomes

The outcomes of interest will be:

(i) Confirmed symptomatic SARS-CoV-2 infection;
(ii) Severe COVID-19, defined as COVID-19 related hospitalisation or death.

3.8.2 Exposures

For estimating the VE of CoronaVac over time, the exposure will be time since the second dose of CoronaVac, with unvaccinated individuals used as the reference group. Time since the second dose of CoronaVac will be defined as the time between vaccination date and symptoms onset.

In the current data, we may have limited follow-up time after the BNT162b2 booster. Therefore, for estimating the VE of the BNT162b2 booster, the exposure will be BNT162b2 booster receipt status, with time since vaccination not considered and unvaccinated individuals used as the reference group.

3.8.3 Potential confounders

The analyses will be adjusted for potential confounders such as age, sex, temporal trend, state of residence, previous infection, pregnancy, postpartum period, and comorbidities. These have been identified as potential confounders on the basis of findings from previous studies (2,5). The temporal trend will be estimated using the time elapsed, in days, between the study start and the date of symptoms onset.

3.8.4 Potential effect modifiers

Age, a potential effect modifier (6), will be considered for performing a subgroup analysis. Other effect modifiers, such as region or state, may be also considered if adequate numbers are available for deriving meaningful statistical inferences.

3.9 Data

3.9.1 Data sources

We will use three administrative datasets with national coverage:

- Programa Nacional de Imunizações (PNI), which holds records of all vaccines administered;
- e-SUS Notifica, which holds records of clinical and laboratory testing of all COVID-19 suspected cases; and
- Sistema de Informação da Vigilância Epidemiológica da Gripe (SIVEP-Gripe), which holds records of all COVID-19 related hospitalisations and deaths.

The datasets described will be deterministically linked using a unique identifier available across all health datasets, in a secure data-management environment within FIOCRUZ. The study researchers will have access to anonymised data only.

Table 1 lists the groupings of variables available for this study by data source. Exposure data are described in the vaccination category. Outcome data are described in the secondary care, mortality, and laboratory testing categories. The remaining categories contain data on potential confounding factors or effect modifiers.

Table 1: Data items/variables and data sources

Data category	Data item	Data source
Socio-demographic	Age	PNI or eSUS Notifica or SIVEP-Gripe
	Sex	PNI or eSUS Notifica or SIVEP-Gripe
Geographic	State of residence	PNI or eSUS Notifica or SIVEP-Gripe
Vaccination	Vaccine type	PNI
	Vaccine dose	PNI
	Vaccination date	PNI
Laboratory testing	RT-PCR test result	eSUS Notifica or SIVEP-Gripe
	Rapid antigen test result	eSUS Notifica or SIVEP-Gripe
	Date of symptoms onset	eSUS Notifica or SIVEP-Gripe
	Date of sample collection	eSUS Notifica or SIVEP-Gripe
Secondary care	Date of COVID-19 hospital admission	SIVEP-Gripe
Mortality	Date of COVID-19 death	SIVEP-Gripe
Comorbidities	Cardiac disease, diabetes mellitus, obesity, immunosuppression, and chronic kidney disease	eSUS Notifica or SIVEP-Gripe
Pregnancy	Pregnancy indicator and postpartum period	eSUS Notifica or SIVEP-Gripe

3.9.2 Consistency and error checking

We will examine the patterns of missingness and check for implausible values (e.g. date of second vaccine dose being earlier than the first) for all analytical variables being used, with a record maintained of reasons for exclusion of any records from analysis. In the case where a variable of interest has high levels of missingness, we may consider using alternative variables that are closely related as a proxy for these missing data.

3.9.3 Sample size

Based on a previous study (2), we anticipate having over 22 million individuals vaccinated with two doses of CoronaVac for our analyses.

4. Statistical analyses

4.1.1 Analytical procedures

We will conduct descriptive analyses to visually inspect trends in vaccination delivery, SARS-CoV-2 infection, and COVID-19 related hospitalisations and deaths, including by age groups. The descriptive analysis of vaccination delivery will include assessing the number of people who have received the two-dose schedule of CoronaVac, the BNT162b2 booster, and the length of follow-up available following the second CoronaVac dose and the BNT162b2 booster.

Generalized additive logistic regression will be used to estimate the odds ratio (OR) comparing the odds of vaccination between cases and controls, and the associated 95% confidence interval. This regression analysis will be adjusted for potential confounders such as age, sex, temporal trend, state of residence, previous infection, pregnancy, postpartum period, and comorbidities. We will compare the standard errors of the VE effect estimates from the fully adjusted model to those from the basic model, adjusted only for age and sex, to check for multicollinearity (7). The temporal trend will be estimated using the time elapsed between the study start and the date of symptoms onset. Temporal trend and age will be modeled as cubic regression spline smooth functions. VE will be estimated as 1-OR and expressed as a percentage.

4.1.2 Subgroup analysis

The analytical procedures described in section 4.1.1 will be applied to conduct a subgroup analysis by age (e.g., 18–59, 60–79, and ≥80 years), with models being estimated for each age group.

4.1.3 Sensitivity analysis

As a sensitivity analysis, we will include SARS-CoV-2 tests based on antigen detection in addition to RT-PCR tests. Antigen test has lower accuracy compared to RT-PCR. However, antigen tests have been progressively replacing the RT-PCR test.

4.2 Corrections for multiple testing

Following previous epidemiological recommendations (8), we will not correct for multiple testing since we are focused on effect estimation rather than hypothesis testing.

4.3 Missing data

We will examine the quantity and patterns of missing data in all covariates. We may use records with no item missingness (complete case analysis) or include a missing category. Missing data will be reported as percentages of total or raw numbers where possible.

4.4 Statistical software

Analyses will be conducted using the R statistical software.

5. Reporting results

5.1 Reporting guidelines and conventions

Results will be reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and Reporting of studies Conducted using Observational Routinely-collected Data (RECORD) guidelines. P-values will be quoted to two decimal places, unless they are less than 0.001 (whereby the p-value will be given as <0.001) or between <0.005 and >0.001, in which case they will be stated to three decimal places. Measures of association will be interpreted based on 95% CIs.

5.2 Dissemination

The analyses will be written in a manuscript and submitted to a peer-reviewed journal. We will also seek to provide feedback to relevant COVID-19 advisory bodies (e.g. the Brazilian Ministry of Health) as appropriate. All codes used in the analyses will be made publicly available via a GitHub repository.

6. Main changes from the first to the second version of the SAP

In the first version of the SAP we considered using cohort and TND study designs, but after methodological discussions with statisticians and epidemiologists, we decided to use only the TND. Compared to the cohort design, the TND has important advantages such as the possibility of controlling for healthcare seeking behaviour, access to testing, and case ascertainment (9). Furthermore, the TND has been more commonly applied and recommended for vaccine effectiveness studies (9). Another relevant change was the inclusion of the objective of estimating the effectiveness of the BNT162b2 vaccine as a booster dose after completion of the two-dose schedule of CoronaVac.

7. References

1. Williams I, Essue B, Nouvet E, Sandman L, Razavi SD, Noorulhuda M, et al. Priority setting during the COVID-19 pandemic: going beyond vaccines. BMJ Glob Heal. 2021;6(1):e004686.
2. Cerqueira-Silva T, Oliveira V de A, Pescarini J, Júnior JB, Machado TM, Ortiz RF, et al. The effectiveness of Vaxzevria and CoronaVac vaccines: A nationwide longitudinal retrospective study of 61 million Brazilians (VigiVac-COVID19). Preprint. 2021;

3. Vandenbroucke JP, Pearce N. Test-Negative Designs: Differences and Commonalities with Other Case-Control Studies with “Other Patient” Controls. Epidemiology. 2019;30(6):838–44.
4. Vandenbroucke JP, Brickley EB, Vandenbroucke-Grauls CMJE, Pearce N. A Test-Negative Design with Additional Population Controls Can Be Used to Rapidly Study Causes of the SARS-CoV-2 Epidemic. Epidemiology. 2020;31(6):836–43.
5. Hitchings MDT, Ranzani OT, Dorion M, D’Agostini TL, de Paula RC, de Paula OFP, et al. Effectiveness of ChAdOx1 vaccine in older adults during SARS-CoV-2 Gamma variant circulation in São Paulo. Nat Commun. 2021 Dec 28;12(1):6220.
6. Pastor-Barriuso R, Pérez-Gómez B, Hernán MA, Pérez-Olmeda M, Yotti R, Oteo-Iglesias J, et al. Infection fatality risk for SARS-CoV-2 in community dwelling population of Spain: nationwide seroepidemiological study. BMJ. 2020;m4509.
7. Greenland S, Daniel R, Pearce N. Outcome modelling strategies in epidemiology: traditional methods and basic alternatives. Int J Epidemiol. 2016 Apr 1;45(2):565–75.
8. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990;1(1):43–6.
9. World Health Organization. Evaluation of COVID-19 vaccine effectiveness: interim guidance, 17 March 2021 [Internet]. 2021. Available from: https://apps.who.int/iris/handle/10665/340301