Statistical Analysis Plan for Comparing COVID-19 Risk in Ubuntu/CoVPN 3008: Analysis of Follow-up since Enrollment or Month 6 Booster Dose through Month 12 or Month 18

(alphabetical order) Peter Gilbert, Yunda Huang, Aaron Hudson, Bo Zhang

Date finalized for signature: March 29, 2024
SAP effective upon last signature

Protocol version 8.0 (June 28, 2023)

SAP Version 2.0

SAP SIGNATURE PAGE

Author(s):

Legal Name	Yunda Huang
Job Title	Associate Professor
Signature & Date	See eTMF signature manifest

Legal Name	Aaron Hudson
Job Title	Assistant Professor
Signature & Date	See eTMF signature manifest

Legal Name	Bo Zhang
Job Title	Assistant Professor
Signature & Date	See eTMF signature manifest

Approvals:

Legal Name	Peter B. Gilbert
Job Title	Professor
Signature & Date	See eTMF signature manifest

SAP Modification History

The version history of, and modifications to, this statistical analysis plan are described below.

Date: Feb. 23, 2024
Protocol version: 8.0
SAP version: 1.0

Date: March 29, 2024
Protocol version: 8.0
SAP version: 2.0
Modifications:

- In Section 5, updated the definition of recent SARS-CoV-2 infection related to the availability of anti-NP titer data (criterion 3).
- In Section 7.1 and 11.4.1, clarified that TB status will not be adjusted for in analysis restricting to PLWoH due to data sparsity in the Cox regression analyses.
- In Section 7.1, added the minimum number of endpoints required to perform inferential analyses of the COVE- or CDC-based COVID-19 endpoint, and to perform covariate-adjusted analyses. These changes are also reflected in Table 2.

Contents

List of Tables

1 Introduction

This SAP focuses on comparisons of risk of COVID-19, and risk of severe COVID-19 (if sufficient number) between the randomized bivalent (mRNA-1273.222) vs. monovalent (mRNA-1273) Month 6 boost arms with endpoints counted up to the Month 12 and Month 18 visits in the Ubuntu/CoVPN 3008 study. Exploratory analyses will also include participants who did not receive the Month 6 boost, and/or received the monovalent boost outside of the monovalent:bivalent randomization scheme.

In the following, Section 2 describes the randomization scheme of the monovalent and bivalent boost arms stratified by study group. Section 3 summarizes different analysis cohorts. Section 4 describes the definitions of the time-to-event endpoints in general. Section 5 defines the baseline and Month 6 SARS-CoV-2 statuses. Section 6 defines primary and secondary study endpoints in detail. Section 7 describes analysis of follow up data only in Part B. Section 8 describes analysis of follow up data in Parts A+B combined. Section 9 summarizes all proposed primary, secondary and exploratory analyses in two tables. Sections 10 & 11 describe details of employed statistical methods for the analysis of study endpoints. Handling of multiplicity and missing data is provided at the end of the document.

2 Randomization and blinding

The Ubuntu study enrolled a total of 14,237 participants into 4 study groups based on their baseline HIV serostatus (positive or negative) and anti-spike SARS-CoV-2 Point-of-Care serology (POCS) status (positive or negative). Prior to protocol version 6.0, for Part A of the study, participants were assigned to receive 1 dose (Month 0), or 2 doses (Month 0 and Month 1) of mRNA-1273 based on their baseline POCS result being positive or negative, respectively. Prior to protocol version 6.0, for Part B of the study, randomization code was generated but all participants received the monovalent mRNA-1273 vaccine at Month 6 due to the bivalent vaccine not being available at the time, and no randomization assignment was actually performed. At the time of implementation of protocol version 6.0 after the bivalent vaccine became available, The Ubuntu study enrolled a total of 14,237 participants into 4 study groups based on their baseline HIV serostatus (positive or negative) and anti-spike SARS-CoV-2 Point-of-Care serology (POCS) status (positive or negative). Prior to protocol version 6.0, for Part A of the study, participants were assigned to receive 1 dose (Month 0), or 2 doses (Month 0 and Month 1) of mRNA-1273 based on their baseline POCS result being positive or negative, respectively. Prior to protocol version 6.0, for Part B of the study, randomization code was generated but all participants received the monovalent mRNA-1273 vaccine at Month 6 due to the bivalent vaccine not being available at the time, and no randomization assignment was actually performed. At the time of implementation of protocol version 6.0 after the bivalent mRNA-1273.222 vaccine became available, the previous randomization code was replaced with the randomization code generated under an updated randomization plan, ensuring that the treatment number assigned at original "randomization" was kept and corresponded to either mRNA-1273 or mRNA-1273.222. This allowed the site to use the randomized assignment to define the Month 6 vaccination for participants receiving the Month 6 vaccination under protocol version 6.0. An updated treatment key was provided to sites with participants who re-consented on protocol version 6.0, for documenting the Month 6 vaccinations only. Only as many treatment assignments as

needed for the number of participants randomized were generated in the treatment table and treatment keys. Participants receiving their Month 6 vaccination under Version 6 were blinded to their assignment to either the monovalent or bivalent vaccine. Participants who received Month 6 vaccination prior to Version 6 were not subject to this randomization. The randomization was done in blocks to ensure balance within each study group among participants who received their Month 6 vaccination under Version 6. More details can be found in a separate blinding management plan (Version 8.0) accessible only to unblinded statisticians prior to the unblinding of the study.

3 Analysis cohorts

The following analysis cohorts will be considered:

- Randomized Month 6 cohort (RM6): this cohort is the primary Part B post-M6 stage analysis cohort, and it includes all FAS participants who received a Month 6 vaccination under the randomization scheme.
- Full Month 6 cohort (FM6): this cohort is the exploratory Part B post-M6 stage analysis cohort, and it includes all FAS participants who received a Month 6 vaccination, regardless of in or outside of the randomization scheme.
- Full per-protocol cohort (FPP): this cohort is the primary Part A+B combined analysis cohort, and it includes all participants who received their M0, M1 (if applicable) and M6 vaccinations as assigned without major protocol violations. All participants in this cohort should have received their Month 6 vaccination according to the randomization scheme.
- Full Analysis Set (FAS): this cohort is the exploratory Part A+B combined analysis cohort, and it includes all participants who enrolled into the study (excluding 235 participants from one site due to data quality concerns).

4 Time-to-event endpoints

For all primary and secondary comparisons and a subset of exploratory analyses, the time origin for time-to-event COVID-19 or severe COVID-19 endpoints is either the day of or 13 days after the Month 6 vaccination visit, counting events either starting 1 day or 14 days after the Month 6 vaccination. For a subset of exploratory analyses that will include information from enrollment until Month 12 or Month 18, the time origin is either the date of enrollment or 13 days after the Month 6 vaccination visit, counting events either starting 1 day after enrolment or 14 days after the Month 6 vaccination. The Cox model approach will always using the calendar time scale and the cumulative incidence approach will always use the study time scale. See more details about these two approaches in later sections. These analyses will address the co-primary objectives, key secondary objectives and exploratory objectives specified in the protocol (Version 8.0).

Different analyses specify different time origins for the COVID-19 or severe COVID-19 (if applicable) time-to-event T; the endpoint date is always the same: the date of the first occurrence of COVID-19 or severe COVID-19 endpoint after the corresponding time-origin and up to the date of Month 12 or Month 18 visit. The failure time T is right-censored by the first event of receipt

of outside vaccination, loss to follow-up, or reaching the upper window of the Month 12 (for the Month 12 analyses) or Month 18 visit (for the final analyses), whichever occurs first.

5 SARS-CoV-2 status at baseline and at Month 6

This section defines baseline and Month 6 SARS-CoV-2 statuses, as well as recent SARS-CoV-2 infection status at Month 6 for the pre- & post-M6 stage combined analyses and for the post-M6 stage analyses. These variables will be used as a stratification variable to define different analysis groups, evaluated as a covariate to adjust for prior exposure, and/or as a potential effect modifier in various analyses.

Baseline SARS-CoV-2 status is defined in the same way as in the Pre-Month 6 stage SAP. Specifically, in addition to the Point-of-Care serology (POCS) testing result, which was used for assigning participants into the 4 study groups at enrolment, the central lab serology (CLS) testing result and the Nucleic Acid Amplification Test (NAAT) result are also used to define the SARS-CoV-2 naïve and non-naïve status in subsequent statistical analyses. Specifically, a participant is said to be baseline SARS-CoV-2 naïve (or negative) if all 3 test are negative. A participant is said to be non-naïve (or positive) if at least 1 of the 3 tests is positive. We will further differentiate between two types of non-naïve participants: (i) those who tested negative for POCS but positive for either CLS or NAAT; and (ii) those who tested positive for POCS. These definitions will be applied after imputing missing baseline CLS and NAAT testing results according to the following imputation scheme. For participants who missed the CLS result but had both POCS and NAAT results, the missing CLS result was imputed as positive if either POCS or NAAT was positive and negative if both POCS and NAAT were negative. For participants who missed the NAAT result but had both POCS and CLS results, the missing NAAT result was always imputed as negative. For participants who missed both CLS and NAAT test results, their missing CLS test result was imputed as positive if POCS result was positive and negative if the POCS result was negative. Their NAAT result was always imputed as negative. This imputation scheme was motivated by the joint distribution of POCS, CLS and NAAT among those who had complete measurements of test results (more than 98% of the study cohort); see Table 1 for various empirical conditional probabilities. We also considered these conditional probabilities separately in the stratum defined by baseline HIV status; results were qualitatively similar and the derived imputation scheme was the same.

Month 6 SARS-CoV-2 status, a new variable for Part B analyses, is defined as hybrid immunity or vaccine immunity. Specifically, a participant is said to have Month 6 hybrid immunity if they have any evidence of prior SARS-CoV-2 infection up to the Month 6 visit as defined by meeting any of the following criteria: 1) baseline overall SARS-CoV-2 status is positive as defined above, 2) at least one NAAT result at or prior to Month 6 is positive, or 3) at least one CLS at or prior to Month 6 is positive. Otherwise, a participant is said to have Month 6 vaccine immunity. In another word, a participant is considered having only vaccine immunity at Month 6 if absent of any evidence of SARS-CoV-2 infection from enrollment up to Month 6. Of note, POCS results are expected to be available from all participants only for their M0 samples; CLS results are expected to be available from all participants for both their M0 and M6 samples.

Lastly, recent SARS-CoV-2 infection status at Month 6, another new variable for Part B analyses,

Table 1: **Panel A**: Empirical probability that CLS or NAAT test result was positive conditional on different configurations of the other two test results among $13,857/14,040$ participants with all three test results measured at baseline. **Panel B**: Empirical probability that CLS or NAAT was positive conditional on POCS test result among $13,929/14,040$ participants with both CLS and POCS test results and among $13,915/14,040$ participants with both NAAT and POCS test results, respectively.

	Value
Panel A	
$P(\text{CLS} = 1 \mid \text{POCS} = 0 \ \& \ \text{NAAT} = 0)$	0.29
$P(\text{CLS} = 1 \mid \text{POCS} = 1 \ \& \ \text{NAAT} = 0)$	0.58
$P(\text{CLS} = 1 \mid \text{POCS} = 1 \ \& \ \text{NAAT} = 1)$	0.80
$P(\text{CLS} = 1 \mid \text{POCS} = 0 \ \& \ \text{NAAT} = 1)$	0.59
$P(\text{NAAT} = 1 \mid \text{POCS} = 0 \ \& \ \text{CLS} = 0)$	0
$P(\text{NAAT} = 1 \mid \text{POCS} = 1 \ \& \ \text{CLS} = 1)$	0.07
$P(\text{NAAT} = 1 \mid \text{POCS} = 1 \ \& \ \text{CLS} = 0)$	0
$P(\text{NAAT} = 1 \mid \text{POCS} = 0 \ \& \ \text{CLS} = 1)$	0.23
Panel B	
$P(\text{CLS} = 1 \mid \text{POCS} = 0)$	0.32
$P(\text{CLS} = 1 \mid \text{POCS} = 1)$	0.59
$P(\text{NAAT} = 1 \mid \text{POCS} = 0)$	0.07
$P(\text{NAAT} = 1 \mid \text{POCS} = 1)$	0.04

is defined as positive or negative. A participant is said to have had a recent SARS-CoV-2 infection within the last ∼6 months (i.e., a new infection during Part A of the study) if any of the following criteria are satisfied: 1) NAAT-negative at baseline and NAAT-positive at anytime after baseline before or at the Month 6 vaccination visit; 2) baseline anti-NP (CLS) sero-negative, but anti-NP sero-positive at Month 6; and 3) baseline anti-NP sero-positive and anti-NP titer at Month 6 is at least 4 times of that at baseline. Of note, due to the anti-NP titer data being not clean/final, criterion 3 is not used in the Month 12 analyses and may be only used in the Month 18 analyses.

6 Study endpoints

This section defines the co-primary COVID-19 and severe COVID-19 endpoints. The same definitions described in the pre-M6 SAP are used. In future SAP versions, more details on other secondary endpoints may be described, including asymptomatic/subclinical SARS-CoV-2 infection, overall SARS-CoV-2 infection regardless of symptomology, as well as new infection confirmed by genetic sequence data.

6.1 Co-Primary COVID-19 endpoint

According to Version 8 of the protocol, all symptomatic COVID-19 endpoints will be evaluated using two case definitions, one used in the COVE study and the other using a case definition based on criteria from the CDC, with the latter being the leading definition given the lighter-symptom

nature of omicron infections observed and anticipated from the study. No multiplicity adjustments are applied to adjust for analyses of multiple COVID-19 endpoints.

6.1.1 COVE-based COVID-19 endpoint definition

The COVE-based primary COVID-19 endpoint is defined as the first occurrence of adjudicated symptom-triggered NAAT-confirmed COVID-19 based on the following criteria (same as in the Moderna mRNA-1273 COVE trial):

- At least TWO of the following systemic symptoms: Fever $\geq 38°C$), chills, myalgia, headache, sore throat, new loss of taste or smell, OR
- At least ONE of the following respiratory signs/symptoms: cough, shortness of breath or difficulty breathing, OR clinical or radiographical evidence of pneumonia; AND
- At least ONE (post-baseline) nasal swab (or respiratory sample, if hospitalized) positive for SARS-CoV-2 by NAAT.

The date of a documented/adjudicated COVID-19 endpoint is the **later** date of a qualifying symptom and the date of the first confirmatory positive NAAT. This (event) date is supplied by the Endpoint Adjudication Committee.

6.1.2 CDC-based COVID-19 endpoint definition

The CDC-based definition of the COVID-19 endpoint is defined as symptomatic NAAT-confirmed COVID-19 based on CDC criteria, where the nucleic acid test NAAT may or may not be triggered by symptoms. The CDC-based COVID-19 endpoint is defined as the first occurrence of symptomatic NAAT-confirmed COVID-19 based on the following criteria:

- At least ONE of the following systemic or respiratory symptoms: Fever ($\geq 38°C$), chills, cough, shortness of breath and/or difficulty breathing, fatigue, muscle and/or body aches [not related to exercise], headache, new loss of taste/smell, sore throat, congestion, runny nose, nausea, vomiting, or diarrhea; AND
- At least ONE (post-baseline) nasal swab (or respiratory sample, if hospitalized) positive for SARS-CoV-2 by NAAT.
- The symptom date and NAAT positive date must be within 14 days of each other. Specifically, if the onset of qualifying symptom(s) occurred prior to the specimen collection date of a NAAT-positive result, then the date of last reported symptom(s) must be after 14 days prior to the NAAT-positive date. If the onset of qualifying symptoms occurred after a NAAT-positive date, then the onset date of qualifying symptom(s) must be within 14 days of the NAAT-positive date.

The date of a COVID-19 endpoint according to the CDC-based definition is the **earlier** date of a qualifying symptom and the date of the first confirmatory positive NAAT that satisfy the criteria stated above.

For the Part B or the Part A+B analyses, although additional NAAT testing was performed at the M1 vaccination visit for study groups 1 & 3, not for study group 2 or 4, the randomization

between the bivalent and monovalent Month 6 boost arms was preserved within each study group. Therefore, the primary analyses in the RM6 cohort comparing between the randomized bivalent and monovalent boost arms will use the CDC1 endpoint that include NAAT testing results at Month 1. The CDC2 endpoints derived to exclude NAAT testing results at Month 1 for comparisons of COVID risk between different groups will only be applicable to the exploratory analyses that include also participants who missed the Month 6 vaccination, or received the monovalent boost outside of the randomization scheme.

6.2 Severe COVID-19 endpoint

The severe COVID-19 endpoint is defined as the first occurrence of a COVE-based COVID-19 having at least one sign, symptom, or other evidence of severe disease (e.g., respiratory failure, shock, intensive care unit admission). The date of a documented/adjudicated severe COVID-19 endpoint is the later date of a symptom and the date of positive NAAT. This (event) date is supplied by the Endpoint Adjudication Committee.

6.3 SARS-CoV-2 infection endpoint

6.3.1 SARS-CoV-2 infection regardless of symptomology

Two definitions are considered for the SARS-CoV-2 any-infection endpoint regardless of symptomology. Definition 1 is the primary definition. Definition 2 is an exploratory definition.

- Definition 1: SARS-CoV-2 any-infection is defined in all participants who are:
  - baseline NAAT negative but tested NAAT positive at at least another post-baseline time-point, either triggered by symptom or performed at routine swabbing visits; OR
  - baseline NAAT positive but detected NAAT positive when either a) there are two consecutive negative NAAT results between the baseline and post-baseline positive swabs, or b) a positive post-baseline NAAT result is found at least 90 days after a single NAAT negative swab. Future analyses incorporating sequence information may also consider multiple occurrences; OR
  - baseline anti-N CLS negative but detected anti-N CLS positive post-baseline; OR
  - baseline anti-N CLS positive and detected anti-N CLS positive with anti-N IgG concentration at least 1.5 times higher than that at baseline. When the quantitative readout from the CLS assays is not available, this last criterion will not be considered.
- Definition 2: SARS-CoV-2 infection is defined only in participants with negative baseline SARS-CoV-2 status as defined in Section 5 based on the 3 tests (NAAT, POCS and CLS). A SARS-CoV-2 infection endpoint is any post-baseline SARS-CoV-2 infection detected positive by anti-N CLS, OR by NAAT (triggered by symptom, routine testing or other reasons).

6.3.2 Asymptomatic/subclinical SARS-CoV-2 infection

TBD.

6.3.3 Definition of a new infection endpoint based on genetic sequence data

TBD.

7 Part B analyses: post-M6 stage

All analyses in the post-M6 stage will start counting events either 1 or 13 days after the Month 6 vaccination date till the Month 12 visit for the Month 12 analyses, and till the final study visit at Month 18 for the final analyses. The Cox model approach will always use the calendar time scale and the cumulative incidence approach will always use the study time scale. These analyses will not adjust for multiple looks of the data (at Month 12 and Month 18) given the nature of the study being not to license a product.

7.1 Primary and secondary analyses comparing randomized bivalent vs. monovalent boost arms

For post-M6 follow-up in Part B of the study, the primary analysis is the comparison of COVID-19 risk between the monovalent and bivalent randomized arms, counting events starting 14 days after the Month 6 vaccination in the Randomized Month 6 (RM6) analysis cohort. When there are at least 7 severe COVID-19 cases, the same analysis will be repeated for the severe COVID-19 endpoint as a co-primary endpoint but without covariate adjustment due to the expected small number of severe COVID-19 endpoints. The secondary analysis involves the same comparison of COVID-19 risk and severe COVID-19 (if applicable), but counting events starting 1 day after the Month 6 vaccination in the RM6 analysis cohort. Both the primary and secondary analyses will only include participants who received the Month 6 vaccination in Version 6 under the study-group-stratified randomization scheme.

Two types of estimation procedures are used in each of the primary and secondary analyses. We consider estimation of hazard ratios using Cox proportional hazards regression, as well as estimation of absolute and relative differences in cumulative incidence of infection using methodology recently proposed by Westling et al. (2023).

Calendar-time-based Cox proportional hazards models will be employed to estimate the hazard ratio (HR) between the bivalent vs. monovalent randomized arms. The model will be stratified by the 4 study groups and region (RSA vs. other), thus allowing for a different baseline hazard for each of the $4 \times 2 = 8$ strata. Since the comparison is between randomized groups, the main analysis will only adjust for a small number of covariates that are deemed important or known to be associated with COVID-19 risk to increase efficiency in statistical inference. In particular, sex-assigned-at-birth (Female vs. Male) and baseline history of TB (yes vs. no) were found associated with CDC endpoints in Part A analyses, and having had a recent infection is expected to be associated with COVID-19 risk. Therefore, sex-assigned-at-birth, history of TB at Month 6, and recent SARS-CoV-2 infection status at Month 6 (see definition in Section 5) will be adjusted for in all primary and secondary analyses comparing randomized bivalent vs. monovalent boost arms. These analyses will also be repeated within people who live with HIV (PLWH) and HIV-negative participants, separately. For analyses among PLWH, CD4 counts ($\leq$ vs. $>$ 500 cells/mm^3 at Month 6) and HIV viral load (detectable vs. not detectable at Month 6) will be further adjusted for in

addition to sex assigned at birth, history of TB and recent infection status. For analyses among HIV-negative participants, a subset of the covariates maybe considered to avoid data sparsity, with recent SARS-CoV-2 status as the leading covariate to consider. Specifically, for Cox regression among HIV-negative participants (PLWoH), we will not adjust for TB status due to data sparsity.

We also make comparisons of covariate-adjusted cumulative incidence curves for each comparison arm. We provide a full description of the covariate-adjustment strategies in Section 11.3.

For both the Cox regression approach and the cumulative incidence approach, if there are fewer than 20 COVE-/CDC-based COVID-19 endpoints, then the corresponding endpoint will only be studied descriptively. In a pre-specified Cox regression analysis, if the number of endpoints is fewer than the number of covariates to be adjusted multiplied by 5, then this analysis will only be conducted descriptively. Only studied descriptively means that numbers of endpoints by treatment arm and covariate strata are reported, but no estimation or testing of treatment effect parameters is performed.

7.2 Exploratory analyses comparing bivalent and monovalent boost arms

For post-M6 follow-up in Part B of the study, exploratory comparisons of COVID-19 and severe COVID-19 risk between the bivalent and monovalent boost arms will consider the Full Month 6 cohort (FM6) and count events starting 14 days after the Month 6 vaccination. Because these comparisons are not between randomized arms, covariate-adjusted study-group-and-region stratified Cox models and covariate-adjusted cumulative incidence analysis approaches will be employed. The same list of covariates described above for the primary and secondary analyses will be considered. Exploratory analyses will be performed inferentially using the cumulative incidence approach, but covariate adjustment will only be made in the presence of enough failure events (because of the stability issue of the resulting estimator); details can be found in Section 11.3.

The same exploratory analyses will be repeated within people living with HIV (PLWH) and HIV-negative participants, separately.

8 Parts A+B analyses: Pre- & post-M6 stage combined

All described analyses below will start counting events after study enrolment or 13 days after the Month 6 vaccination, based on follow-up till the Month 12 visit and the Month 18 visit. Again, the Cox model approach will always using the calendar time scale and the cumulative incidence approach will always use the study time scale. These analyses will not adjust for multiple looks of the data.

8.1 Primary analysis in the full per-protocol cohort

For analyses including information from both the pre- & post-M6 stages, the primary analysis is the comparison of COVID-19 risk between the bivalent and monovalent randomized arms in the full per-protocol (FPP) analysis cohort as defined in Section 3, using 13 days after the Month 6 vaccination as the time origin and counting events starting 14 days after the Month 6 vaccination. The FPP cohort includes a subset of participants in RM6 who received the Pre-Month 6 and Month

6 vaccinations according to the protocol without any major protocol violations. When there are at least 7 severe COVID-19 cases, the same analysis will be repeated for the severe COVID-19 endpoint as a co-primary endpoint.

The same covariate-adjusted, calendar-time-based Cox proportional hazards models as described in Section 7.1 will be employed to estimate the hazard ratio (HR) between the bivalent vs. monovalent randomized arms. The same analysis will be repeated for people living with HIV (PLWH) and for HIV-negative participants, separately. The same covariate-adjusted cumulative incidence curves as described in Section 7.1 for each comparison arm will be provided.

8.2 Exploratory analyses

For follow-up including both the pre- & post-M6 stages, two categories of exploratory analyses will be considered. One refers to the exploratory comparisons of COVID-19 and severe COVID-19 between the bivalent and monovalent arms including all participants who received a Month 6 vaccination either in or outside of the randomization scheme. The other refers to the exploratory comparisons of COVID-19 and severe COVID-19 between the hybrid vs. vaccine immunity groups by Month 6 vaccination type (bivalent, monovalent or none) including all enrolled participants. Of note, these exploratory analyses may not be all performed for the primary publication of Part B data.

8.2.1 Exploratory analyses comparing bivalent and monovalent boost groups

For follow-up including both the pre- & post-M6 stages, exploratory analyses of COVID-19 and severe COVID-19 between the bivalant vs. monovalent boost groups will be performed in the full Month 6 (FM6) analysis cohort including all participants who received the Month 6 vaccination either in or outside of the randomization scheme, using Month 6 vaccination date as the time origin and counting events starting 1 day after Month 6 vaccination. Because these comparisons are not between randomized arms, covariate-adjusted study-group-and-region-stratified Cox models using calendar time-scale and covariate-adjusted cumulative incidence analysis approaches will be employed as described above for the primary and secondary analyses that compare between bivalent and monovalent groups.

The same analysis may be repeated for PLWH and for HIV-negative participants, separately. For analyses among PLWH, in addition to covariates described above, baseline CD4 counts ($\leq$ vs. $>$ 500 cells/mm^3) and baseline HIV viral load (detectable vs. not detectable) will be further adjusted for. For analyses among HIV-negative participants, a subset of the covariates maybe considered to avoid data sparsity, with recent SARS-CoV-2 status as the leading covariate to consider.

8.2.2 Exploratory analyses comparing hybrid and vaccine immunity groups among PLWH

For analyses including information from both the pre- & post-M6 stages, exploratory analyses of COVID-19 and severe COVID-19 between hybrid vs. vaccine immunity groups will consider two definitions of hybrid/vaccine immunity groups. The first is based on information at baseline and follows the same definition used in the Part A SAP for AG1 (Vaccine Immunity) and AG2.1 (Hybrid Immunity) and repeated below for easy reference.

1. AG1 (PLWH Vaccine Immunity; HIV+,SARS2-,2d) = PLWH; baseline SARS-CoV-2 naïve (negative for POCS, CLS, and NAAT); assigned 2 vaccinations pre-Month 6
2. AG2.1 (PLWH Hybrid Immunity; HIV+,SARS2+,1d) = PLWH; baseline POCS positive; assigned 1 vaccination pre-Month 6

The second definition of hybrid/vaccine immunity groups is based on a combination of baseline hybrid/vaccine immunity group information, Month 6 SARS-CoV-2 status and Month 6 recent SARS-Cov-2 infection status (see definitions in Section 5). Specifically, Month 6 Vaccine Immunity group includes a subset of AG1 participants who had no evidence of SARS-CoV-infection before or at Month 6 (i.e., Month 6 SARS-CoV-2 status = vaccine immunity); Month 6 Hybrid Immunity group includes a subset of AG2.1 participants who had no evidence of recent SARS-CoV-2 infection at Month 6 (i.e., Month 6 recent infection status = negative).

Comparing baseline hybrid vs.vaccine immunity groups among PLWH

For this exploratory analysis of COVID-19 and severe COVID-19 between baseline hybrid vs. vaccine immunity groups using the first definition described above, the same analysis will be repeated for three subsets of participants: 1) participants who received the monovalent Month 6 booster; 2) participants who received the bivalent Month 6 booster; and 3) participants who did not receive the Month 6 vaccination. An analysis will be performed only if there are at least 7 endpoints in a given subset of participants. These analyses seek to understand if participants' COVID-19 and severe COVID-19 risk differs between baseline hybrid and vaccine immunity groups, counting events starting 1 day after enrollment till Month 12 or Month 18 (as opposed to Month 6 in the Part A analyses). All covariates considered in the Part A analyses of hybrid vs. vaccine immunity will be considered. In addition, Month 6 vaccination will be included as a time-varying covariate using the Cox model approach. The cumulative incidence approach will not be considered for these exploratory analyses.

Comparing Month 6 hybrid vs.vaccine immunity groups among PLWH

For this exploratory analysis of COVID-19 and severe COVID-19 between Month 6 hybrid vs. vaccine immunity groups based on the 2nd definition described above, the same analysis will be repeated for two subsets of participants: 1) participants who received the monovalent Month 6 booster; and, 2) participants who received the bivalent Month 6 booster. An analysis will be performed only if there are at least 7 endpoints in a given subset of participants. These analyses seek to understand if participants' COVID-19 and severe COVID-19 risk after the Month 6 vaccination differs between hybrid and vaccine immunity groups defined at Month 6, counting events starting 14 days after the Month 6 vaccination date. All covariates considered in the Part A analyses of hybrid vs. vaccine immunity will be considered, except that the covariate values are defined at Month 6 whenever applicable. Only the Cox model approach will be considered for these analyses; the cumulative incidence approach will not be considered.

9 Summary of primary, secondary and exploratory analyses

Table 2 summarizes the target population, study endpoints, and time frame for various primary and secondary analyses. Table 3 further summarizes the same information for various exploratory

analyses.

	Primary analysis I (applicable when # COVID-19 endpoints is at least 20)
Target population	Participants randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	COVE-/CDC-based COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
	Primary analysis II (applicable when # COVID-19 endpoints is at least 20)
Target population	Full per-protocol participants randomized to monovalent or bivalent under Protocol Version 6 (FPP)
Endpoint	COVE-/CDC-based COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
	Primary analysis III (applicable when # severe endpoints is at least 7)
Target population	Participants randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	Severe COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
	Primary analysis IV (applicable when # severe endpoints is at least 7)
Target population	Full per-protocol participants randomized to monovalent or bivalent under Protocol Version 6 (FPP)
Endpoint	Severe COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
	Secondary analysis I (applicable when # COVID-19 endpoints is at least 20)
Target population	Participants randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	COVE-/CDC-based COVID-19 endpoints counted 1 day after M6 vaccination
Time frame	M6 vaccination to M12/M18 visit
	Secondary analysis II (applicable when # severe endpoints is at least 7)
Target population	Participants randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	Severe COVID-19 endpoints counted 1 day after M6 vaccination
Time frame	M6 vaccination to M12/M18 visit
	Secondary analysis III
Target population	PLWH/PLWoH randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
	Secondary analysis IV
Target population	PLWH/PLWoH randomized to monovalent or bivalent under Protocol Version 6 (RM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after M6 vaccination
Time frame	M6 vaccination to M12/M18 visit
	Secondary analysis V
Target population	Full per-protocol PLWH/PLWoH randomized to monovalent or bivalent under Protocol Version 6 (FPP)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 14 days after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit

Table 2: Summary of target population and study endpoints for primary and secondary analyses. All analyses compare the bivalent vs. monovalent Month 6 boost arms.

10 Parameters of interest to estimate

10.1 Cumulative-failure-time-based parameters of interest

Let T be the failure time using the study time scale (based on one of the study time origins defined above), t_0 a fixed time point of interest after the given time origin, and W baseline covariates. For a given analysis, we denote by A a binary indicator of belonging to either of strata which define the comparison groups of interest. In the primary analysis A is indicator of receiving a bivalent or monovalent M6 booster (in analyses), and it is similarly defined elsewhere. We denote by $T(a)$ the counterfactual failure time if $A = a$.

In a cumulative-incidence-based analysis, parameters of interest are the marginalized failure time

Comparing bivalent vs. monovalent booster	
Exploratory analysis I	
Target population	FAS participants receiving a M6 vaccine regardless of randomization scheme (FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after M6 vaccination
Time frame	M6 vaccination to M12/M18 visit
Exploratory analysis II	
Target population	FAS PLWH/PLWoH receiving a M6 vaccine regardless of randomization scheme (FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after M6 vaccination
Time frame	M6 vaccination to M12/M18 visit
Comparing baseline hybrid vs. vaccine immunity among PLWH	
Exploratory analysis III	
Target population	AG1 and AG2.1 receiving a bivalent M6 vaccine regardless of randomization scheme (AG1 and AG2.1 in FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after enrollment
Time frame	Enrollment to M12/M18 visit
Exploratory analysis IV	
Target population	AG1 and AG2.1 receiving a monovalent M6 vaccine regardless of randomization scheme (AG1 and AG2.1 in FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after enrollment
Time frame	Enrollment to M12/M18 visit
Exploratory analysis V	
Target population	AG1 and AG2.1 not receiving M6 vaccination regardless of randomization scheme (AG1 and AG2.1 in FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 1 day after enrollment
Time frame	Enrollment to M12/M18 visit
Comparing Month 6 hybrid vs. vaccine immunity among PLWH	
Exploratory analysis VI	
Target population	AG1 and AG2.1 receiving a monovalent M6 vaccine regardless of randomization scheme (AG1 and AG2.1 in FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 14 day after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit
Exploratory analysis VII	
Target population	AG1 and AG2.1 receiving a bivalent M6 vaccine regardless of randomization scheme (AG1 and AG2.1 in FM6)
Endpoint	COVE-/CDC-based/Severe (when applicable) COVID-19 endpoints counted 14 day after M6 vaccination
Time frame	13 days post M6 vaccination to M12/M18 visit

Table 3: Summary of target population and study endpoints for exploratory analyses

cumulative distribution functions $E_W[P(T \leq t_0 \mid A = a, W)]$ over a range of fixed time points t_0, and the contrasts $\theta^{\text{diff}}(t_0) = E_W[P(T \leq t_0 \mid A = 1, W)] - E_W[P(T \leq t_0 \mid A = 0, W)]$ and $\theta^{rr}(t_0) = E_W[P(T \leq t_0 \mid A = 1, W)]/E_W[P(T \leq t_0 \mid A = 0, W)]$. When the treatment assignment A is randomized, e.g., in Primary Analysis I of the randomized cohort RM6, the statistical parameter $E_W[P(T \leq t_0 \mid A = a, W)]$ coincides with the causal parameter $E_W[P(T(a) \leq t_0 \mid W)]$ and is the ultimate estimand of interest. In exploratory analyses involving non-randomized comparison groups (i.e., exploratory analyses I, II, and III in Table 3), the statistical parameter $E_W[P(T \leq t_0 \mid A = a, W)]$ equals the causal parameter $E_W[P(T(a) \leq t_0 \mid W)]$ under standard causal identification assumptions, and $\theta^{\text{diff}}(t_0)$ and $\theta^{rr}(t_0)$ are two versions of an average treatment effect.

Even if the causal assumptions fail, the parameter $E_W[P(T \leq t_0 \mid A = a, W)]$ may still be of greater scientific interest than the unadjusted survival probability. The first reason for this is that this parameter allows for the adjustment of covariates related to both A and T. As a result, this parameter can be interpreted as the average probability that $T \leq t_0$ in a hypothetical population of participants with $A = a$ but with a distribution of the covariate vector W identical to that in the target population. The second reason is that adjusting for W allows the relaxation of the marginal independent censoring assumption $T(a) \perp C(a)|A$ to a conditional independent censoring assumption $T(a) \perp C(a)|A, W$. This relaxation can be important in contexts where the event and censoring times may be dependent, but the recorded covariate vector W at least partly explains the dependence between them.

10.2 Selection of the final time point t_0

For analyses using study enrolment as the time origin, and for a given group comparison, the fixed time point t_0 is chosen to be minimum of 365 (or 540) days and the latest possible time point when stable estimation using follow-up data through the Month 12 (or Month 18) visit can be reasonably assured. For analyses using 13 days after the Month 6 vaccination as the time origin, and for a given group comparison, the fixed time point t_0 is chosen to be minimum of 165 (or 365) days and the latest possible time point when stable estimation using follow-up data through the Month 12 or Month 18 study visit can be reasonably assured. Identifying the minimum t_0 for which stable estimation can be assured is operationalized by selecting the latest point for which the standard error for the log risk ratio satisfies

$$\left| \frac{\log(3)}{\widehat{SE}(t)} \right| \geq z^{1-\alpha/2},$$

where $\widehat{SE}(t)$ denotes the estimated standard error for the log of the cumulative risk ratio at time t and $z^{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution. The rationale is we would only consider an estimate valuable if we have enough precision to conclude that a large cumulative risk ratio (of size 3.0 or larger) is statistically significant at the level α. Study participants with an endpoint observed after t_0 are considered to have censored outcomes.

It is expected that, for analyses using the severe COVID-19 endpoint, the standard error may be very large for all time points due to a low event rate. Therefore, for analyses using study enrolment as the time origin, the time point t_0 is fixed at 365 (or 540) days for analyses through Month 12 (or Month 18), and for analyses using 13 days after the Month 6 vaccination as the time origin, t_0 is chosen to be 165 (or 365) days for analyses through Month 12 (or Month 18). That is, no data-adaptive selection of t_0 is performed when using the severe COVID-19 endpoint.

10.3 Cox-regression-based parameters of interest

We consider parameters of interest defined under the assumption of a proportional hazards model on a calendar time scale. Cox regression allows for flexible nonparametric modeling of the baseline hazard as a function of calendar time, acknowledging that the COVID-19 epidemic undergoes unpredictable changes over calendar time.

Suppose there are J baseline strata, $j = 1, \cdots, J$. The proportional hazards model can be written as

$$\lambda_j(t \mid A, W) = \lambda_{0j}(t) \exp\left\{\beta A + \gamma^T W\right\} \tag{1}$$

for a binary treatment assignment A. In addition to the parameter β in Model (1), parameters γ of variables like an indicator of having an infection during Part A of the study (see Section 7.1 for details) or CD4 counts/viral load will be reported in case they are of interest.

The β coefficients for the calendar-time-based Cox models are the coefficients for which results are reported; results for the β coefficients for the study time Cox modeling are not reported, to avoid too many results and potential confusion. Therefore, the use of the Cox model on the study/enrolment

time scale is solely as an intermediate step in covariate-adjusted cumulative incidence estimation, given that the cumulative incidence results use study/enrolment time.

For cumulative incidence based analyses discussed in Section 10.1, a final time point t_0 was selected and failure times beyond t_0 were right-censored at t_0. The calendar-time Cox models do not use this fixed value of t_0; rather calendar-time Cox models consider all follow-up up through the M6 vaccination, subject to right-censoring described in Section 4.

11 Statistical methods for estimation/inference

11.1 Set of statistical analyses

The data analyses are set up differently depending on the analysis goal and analysis cohort.

First, the approach to comparing the bivalent and monovalent booster relative efficacy based on the analysis cohort consisting of both PLWH and PLWoH will be described. Analyses that fall under this category include:

1. Primary analyses I/II and III/IV (when applicable);
2. Secondary analyses I and II (when applicable);
3. Exploratory analysis I.

Second, the approach to comparing the bivalent and monovalent booster relative efficacy but restricted to PLWH or PLWoH will be described. Analyses that fall under this category include:

1. Secondary analyses III, IV and V;
2. Exploratory analysis II.

Lastly, the exploratory analyses assessing if the COVID-19 incidence after receiving a M6 booster (bivalent/monovalent) or not receiving a M6 booster differed between PLWH with hybrid (Analysis Group 2-1) or vaccine-induced immunity (Analysis Group 1) will be described. Analyses that fall under this category include:

1. Exploratory analyses III, IV and V.

The following statistical approaches are employed when evaluating the bivalent vs. monovalent booster relative efficacy.

1. Estimation of $E_W[P(T(a) \leq t_0 \mid W)]$, $a = 0, 1$, $\theta^{\text{diff}}(t_0)$, $\theta^{rr}(t_0)$, using nonparametric methods.
2. Estimation of the same parameters as estimated in 1 above except using semiparametric Cox modeling with T being the study time. While this approach also provides estimates of β coefficients, they are not used for results so as to avoid complications in interpretations when also reporting results for β coefficients for the calendar-time Cox modeling noted next.
3. Cox proportional hazards modeling with failure time on the calendar time scale and exposure variable the booster indicator (bivalent vs. monovalent), for estimation of β coefficients (this analysis does not estimate marginalized cumulative failure time parameters).

11.2 Hypothesis tests

Each nonparametric method described above provides 95% confidence intervals about the parameters of interest as well as two-sided p-values for testing the following null hypotheses of interest (with t_0 fixed at the latest time point of interest):

$$(A \text{ dichotomous}) \quad H_0: \quad \theta^{\text{diff}}(t_0) = 0, \quad \theta^{\text{rr}}(t_0) = 1 \tag{2}$$

Each Cox modeling method provides two-sided p-values for testing the following null hypotheses of interest:

$$(A \text{ dichotomous}) \quad H_0: \quad \beta = 0 \tag{3}$$

A partial likelihood ratio test will be used for testing each of the Cox model null hypotheses.

11.3 Cumulative failure time approach

Below, we first describe our general approach for estimating the cumulative incidence parameter, and we subsequently provide details describing how this approach is applied for each of the contrasts of interest.

We estimate the marginalized cumulative incidence function $\Gamma^a(t) = E_W[P(T \leq t|A = a, W)]$ using the strategy proposed by Westling et al. (2023). Estimating the marginalized cumulative incidence function requires estimation of the following three nusiance functions:

1. The conditional probability of the outcome by or prior to time t, given analysis group and covariates, $\eta_1(t, a, w) = P(T \leq t|A = a, W = w)$
2. The conditional probability of censoring by or prior to time t, given analysis group and covariates, $\eta_2(t, a, w) = P(C \leq t|A = a, W)$ (where C is the censoring time)
3. The conditional probability of analysis group given the covariate vector, $\pi(a|w) = P(A = a|W = w)$ also referred to as the propensity score

The approach we consider allows for us to supply nonparametric estimators for each of the above nuisances, which could be obtained using a flexible machine learning procedure (e.g., the Super Learner). In this case, our estimator of $\Gamma^a(t)$ is fully nonparametric, and resulting inferences are valid even if one of the the nuisance parameters is not estimated well. However, it can be advantageous to also consider an approach based on parametric modeling of the nuisances in consideration of various practical issues. In particular, nonparametric estimators for the survival function can be computationally intensive and have long run times when the sample size is large; they can also be computationally unstable when the number of endpoints is small, and they can have higher variance than more parametric methods. In view of these considerations, we consider application of both parametric and nonparametric approaches for nuisance estimation.

Parametric modeling approach: We estimate the conditional survival functions η_1 and η_2 using a Cox proportional-hazards model in conjunction with the Breslow estimator for the baseline hazard. Our Cox regression models include linear terms for each covariate, and no interaction

or higher order terms (e.g., splines) unless otherwise specified. In all analyses of groups that were randomized to receive the bivalent or monovalent booster, π is estimated using the sample mean. This is valid because the treatment assignment did not depend on covariates. For other comparisons, we estimate π using logistic regression, and similarly, and we only include linear terms unless otherwise specified.

Nonparametric modeling approach: The conditional probabilites of survival and censoring are estimated using a super learner. For the survival functions η_1 and η_2, the super learner library consists of the following estimators: (1) Kaplan-Meier estimator; (2) Cox proportional hazard estimator; (3) generalized additive models; (4) survival random forests. In all comparisons of groups that were randomized to receive the bivalent or monovalent booster, π is estimated using the sample mean. Otherwise, the propensity score is estimated using a super learner, with library: (1) the sample average of A – i.e., an intercept-only model; (2) logistic regression model; (3) generalized additive models; (4) random forests. Cross-fitting is performed in order to avoid assumptions that would disallow the use of flexible estimators that do not satisfy complexity (Donsker) conditions.

Due to the low rate of severe COVID-19, no covariate adjustment will be performed for analyses using the severe COVID-19 endpoint.

11.3.1 Primary and Secondary Analysis

In what follows, we describe our methods for covariate adjustment for the cumulative incidence approach for each of the primary and secondary analyses in Table 2. Covariate adjustment is only to be used for non-severe COVID-19 endpoints as the anticipated number of severe COVID-19 endpoints is low, and this could lead to instability of the covariate-adjusted estimates.

Parametric modeling
For all analyses comparing the monovalent and bivalent arms, we use the following variables to produce the Cox proportional hazards fit of η_1 and η_2: recent SARS-CoV-2 infection status at Month 6 and study group.

Nonparametric modeling
For all analyses comparing the monovalent and bivalent arms, we use the following variables to produce the Cox proportional hazards fit of η_1 and η_2: recent SARS-CoV-2 infection status at Month 6 and study group.

Main analysis vs sensitivity analysis
Of the strategies described above, the parametric modeling approach is treated as the main analysis strategy, and the nonparametric strategy is treated as a method sensitivity analysis.

11.3.2 Exploratory Analyses

We only use the cumulative incidence approach to address Exploratory Analyses I and II in Table 3. Other exploratory analyses (III through VII) are conducted only using the Cox modeling strategy, which is described fully in Section 11.4.3.

We now describe our covariate adjustment strategies for the cumulative incidence approach for the exploratory analyses I and II. While we considered two different covariate adjustment strategies for

the primary and secondary analyses, for the exploratory analyses, we only consider the parametric modeling strategy with binary covariates. Moreover, the same set of covariates are used as in the primary and secondary analysis.

11.4 Cox modeling approach

Advantages of the Cox modeling approach include (1) automatically includes all follow-up without somewhat tricky choices about the last time point t_0 for cumulative failure time analyses; (2) stable modeling approach in the presence of small numbers of events; (3) we have good *a priori* understanding of how this method should behave; in addition for the calendar-time scale analyses (4) it advantageously accommodates scenarios where the baseline hazard depends strongly and unpredictably on calendar time. Its disadvantages are the usual ones with a strong parametric modeling assumption, including that quantifying uncertainty with a 95% confidence interval about the conditional hazard ratio communicates more precision than we actually have. PH assumptions will be assessed using the Schoenfeld residuals (Grambsch and Therneau, 1994).

11.4.1 Primary analyses

Primary analysis I compares the COVE and CDC-based COVID-19 endpoints among participants randomized to a monovalent or bivalent M6 booster (RM6), counting event starting 13 days post M6 vaccination. To achieve this within a calendar-time-scale Cox regression framework, we consider the following time origin and associated analysis.

Let τ_0 denote the calendar time when the first RM6 participant completed the booster dose. For participant i, the time origin equals the number of days between 13 days after the participant's booster dose and τ_0. At any calendar time t, the risk set consists of RM6 participants who are at least 14 days post the M6 booster. We consider the following Cox regression model:

$$\lambda_{ji}(t \mid A_i, W_i) = \lambda_{0j}(t) \exp\left\{\beta A_i + \gamma^T W_i\right\} I(t > \tau_i^{\text{D14}}), \tag{4}$$

where $\lambda_{0j}(t)$ is a stratum-j-specific baseline hazard, τ_i^{D14} is Day 14 post M6 booster, $A_i = 1$ if participant i received a bivalent booster and $A_i = 0$ if a monovalent booster. When analyzing data under model (4), we will report $\exp(\beta)$ which measures the hazard ratio conditional on baseline covariates W when comparing bivalent to monovalent M6 booster.

The Cox regression will be conducted using the following stratification and covariate adjustment strategies:

- Stratification: the analysis will be stratified by 4 study groups and the geographic region (RSA or not). Therefore, a total of $4 \times 2 = 8$ strata will be considered.
- Covariate adjustment: the analysis will adjust for sex assigned at birth, history of TB, and a binary indicator of having had an infection during Part A of the study as defined in Section 7.1. The history of TB will not be adjusted in analysis restricting to PLWoH because of the small sample size.

Primary analysis II will be based on the time origin and stratification/covariate adjustment strategies as described above but restricted to the full per-protocol participants (FPP).

Primary analyses III and IV, when applicable, will be based on the same time origin as described above; however, no stratification or covariate adjustment will be conducted because of the expected small number of endpoints.

11.4.2 Secondary comparisons

Secondary analyses I and II counts event starting 1 day after the M6 vaccination. To this end, the following Cox regression model will be considered:

$$\lambda_{ji}(t \mid A_i, W_i) = \lambda_{0j}(t) \exp\left\{\beta A_i + \gamma^T W_i\right\} I(t > \tau_i^{\text{M6}}), \tag{5}$$

where τ_i^{M6} is M6 booster time of a participant i and the other quantities are defined as those described in Section 11.4.1. We will report the parameter $\exp(\beta)$ which measures the hazard ratio conditional on covariates W.

The same stratification and covariate adjustment strategies as described in Section 11.4.1 will be used in secondary analysis I. Secondary analysis II will not be stratified or adjust for baseline covariates because of the small expected number of endpoints.

Secondary analysis III will be subgroup analysis restricted to RM6 participants who live with HIV (PLHW) or without HIV (PLWoH), counting events 13 days post M6 booster. Analysis will be based on the same Cox regression model (4) and use the same stratification strategy. In addition to the sex, TB history, and an indicator of infection during Part A, analysis restricted to PLWH will further adjust for an indicator of CD4+ cell counts (≤ 500 or not) and an indicator of viral load (detected or not).

Secondary analysis IV is analogous to the secondary analysis III. Secondary analysis IV will be based on the Cox regression model (5) and adopt the same stratification and covariate adjustment strategies as for the secondary analysis III.

Finally, secondary analysis V is analogous to the secondary analysis III except that the analysis cohort will be the FPP (as opposed to RM6) PLWH/PLoW.

11.4.3 Exploratory comparisons

Exploratory analysis I is analogous to primary analysis I and secondary analysis I except that the analysis cohort will be the full M6 cohort (FM6). The same stratification and covariate adjustment strategies used in secondary analysis I will be adopted here.

Exploratory analysis II is analogous to the secondary analysis III except that the analysis cohort will be the FM6 PLWH/PLWoH. The same stratification and covariate adjustment strategies used in secondary analysis III will be adopted here.

Exploratory analysis III seeks to compare the COVID-19 incidence among participants in Analysis Group 1 vs. Analysis Group 2-1 (defined at baseline/enrollment) counting events from enrollment. To this end, the following Cox regression model is considered:

$$\lambda_{ji}(t \mid A_i, W_i) = \lambda_{0j}(t) \exp\left\{\beta A_i + \beta_1 Z_i(t) + \beta_2 A_i Z_i(t) + \gamma^T W_i\right\} I(t > \tau_i^{\text{enroll}}), \tag{6}$$

with $A_i = 1$ if participant i is in AG 2-1 and 0 if in AG 1, and $Z_i(t)$ is a time-varying indicator of receiving a M6 booster vaccine. The analysis will consider the same stratification and covariate adjustment strategy as specified in the pre-M6 SAP. Key parameters of interest include $\exp\{\beta\}$, $\exp\{\beta_1\}$ and $\exp\{\beta_2\}$.

Exploratory analysis IV and V are analogous to exploratory analysis III and will adopt the same stratification and covariate adjustment strategies.

Exploratory analysis VI and VII are analogous to the hybrid vs. vaccine-induced comparison as specified in the pre-M6 SAP, except that the analysis will define the hybrid vs. vaccine immunity status based on M6 and count events 13 days post M6 vaccination. The analyses will adopt the same stratification and covariate adjustment strategies as its pre-M6 counterpart.

12 Multiplicity adjustment

For all analyses, point estimates, 95% CIs and nominal p-values will be displayed in forest plots without multiplicity adjustment. No multiple hypothesis testing will be done in any described analyses.

13 Handling missing data

A small number (about 5%) of PLWH participants are missing CD4 cell count and/or HIV-1 viral load at enrollment or at Month 6. In order to adjust for CD4 cell count and HIV VL in various analyses and retain these participants, hard-imputation is used to assign a fixed CD4 cell count value and VL value for each of the participants with missing data. Specifically, for baseline values, within each study group (Group 1 or 2), HIV anti-retroviral therapy (ART) status (Yes, No or Missing) and Sex (Female or Male), missing baseline CD4 cell count is imputed to be the median CD4 count of those with observed baseline data. The same imputation is done for missing baseline VL. For Month 6 values, missing data will be imputed as the baseline observed or imputed value for CD4 cell count and VL. Missing data of history of TB at Month 6 is defined as negative.

References

Grambsch, P. M. and Therneau, T. M. (1994), "Proportional hazards tests and diagnostics based on weighted residuals," *Biometrika*, 81, 515–526.

Westling, T., Luedtke, A., Gilbert, P. B., and Carone, M. (2023), "Inference for treatment-specific survival curves using machine learning," *Journal of the American Statistical Association*, 1–26.

Signature Page for SCHARP-TMF-118029 v2.0

Reason for signing: Approved	Name: Bo Zhang Role: I am an author of the document. Date of signature: 01-Apr-2024 19:29:49 GMT+0000

Reason for signing: Approved	Name: Aaron Hudson Role: I am an author of the document. Date of signature: 01-Apr-2024 19:56:27 GMT+0000

Reason for signing: Approved	Name: Peter Gilbert Role: I reviewed and approved the document and am signing in lieu of or on behalf of another individual. Date of signature: 01-Apr-2024 21:23:53 GMT+0000

Reason for signing: Approved	Name: Yunda Huang Role: I am an author of the document. Date of signature: 02-Apr-2024 00:55:44 GMT+0000

Signature Page for SCHARP-TMF-118029 v2.0