Statistical analysis plan (SAP) for the Dupilumab Study

Version 6

March 29, 2022

1.0 Introduction

1.1 Study objectives: The primary objective is to determine the efficacy of dupilumab use plus standard of care management in patients hospitalized with moderate to severe COVID-19 infection compared to placebo plus standard of care management assessed by the proportion of patients alive and free of invasive mechanical ventilation at 28 days. The secondary objectives are to: (1) evaluate the safety of Dupilumab use in terms of eosinophilia in patients hospitalized with moderate to severe COVID-19 infection, (2) compare clinical outcomes on Dupilumab versus placebo across a range of secondary clinical safety and efficacy domains, and (3) evaluate the immunologic and biologic end points of inhibition of type 2 inflammation.

1.2 Study design: This is a randomized, double-blind, placebo-controlled, superiority phase IIa trial to assess the safety and efficacy of dupilumab use in hospitalized patients with moderate to severe COVID-19 infection. A total of 40 eligible subject will be enrolled and randomized in a 1:1 ratio to receive either dupilumab or placebo, stratifying on the disease severity measured by the required oxygen ≤ 15L or > 15L by nasal cannula. Both arms will receive standard of care management per current National Institutes of Health (NIH) COVID-19 treatment guidelines (https://covid19treatmentguidelines.nih.gov/) in addition to their randomized treatments. Patients will be followed prospectively for up to 60 days after enrollment.

1.3 Database Sources: Completed forms and electronic data will be entered into the data management system in the REDCap platform. Only authorized individuals shall have access to electronic CRFs.

1.4 Randomization:

In this randomized, double-blind, placebo-controlled Phase IIa trial, a total of 40 eligible subjects will be enrolled and randomized in a 1:1 ratio to receive either dupilumab or placebo, stratifying on the disease severity, which is measured by the required oxygen ≤ 15L or > 15L by nasal cannula. Both arms will receive standard of care management per current National Institutes of Health (NIH) COVID-19 treatment guidelines (https://covid19treatmentguidelines.nih.gov/) in addition to their randomized treatments.

Since this is a small trial, simple randomization procedure may result by chance in the compositions of the study arms being markedly different with respect to factors that may affect the outcome measures in the trial, or markedly unequal numbers of participants may be recruited to each arm. Further, disease severity measured by the required oxygen ≤ 15L or > 15L by nasal cannula may impact on the outcomes differentially. Thus, to ensure the

balance of treatment allocation in disease severity, the blocked randomization with block size of 4 will be performed, stratifying based on disease severity: patients who are on 15L or less of oxygen by nasal cannula (severity group A) and those requiring more than 15L of oxygen by nasal cannula including any noninvasive ventilation measures (severity group B). Thus, separate randomization lists will be generated for mild and severe patients, using the Proc Plan procedure in SAS, and the complete randomization lists will be sent to UVa Pharmacy only to ensure the blindness for others.

Example list for mild disease group:

Disease Severity	Order	Treatment Arm	Subject Initial	SubjectID	Date of Randomization	Randomized By
Mild	101	A				
Mild	102	A				
Mild	103	B				
Mild	104	B				
Mild	105	B				
Mild	106	A				
Mild	107	A				
Mild	108	B				

1.5 Study population:

Adult patients (18 years of age or older) who are hospitalized with a positive RT-PCR for SARS-CoV-2 within the last 14 days, with illness duration within the last 14 days, and evidence of moderate to severe COVID-19 infection as defined by NIH COVID-19 Severity Categorization (see the Appendix of study protocol):

- Moderate illness: Individuals who show evidence of lower respiratory disease during clinical assessment or imaging and who have saturation of oxygen $SpO_2 \geq$ 94% on room air at sea level.
- Severe illness: Individuals who have SpO_2 <94% on room air at sea level, a ratio of arterial partial pressure of oxygen to fraction of inspired oxygen (PaO_2/FiO_2) <300 mm Hg, respiratory frequency >30 breaths/min, or lung infiltrates >50%.

Inclusion Criteria

- Male or female 18 years of age or older at the time of enrollment.
- Patients hospitalized with a positive RT-PCR for SARS-CoV-2 within the last 14 days, with illness duration within the last 14 days, and evidence of moderate to severe COVID-19 infection as defined by NIH COVID-19 Severity Categorization (8):
  -Moderate illness: Individuals who show evidence of lower respiratory disease during clinical assessment or imaging and who have saturation of oxygen SpO2≥ 94% on room air at sea level.
  -Severe illness: Individuals who have SpO_2 <94% on room air at sea level, a ratio of

arterial partial pressure of oxygen to fraction of inspired oxygen (PaO_2/FiO_2) <300 mm Hg, respiratory frequency >30 breaths/min, or lung infiltrates >50%.

- Patient and/or legally authorized representative is willing and able to provide written informed consent and comply with all protocol requirements.
- Patients with hematologic malignancies or solid tumors are eligible.
- Patients with autoimmune disorders are eligible.
- Patients with immunodeficiency and organ or stem cell transplant recipients are eligible.
- Patients with acute or chronic renal injury/failure are eligible.
- Patients with neutropenia/lymphopenia are eligible.
- Patients with elevated liver function tests are eligible.
- Women who are not taking contraception are eligible.
- Patients who are currently or have recently received steroids and/or remdesivir are eligible.
- Patient agrees to not participate in another clinical trial for the treatment of COVID-19 through end of study period.

Exclusion Criteria

- Patients who do not require inpatient admission for COVID-19 infection.
- Patients who require invasive mechanical ventilation at time of enrollment.
- A pre-existing condition or use of a medication that, in the opinion of the site investigator, may place the individual at a substantially increased risk due to study participation.
- Pregnancy or breast feeding (lactating women who agree to discard breast milk from day 1 until two weeks after the last study product is given are not excluded).
- Allergy to Dupilumab or its excipients.
- Current acute parasitic helminth infection or history of chronic parasitic infection.
- History of ocular scleritis, uveitis, keratitis or recent (<6 months) eye injury (chemical or traumatic), infection or vascular occlusion.
- Have received any live vaccine (that is, live attenuated) within 4 weeks before screening, or intend to receive a live vaccine (or live attenuated) during the study. *Note: Use of non-live (inactivated) vaccinations is allowed for all subjects*.

2.0 Outcomes

2.1 Primary outcome: The primary efficacy endpoint of this double-blind, placebo-controlled phase IIa trial is a binary indicator for whether a patient would be alive and free of invasive mechanical ventilation at 28 days. The outcome will be summarized as the proportion of patients alive and free of invasive mechanical ventilation at 28 days in those receiving dupilumab in addition to standard of care (treatment arm) compared to those receiving solely standard of care management (placebo arm).

2.2 Secondary outcomes:

2.2.1 **Eosinophilia as the safety outcome**: The safety endpoint of this study is a binary indicator for whether a patient would have eosinophilia, which is defined as an absolute eosinophil count > 0.6 k/µl at ≥ 1 measurement throughout the study period. The outcome will be summarized as the proportion of patients with eosinophilia for treatment arm and placebo arm respectively. Complete blood counts with differentials

and complete metabolic panels will be measured on Day 0, 2, 5, 7, 14, 28 and 60. Day 7, 28 and 60 optional, although recommended, if patient discharged within the time frame.

2.2.2 **Other secondary safety and efficacy outcomes**:

- Cumulative incidence (defined as number of new events divided by the total number of individuals in the population at risk for the time interval) of grade 3 and 4 adverse events, serious adverse events or death including: injection site reactions, eye/eyelid inflammation, conjunctivitis, oropharyngeal pain, insomnia, tooth ache, gastritis, arthralgia, bacterial pneumonia, herpes viral infection, hypereosinophilic syndrome, hypersensitivity reaction.
- Prevalence of B.1.1.7 and B.1.351 and other SARS-CoV-2 lineages in study cohort: Day 0.
- Plasma total Immunoglobulin E (IgE) levels: Day 0 and 14.
- Plasma inflammatory markers (C-reactive protein and ferritin): Day 0, 7 and 14. Day 7 optional if patient discharged.
- Plasma cytokine levels including TARC (CCL17), YKL40, eotaxin 3 (CCL26), IL-13, IL-4, Arg1, Hyaluronan, soluble ST2: Day 0, 2, 5, 7, 14, and 28. Day 7 and 28 optional if patient discharged.
- Change in PaO2/SaO2 to FiO2 ratio: Day 0, 2, 5, 7 (if patient remains inpatient throughout)
- All-cause mortality rate at 28 days.
- Hospital length of stay (LOS)
- ICU LOS
- Proportion of patients alive and free of invasive respiratory failure at 28 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 60 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 90 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 120 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 150 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 180 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 270 days.
- Proportion of patients alive and free of invasive mechanical ventilation at 360 days.
- National Institute of Allergy and Infectious Diseases (NIAID) 8-point ordinal scale (see Appendix): Day 0, 2, 5, 7, 14, 28, 60, 180 and 360.
- Need for vasopressors
- Need for renal replacement therapy
- Need for extracorporeal membrane oxygenation (ECMO)

3.0 General considerations

3.1 Analysis populations: Each primary and secondary efficacy outcome measure will be analyzed for the intent-to-treat (ITT) and evaluable population. The intent-to-treat principle is that participants will be analyzed in the groups to which they were randomized, regardless of whether they received the randomized study medication. Another analysis population of interest is the as-treated population. The as-treated population will consist of only participants who were inducted onto study medication, even one dose of drug. Safety outcome will be primarily analyzed for the as-treated population.

Major differences in the results for the ITT and as-treated populations, if any, will be further explored. While there is every intention to be complete in describing the analyses to be performed, it is not possible to anticipate every contingency, and some adjustments may be required to meet constraints posed by the structure of the data. Constraints such as non-linearity, non-normality, etc. may lead to different but more appropriate approaches to analysis.

3.2 Testing of Distribution assumptions: Empirical distributions of all variables will be visually inspected to detect outliers. The underlying proposed statistical methods for each analysis will be examined, primarily through inspection of graphical displays, standardized residuals, or influence diagnostics. Where appropriate and/or necessary, transformations will be utilized for continuous measures such as biomarker data, or categorized biomarkers will be analyzed.

3.3 Interim analysis: there is no formal scheduled interim analysis with respect to the primary efficacy outcome and thus the sample size estimation would not be altered. However, analysis for some selected biomarkers will be performed after the 14^{th} day of the 20th patient enrolled. Particularly, the proportion and rate of change of C- reactive protein and ferritin levels will be calculated from day 0 and day 14 time points to determine efficacy trends between the two treatment groups. Additionally, percentage of eosinophilia will be calculated between the two groups to determine safety trend. Data will be analyzed by a third party and remain blinded to the study investigators. We expect greater changes in C-reactive protein and ferritin for patients receiving Dupilumab than those receiving placebo and no significant differences in eosinophilia between the two groups.

3.4 Software: All analyses will be performed utilizing SAS® version 9.4. All statistical tests will be conducted at the 5% Type I error rate (two-sided). When multiple tests are conducted, the chance of finding a significant difference in one of the tests, when in fact no difference exists, is greater than the stated Type I error rate. The investigators are aware of the multiple testing issues and will interpret results with caution and use confidence intervals where possible.

4.0 Analysis of demographic and baseline data:

The demographic variables for this study include age, gender, race, ethnicity, and BMI, etc. The baseline clinical characteristics include disease severity, symptoms, comorbid conditions, days

from COVID-19 symptom onset, and baseline blood lab results, etc.

Descriptive statistics for baseline and demographic variables will be presented for patients randomized to each of the treatment arms and overall. Descriptive statistics will include N, mean, standard deviation, median, 25th and 75th percentiles, minimum and maximum for continuous variables and proportions and percentages for categorical variables. In general, randomization is expected to produce balance at baseline between the two arms of the trial, formal statistical comparisons of treatment groups with respect to baseline characteristics will not be necessary. However, since this is a relatively small study, the differences in baseline characteristics between the treatment arms will be evaluated exploratorily for better understanding the study population. The updated CONSORT statement no longer recommends formal testing of statistical significance of differences between baseline characteristics.

5.0 Efficacy analyses:

5.1 General analysis strategy: The difference between dupilumab and placebo arms in outcome measures will be primarily analyzed under ITT principle, regardless of patient's use of rescue medications, protocol violations, or investigational product discontinuation. The difference will be tested initially with Chi-square for categorical measures (e.g., the primary efficacy endpoint) and two-sample t-test or Wilcoxon rank sum test for continuous measures (e.g., biomarkers). Due to the relatively small sample size in the study, the adjusted treatment effect of dupilumab will be estimated exploratively in the generalized linear regression for the primary and secondary outcomes to account for the potential impact of patient characteristics (e.g., age, sex, race and co-morbidities). Data transformation will be considered for the outcomes with skewed distribution such as hospital length of stay (LOS) and ICU LOS. The trajectory of each immune biomarker over the follow-up time points will be plotted to characterize the change of immune responses over time, and difference in biomarker (or transformed biomarker to improve normal distribution, if necessary) trajectories between the two study arms will be assessed in the mixed effects model.

Since this is intended to be an early stage, proof of concept study for which the primary goal is to inform future trials rather than to perform formal hypothesis testing, no multiplicity adjustment is planned. In addition, machine learning methods such as random forests algorithm will be applied to fully evaluate the treatment effect and predictability of dupilumab and other important predictors on the outcomes non-parametrically.

5.2 Primary efficacy outcome

The primary efficacy endpoint of this study is a binary indicator for whether a patient would be alive and free of invasive mechanical ventilation at 28 days. The outcome will be summarized as the proportion of patients alive and free of invasive mechanical ventilation at 28 days and the difference in the proportions between the two study arms will be evaluated with Chi-square test. The odds ratio of being alive and free of invasive mechanical ventilation at 28 days for Dupilumab vs. placebo and its 95% confidence interval (CI) will be estimated in logistic regression, adjusting for potential baseline characteristics if necessary.

5.3 Secondary outcomes

5.3.1 **Eosinophilia as the safety outcome**: Eosinophilia is a binary indicator for a patient

to have an absolute eosinophil count > 0.6 k/μl at ≥ 1 measurement based on the blood counts with differentials and complete metabolic panels on Day 0, 2, 5, 7, 14, 28 and 60 (Data on Day 7, 28 and 60 would be optional). The outcome will be summarized as the proportion of patients with eosinophilia and the difference in the proportions between the two study arms will be evaluated with Chi-square test. The odds ratio of having eosinophilia for Dupilumab vs. placebo and its 95% CI will also be estimated in logistic regression, with adjustment for other covariates.

5.3.2 **Other secondary safety and efficacy outcomes**:

- Cumulative incidence rate of adverse events (injection site reactions, eye/eyelid inflammation, conjunctivitis, oropharyngeal pain, insomnia, tooth ache, gastritis, arthralgia, bacterial pneumonia, herpes viral infection, hypereosinophilic syndrome, hypersensitivity reaction, other grade 3 or 4 adverse events, serious adverse events or death) is defined as number of adverse events divided by the total number of individuals in the population at risk for the time interval. The rate ratio of adverse events for Dupilumab vs. placebo and its 95% CI will be analyzed in Poisson regression, in which the follow-up time interval will be considered as the offset. Adjusted analysis for other covariates will be performed, as necessary. Mortality at 60 days will be analyzed as the time to event response. Survival probabilities will be estimated by the Kaplan-Meier method, and the difference in the mortality risk between the two treatment groups will be evaluated by the log-rank test and further in Cox regression, adjusting for the baseline covariates. For the analysis of mortality over the first year, additional adjustment for dynamic mechanical ventilation effect will be considered by including being on ventilation as a time-vary covariate in the Cox regression.
- Plasma total Immunoglobulin E (IgE) levels on Day 0 and 14 will be considered as continuous responses, and difference between treatment arms will be evaluated by t-test or Wilcoxon test initially, and adjusted difference and its 95% CI will be estimated from linear regression. Further, changes of IgE from Day 0 to Day 14 will also be explored and the difference between the two study arms will be analyzed similarly.
- Plasma inflammatory markers (C-reactive protein and ferritin) on Day 0, 7 and 14 (Day 7 would be optional if patient discharged) will be analyzed as longitudinal responses in the linear mixed effects model with Dupilumab treatment as a fixed effect. Since this is a randomized study, significant effect of treatment on intercept is not expected. Thus, the longitudinal analysis will focus on the slope difference between the treatment arms via interaction of treatment and time in the model. Since there are only 3 measures for each subject, only random slopes will be specified as *a prior*. Model fitting indices will be examined carefully to ensure the validity of estimated parameters. If Day 7 data won't be available for all patients, these biomarkers between Day 0 to Day 14 will be analyzed as the IgE above.
- Plasma cytokine levels including TARC (CCL17), YKL40, eotaxin 3 (CCL26), IL-13, IL-4, Arg 1, Hyaluronan, soluble ST2 on Day 0, 2, 5, 7, 14, and 28 (Day 7 and 28 optional if patient discharged) will be analyzed as longitudinal responses

in the linear mixed effects model with Dupilumab treatment as a fixed effect, and with appropriate transformation of biomarkers if necessary. The trajectories of these biomarkers will be carefully evaluated, and the slope difference between the treatment arms will be evaluated by the interaction of treatment and time. Additional curvature in trajectories will also be explored and captured whenever possible in the linear mixed effects model.

- Change in PaO_2/SaO_2 to FiO_2 ratio on Day 0, 2, 5, 7 (if patient remains inpatient throughout) will be analyzed as longitudinal responses in the linear mixed effects model, similar methods to that for inflammatory biomarkers and cytokine biomarkers as described above.
- All-cause mortality at 28 days will be analyzed as the binary response, using the same methods as that for the primary efficacy outcome and eosinophilia.
- Hospital length of stay (LOS) and ICU LOS between treatment arms will be analyzed as continuous measures using t-test or Wilcoxon test initially, and adjusted difference and its 95% CI will be estimated from linear regression. Since the distributions of hospital LOS and ICU LOS are often skewed, the log-transformed LOS data will be considered in the analyses.
- The response of being alive and free respiratory failure at 28 days is a binary response and will be analyzed using the same methods as that for the primary efficacy outcome.
- The response of being alive and free invasive mechanical ventilation at 60 days is a binary response and will be analyzed using the same methods as that for the primary efficacy outcome.
- National Institute of Allergy and Infectious Diseases (NIAID) 8-point ordinal scale on Day 0, 2, 5, 7, 14, 28, 60, 180 and 360 will be considered as longitudinal data. The distributions of the data will be examined visually first and further collapsing will be considered. Collapsed data will be analyzed in the generalized linear mixed effect model or multinomial logistic regression with the generalized estimating equation approach.
- Proportion of patients alive and free of invasive mechanical ventilation at 60, 90, 120, 150, 180, 270 and 360 days will be estimated from the Kapan-Meier survival for the time to mechanical ventilation-free survival over the first year. The difference in these proportions between the two treatment groups will be evaluated by the log-rank test and further in Cox regression adjusting for the baseline covariates. For the analysis of mortality over the first year, additional adjustment for dynamic mechanical ventilation effect will be considered by including being on ventilation as a time-vary covariate in the Cox regression.
- Need for vasopressors is a binary response and will be analyzed using the same methods as for the primary efficacy outcome.
- Need for renal replacement therapy is a binary response and will be analyzed using the same methods as for the primary efficacy outcome.
- Need for extracorporeal membrane oxygenation (ECMO) is a binary response and will be analyzed using the same methods as for the primary efficacy outcome.

6.0 Other safety analyses:

Analysis of adverse event (AE) data will primarily be descriptive based on CTCAE coding of events. The proportion of subjects experiencing a severe AE and the proportion experiencing a Grade 3 or higher will be recorded. AE will be compared to published data.

7.0 Sample size estimation and power analysis:

This trial is intended to be an early stage and proof of concept study with the primary goal to collect reliable data and obtain sufficient evidence for future trials. Thus, the emphasis of the study is an exploratory estimation of effect size and incidence of adverse events in terms of eosinophilia; it is not considered as a pivotal trial for efficacy at this time. Therefore, a larger significance level (one-sided alpha=0.1), and the sample size is pre-selected based on the feasibility. Given the current decline trend in the incidence of COVID-19 positive cases, we anticipate to enroll 40 patients realistically into this randomized clinical trial, which shall be sufficient for us to collect reliable and interpretable data for proof of concept to inform future trials. The COVID-19 hospitalization data from our University of Virginia Health System between March 2020 and April 6, 2021 showed that 79.5% of COVID-19 inpatients were alive and free of mechanical ventilation at 28 days under the usual care. With 40 patients that are deemed to be feasible, we would expect Dupilumab to improve a margin of 17.7% in the proportion of patients alive and free of mechanical ventilation at 28 days. That is, it is anticipated that 97.2% of patients being alive and free of ventilation at 28 days in the Dupilumab arm. The power analysis is conducted using the one-sided Score test (Farrington & Manning) with 75% power and 10% type I error.

8.0 Missing data

Since patients in this study will be closely followed for mortality and status on mechanical ventilation, the primary outcome of being alive and free of ventilation is likely captured. Dropout from the study will be carefully evaluated for its potential informativeness. A sensitivity analysis will be performed to examine the impact on the proportion and its confidence limits. In particular for the primary endpoint, individuals who withdraw will be considered events if their statuses on day 28 are unknown.

Other outcome variables (clinical conditions and laboratory biomarkers) will have missing data due to dropout from treatment and/or from study participation, or due to missed outpatient visits if patients have been discharged from the hospital. The generalized linear model, or mixed effects model frameworks that will be used for analyses, works with what data are gathered, and assumes missing data are missing at random. For selected secondary outcome analyses, sensitivity analyses will be considered to examine the stability of estimated treatment effects in the face of departures from the assumption of missing at random.