STATISTICAL ANALYSIS PLAN

for

ACTIV-4B: COVID-19 Outpatient Thrombosis Prevention Trial

A multicenter adaptive randomized placebo-controlled platform trial evaluating the efficacy and safety of antithrombotic strategies in COVID-19 adults not requiring hospitalization at time of diagnosis

May 31, 2021

1 PRIMARY AND SECONDARY AIMS OF THE TRIAL

The **primary aim** of the trial is to compare the effects of treatment with (i) anticoagulation at prophylactic doses; (ii) anticoagulation at therapeutic doses; (iii) antiplatelet therapy; and (iv) placebo relative to each other on the primary composite endpoint of symptomatic deep venous thrombosis, pulmonary embolism, arterial thromboembolism, myocardial infarction, ischemic stroke, need for hospitalization for cardiovascular/pulmonary events, and all-cause mortality for up to 45 days after initiation of assigned treatment among COVID-19 patients not requiring hospitalization at time of diagnosis who are aged $\geq$ 40 years and < 80 years.

The **secondary aims** of the trial are:

1. to compare the effects of treatment with (i) anticoagulation at prophylactic doses; (ii) anticoagulation at therapeutic doses; (iii) antiplatelet therapy; and (iv) placebo relative to each other on the following secondary endpoints up to 45 days after initiation of assigned treatment among COVID-19 patients not requiring hospitalization at time of diagnosis who are aged $\geq$ 40 years and < 80 years:
   - need for hospitalization for cardiovascular/pulmonary events
   - venous thromboembolism including symptomatic DVT and PE.
   - arterial thrombotic events including MI, ischemic stroke, and arterial thromboembolism.
   - all-cause mortality.
   - mortality without antecedent hospitalization.
   - the time-to-event for the primary composite endpoint of symptomatic deep venous thrombosis, pulmonary embolism, arterial thromboembolism, myocardial infarction, ischemic stroke, need for hospitalization for cardiovascular/pulmonary events, and all-cause mortality
   - a clinical rank-based score.

2. to compare the effects of treatment with (i) combined prophylactic and therapeutic doses of apixaban with (ii) placebo for the primary endpoints for efficacy and for safety**.**

3. to test whether D-dimer and/or hsCRP modify the treatment effect of assigned treatment on the trial primary and secondary outcomes.

The **safety aims** of the trial are to compare the effects of treatment with (i) anticoagulation at prophylactic doses; (ii) anticoagulation at therapeutic doses; (iii) antiplatelet therapy; and (iv) placebo relative to each other on bleeding outcomes for up to 45 days after initiation of assigned treatment and after an additional 30 days of safety follow up (day 75) among COVID-19 patients not requiring hospitalization at time of diagnosis who are aged $\geq$ 40 years and < 80 years.

- ISTH major bleeding
- ISTH clinically relevant non-major bleeding (CRNMB).
- development of disseminated intravascular coagulation (DIC)

2. STUDY DESIGN

2.1 POPULATION

The trial eligibility criteria for randomization are listed below.

Inclusion Criteria

- Age between 40 and 80 years inclusive
- Documentation of PCR or antigen test positive symptomatic COVID-19 infection in the past 14 days
- Ability to be contacted by telephone or other electronic methods of communication
- Negative pregnancy test for women of child bearing potential

Exclusion Criteria

- Indication for therapeutic anticoagulation (mechanical heart valve, AF, APS)
- Indication for single or dual antiplatelet therapy
- Lactating
- Primary brain tumor or acute leukemia
- Bleeding risk defined as hospitalization in the past 2 months for:
    - bleeding due to ulcer or GI tract disease
    - major surgery
    - stroke
    - intracranial hemorrhage
- Platelet count < 100,000 per microliter (can be obtained after randomization)
- Calculated creatinine clearance < 30 ml/min (can be obtained after randomization)
- Ever hospitalized after diagnosis of COVID-19
- Concomitant need for strong inducers/inhibitors of p-gp and CYP3A4
- SARS-CoV-2 PCR or antigen test more than 14 days prior to randomization
- Unable to give written informed consent

2.2 INTERVENTIONS

Assigned Intervention Groups: Participants will be randomized at a 1:1:1:1 ratio to the four treatment groups using a permuted block design.

Group	Treatment	Dose AM	Dose PM	Duration
1.	Apixaban	2.5 mg	2.5 mg	45 days
2.	Apixaban	5.0 mg	5.0 mg	45 days
3.	Aspirin	81 mg	Placebo	45 days
4.	Placebo	Placebo	Placebo	45 days

For randomized participants, treatment duration will be 45 days unless a primary, secondary, or safety outcome occurs before 45 days in which case treatment may be stopped for clinical reasons.

2.3 OUTCOMES AND TIMING

The **primary endpoint** is the binary (yes/no) composite efficacy endpoint indicating that any of the following events occurred: symptomatic deep venous thrombosis, pulmonary embolism, arterial thromboembolism, myocardial infarction, ischemic stroke, need for hospitalization for cardiovascular/pulmonary events, and all-cause mortality.

Primary treatment comparisons will be conducted in the sample of **randomized participants who initiate treatment and have at least one follow-up contact.** A follow-up visit includes a contact where patient-reported outcomes or site-reported outcomes about patient status are collected. In this modified intention to treat (mITT) sample of patients, endpoint events will be tabulated from initiation of assigned therapy through 45 days after treatment initiation. Additional analyses will be conducted in the sample of **all randomized participants**. In the complete randomized sample, follow-up will begin at the time of randomization.

A **key secondary endpoint** is the Kaplan Meier time-to-event estimate of the cumulative risk of the primary composite endpoint. For treatment comparisons among randomized participants who initiate treatment, the cumulative risk 45-days after initiation of assigned therapy will be estimated, and for analyses among all randomized participants, the cumulative risk 45-days after randomization will be estimated.

The **secondary endpoints** are:

- Hospitalization for cardiovascular/pulmonary events
- Venous thromboembolism, a composite of symptomatic DVT and PE.
- Symptomatic DVT
- Pulmonary embolism
- Arterial thrombotic events, a composite of MI, ischemic stroke and arterial embolism
- Myocardial infarction
- Ischemic stroke
- Arterial thromboembolism
- All-cause mortality
- Mortality without antecedent hospitalization

Timing for the secondary endpoints is the same as what is described for the primary composite efficacy endpoint.

An **exploratory tertiary endpoint** is a clinical rank-based score. This clinical rank-based score is defined as the worst category accomplished during the 45-day treatment period (i.e. starting at treatment initiation) using the numeric rankings from best (score=1) to worst (score=9):

1. No clinical event (i.e. no study endpoint, safety endpoint or urgent/emergent health care encounter). A minor bleed that does not involve seeking medical attention is not a trial safety endpoint and hence is counted in this category.
2. Non-fatal bleeding that requires medical attention but not a hospital admission

3. Non-fatal event that is one of the composite primary events that requires an urgent care center visit or emergency room visit but not a hospital admission (e.g. this includes a DVT or pulmonary embolism that do not result in a hospital admission)
4. Non-fatal hospitalization for bleeding event or cardiovascular/pulmonary event not including stroke, MI, pulmonary embolism or DVT.
5. Non-fatal hospitalization for DVT
6. Non-fatal hospitalization for PE
7. Non-fatal hospitalization for MI
8. Non-fatal hospitalization for stroke
9. Death

The **safety endpoints** are:

- Major bleeding (ISTH major bleeding)
    - Drop in hemoglobin of 2 gm/dl attributed to bleeding and
    - Requiring transfusion of 2 or more units
    - Bleeding in a critical site which includes hemorrhagic stroke and intracranial hemorrhage
    - Fatal bleeding
- Mild bleeding (ISTH CRNMB): Non-major clinically relevant bleeding is defined as overt bleeding not meeting the criteria for major bleeding but associated with medical intervention, unscheduled contact (visit or telephone call) with a physician, (temporary) cessation of study intervention, or associated with discomfort for the participant such as pain or impairment of activities of daily life.
- Development of disseminated intravascular coagulation (DIC)

For the treatment comparisons among randomized participants who initiate treatment and have at least one follow-up contact, safety endpoints will be tabulated from initiation of assigned therapy through 45 days after treatment initiation and through 75 days after treatment initiation. For analyses involving all randomized participants, safety outcomes will be tabulated from time of randomization through 45 days after randomization and up to 75 days after randomization.

2.4 POWER AND SAMPLE SIZE

We determined the samples sizes required to provide 80% and 90% power to detect a relative reduction of 33% in the 45-day primary outcome event rates between two assigned treatment groups using chi-square statistic with one-sided test with alpha=0.025. Based on these estimates, we proposed a total sample of N=7000 patients with N=1750 patients assigned to each of the four treatment arms. Assuming a placebo event rate of 8.0%, a trial with N=1750 patients in each arm will have 80% power to detect superiority of apixaban 5.0 mg to placebo when there is a 30% relative reduction in risk (i.e. 8.0% vs. 5.62%) and 90% power with a 34% relative reduction (i.e. 8.0% vs. 5.28%). Assuming an event rate of 6.0% with aspirin, a trial with N=1750 patients in each arm will have 80% power to detect superiority of apixaban 5.0 mg to aspirin when there is a 34% relative reduction in risk (i.e. 6.0% vs. 3.94%) and 90% power with a 39% relative reduction (i.e. 6.0% vs. 3.65%).

Since we hypothesize that the active treatments will be beneficial, we estimated that the overall primary composite efficacy endpoint event risk in the trial (i.e. all treatment groups combined) will be this population is approximately 7.0%. We also estimated the overall bleeding event rate will be approximately 1.0%. With a total of N=7000 patients, we therefore assumed that we will observe approximately 490 patients with primary endpoint events and 70 with bleeding events.

2.5 ENDPOINT ADJUDICATION

All events suggestive of the primary composite endpoint will be adjudicated by a central independent Clinical Endpoints Committee (CEC). All of the suspected endpoint events will be classified by the CEC so that each of the defined secondary endpoint events will be adjudicated. All suspected bleeding events deemed to be clinically relevant non-major bleeding (CRNMB), major bleeding or DIC will be adjudicated by the CEC. The specific endpoints are each defined in the ACTIV-4B Clinical Endpoints Committee Charter.

The options available to the adjudicators include the ability to confirm an event (if there is sufficient clinical information to support the endpoint), disconfirm the event (if there is sufficient clinical information to disconfirm the endpoint), or mark the event as having "insufficient evidence" for confirmation (or for disconfirmation).

2.6 BLINDING

The ACTIV-4B Outpatient trial is double blinded. All study participants, clinical investigators, staff who collect data, medical monitors who classify adverse events, and Clinical Endpoints Committee members who adjudicate study endpoints are blinded to treatment assignment. Only the central data management and the unblinded statistical team have access to the treatment assignments. Protocols have been developed so that treatment assignment may be revealed in cases of clinical emergencies.

2.7 ANALYSIS POPULATIONS

Modified Intention to Treat (mITT) Population: The primary treatment comparisons will be based on the modified intention to treat (mITT) principle. The mITT analyses will include only **randomized participants who initiate their assigned treatment regimen and for whom there is at least one follow-up visit**. A follow-up visit includes a contact where patient-reported outcomes or site-reported outcomes about patient status are collected. Analyses in this sample will be conducted based on the randomly assigned treatment starting at the time of treatment initiation.

All Randomized Patients: The population of all randomized participants will be evaluated from time of randomization onward. The combined population will be used to estimate the overall risk of events in this patient cohort. Secondary treatment comparisons will be conducted using the intention to treat (ITT) principle based on the randomly assigned treatment group starting at the time of randomization.

Per Protocol Population: The group of randomized participants who report taking ≥70% of their pills per week for ≥5 weeks or until the time of a primary outcome or safety event occurred

will be considered adherent to the trial treatment regimen and will be included in the per protocol sample. Analysis of this per protocol group will begin at the time of treatment initiation. mITT participants who adhered to their assigned treatment and have complete 45-day follow-up or complete follow-up up to the time of a hospitalization or a death will be included in the analysis of the per protocol group.

2.8 HANDLING MISSING DATA

Missing Outcome Data: The primary endpoint for the ACTIV-4B Outpatient trial is a binary (yes/no) outcome indicating whether any of the listed events occurred within 45 days of treatment initiation. Hence, our primary analysis will be conducted in the sample for whom 45-day outcome data is available (45 day follow-up completed or a death). We expect that the percentage of participants with missing 45-day outcome data due to withdrawal or loss to follow-up will be small (<5%) and that the probability of missing data will be similar across the four treatment arms and will be weakly associated with the missing endpoint.

We will compare those participants with and without missing endpoint data by treatment group and by baseline demographic features. We will also present the likelihood that missing data would change the conclusions about the treatment effects using a tipping point analyses. By systematically and comprehensively varying assumptions about the missing outcomes in the four treatment arms, we will explore the whether the conclusions change. We will allow assumptions about the missing outcomes in the four treatment arms to vary independently, including scenarios where dropouts on active drugs tend to have worse outcomes than dropouts on control. This approach is consistent with recent FDA guidelines (E9(R1)-Statistical-Principles-for-Clinical-Trials attached).

The secondary endpoint, the Kaplan-Meier estimate of the cumulative risk of a primary endpoint at 45 days, appropriately accounts for variable follow-up time under the assumption of non-informative censoring. Analysis of this time to event outcome among all randomized or mITT participants (i.e. those with and without missing 45-day data) will provide further insight about the robustness of the trial conclusions based on the primary composite efficacy endpoint.

Missing Adjudication Data: Each suspected specified efficacy and bleeding event is adjudicated by the CEC based on medical records. In the rare case when medical records cannot be obtained, the adjudicators will review all available information including narratives from the local principal investigator or members of the study team who reported the event. The adjudicators will use all available information to classify the event as confirmed, disconfirmed or insufficient evidence. The primary analyses will be based on confirmed events. A secondary sensitivity analysis will be based on events that are confirmed and those that have insufficient evidence.

Missing Covariate Data: Data will not be removed from the primary analyses due to missing covariate data. Variables that have <10% missing data will be imputed using single imputation. Categorical variables that have > 10% missing data will include a category for missing data. Continuous variables that have > 10% missing data will be imputed using single imputation and a missing indicator variable will be added to the model.

3. STATISTICAL ANALYSIS PLAN

3.1 BASELINE DESCRIPTIVE STATISTICS

The distribution of demographic, clinical history, medications and biomarker variables will be examined and transformations will be applied as needed. Baseline characteristics will be examined for all randomized participants, for the entire mITT sample, and by assigned treatment group within the mITT group. Variables will be summarized using mean, standard deviation or median (first and third quartile) for continuous variables and frequency (percentage) for categorical variables. No test of significance levels will be reported for baseline variables.

3.2 RETENTION ANALYSES

The proportion of mITT participants who withdraw or are lost to follow-up before 45-days of follow-up will be tabulated overall and by treatment group. Baseline characteristics of patients with missing primary outcome data will be compared to those with complete data.

3.3 ADHERENCE ANALYSES

The proportion of mITT participants overall and in each assigned treatment group who have interrupted treatment permanently or temporarily and the reason for interruption will be described. We will present the proportion of participants who took ≥70% of their pills as prescribed for ≥5 weeks or until a clinical event occurred in the overall mITT sample and stratified by assigned treatment group.

3.4 PRIMARY OUTCOME ANALYSIS

The primary analyses will be conducted in the mITT sample based on the randomly assigned treatment starting at the time of treatment initiation. Participants who complete at least one follow-up visit after starting their assigned drug treatment will be included in the analysis.

The odds of the primary composite efficacy endpoint in the mITT sample will be modeled using a logistic regression model defined as:

$$\text{Log}\,(p/1\text{-}p) = \beta_0 + \beta_1\text{Apixiban5.0} + \beta_2\text{Apixiban2.5} + \beta_3\text{Asipirin} + \beta_4\text{Non-US} + \beta_5\text{age} + \beta_6\text{female} + \beta_7\text{BlackNH} + \beta_8\text{Hispanic} + \beta_9\text{OtherRE} + \beta_{10}\text{D-Dimer} + \beta_{11}\text{HsCRP} + \beta_{12}\text{Weight} + \beta_{13}\text{CrClearance}$$

Race/ethnicity will be defined as white non-Hispanic, black non-Hispanic, Hispanic, and Other/unknown race/ethnicity. White non-Hispanic race/ethnicity will serve as the references group.

The placebo treatment group will serve as the “reference” treatment group in this model, and we will test whether the coefficient for each active treatment group relative to the reference placebo group is equal to 0 using a two-sided test with alpha=0.05. Other pairwise treatment

comparisons (apixaban 5.0 versus apixaban 2.5, apixaban 5.0 versus aspirin, apixaban 2.5 versus aspirin) will be conducted, and the effect of treatment with apixaban (i.e. the combined group including both apixaban 5.0 and apixaban 2.5) will be compared with placebo.

If the number of mITT participants with a primary composite efficacy endpoint event in the 45-days after treatment initiation is low, a logistic regression model with a reduced number of covariates must be used in order to have adequate degrees of freedom for valid estimation. Below are the planned models to be used for the primary treatment comparison under the scenarios that the number of patients with primary endpoint events is <30, 30-49, and ≥50. The same treatment contrasts will be computed in the reduced models as described for the full model.

If the number of mITT participants with a primary endpoint event is <30, an unadjusted logistic regression model will be used as the primary model to assess the effect of assigned treatment.

$$\text{Log } (p/1\text{-}p) = \beta_0 + \beta_1 \text{Apixiban5.0} + \beta_2 \text{Apixiban2.5} + \beta_3 \text{Asipirin}$$

If the number of mITT participants with a primary endpoint event is 30-49, a logistic regression model adjusting only for age and D-dimer level will be used as the primary model to assess the effect of assigned treatment.

$$\text{Log } (p/1\text{-}p) = \beta_0 + \beta_1 \text{Apixiban5.0} + \beta_2 \text{Apixiban2.5} + \beta_3 \text{Asipirin} + \beta_4 \text{age} + \beta_5 \text{D-Dimer}$$

If the number of mITT participants with a primary endpoint event is ≥50, the full logistic regression model will be used as the primary model to assess the effect of assigned treatment.

$$\begin{aligned}\text{Log } (p/1\text{-}p) = {} & \beta_0 + \beta_1 \text{Apixiban5.0} + \beta_2 \text{Apixiban2.5} + \beta_3 \text{Asipirin} + \beta_4 \text{Non-US} + \beta_5 \text{age} + \beta_6 \text{female} \\ & + \beta_7 \text{BlackNH} + \beta_8 \text{Hispanic} + \beta_9 \text{OtherRE} + \beta_{10} \text{D-Dimer} + \beta_{11} \text{HsCRP} + \beta_{12} \text{Weight} \\ & + \beta_{13} \text{CrClearance}\end{aligned}$$

In addition to the primary logistic regression analysis, the unadjusted estimated risk of the primary composite efficacy endpoint in each treatment group (i.e. # of participants with an event / # of participants in the group), and the pairwise relative risks and absolute risk differences with 95% confidence intervals will be calculated and presented.

An ITT and a per protocol analysis will be conducted by running the multivariable adjusted logistic regression model and the unadjusted risk estimates using the corresponding sample and relevant exposure time.

3.5 KEY SECONDARY OUTCOME ANALYSES

Kaplan-Meier cumulative incidence curves will be created for the primary composite efficacy endpoint up to 45 days after treatment initiation stratified by assigned treatment group in the mITT sample. Log-rank statistics will be computed to compare the time to event estimates over time among the four treatment groups. The estimated cumulative risk at 45-days and the 95% confidence interval for the estimated cumulative risk at 45-days will be determined for each

treatment group. Pairwise differences and 95% confidence intervals for differences will be computed. The combined group of prophylactic and therapeutic doses of apixaban will be compared with placebo.

Kaplan-Meier cumulative incidence curves will be created for the primary composite efficacy endpoint up to 45 days after randomization for the complete randomized sample overall and stratified by assigned treatment group (ITT). Log-rank statistics will be computed to compare the time to event estimates over time among the four assigned treatment groups. The estimated cumulative risk at 45-days and the 95% confidence interval for the estimated cumulative risk at 45-days will be determined for the overall group and for each treatment group.

3.6 SECONDARY OUTCOME ANALYSES

For each defined secondary outcome event, the unadjusted risk of the endpoint in each treatment group (i.e. # of participants with the specified event / # of participants in the group) and pairwise relative risks and the absolute risk difference between treatment groups will be calculated with their 95% confidence intervals. In addition, the effect of treatment with apixaban (i.e. prophylactic and therapeutic groups combined) will be compared with placebo.

3.6 EXPLORATORY TERTIARY OUTCOME ANALYSES

The distribution across the 9 categories of the clinical rank-based score and the median and 25^{th} and 75^{th} percentile will be present each treatment group in the mITT sample. Kruskal-Wallace tests will be used to compare the distribution of the clinical rank-based score among the assigned treatment groups in the mITT sample. Pairwise comparisons with Wilcoxon rank sum statistics will be conducted to determine if one treatment has a “better” outcome relative to another.

3.7 SAFETY ANALYSES

The risk and 95% CI of each defined safety endpoint event and the composite of any defined safety endpoint event (major bleeding, CRNMB or DIC) at 45 and at 75 days will be computed for each treatment group in the mITT. Pairwise relative risks and absolute risk differences between treatment groups will be calculated with their 95% confidence intervals.

3.8 SUBGROUP ANALYSES AND EFFECT MODIFICATION

A select number of subgroup variables have been specified a priori:

- D-dimer (<1.0 ULN, [1.0-2.0 ULN), [2.0-3.0 ULN), ≥ 3.0 ULN, or by quartiles if needed)
- CRP by quartiles based on the data
- Age (<60 years, ≥60 years)
- Sex
- Race/ethnicity (White non-Hispanic, Black non-Hispanic, Hispanic, other)
- Renal function (creatinine clearance 30-49 ml/min, 50-90 ml/min, >91 ml/min)

The risk of the primary composite efficacy endpoint outcome and the 45-day risk of major bleeding with 95% confidence intervals will be estimated in each treatment group within each subgroup for the mITT sample. Evidence of effect modification of the treatment effectiveness by subgroup will be tested by creating a logistic regression model including the subgroup variable, treatment assignment, and the interaction between the subgroup variable and treatment assignment. If a subgroup variable is inherently continuous (i.e. D-dimer, CRP and age), these variables will be appropriately transformed as needed to approximate a normal distribution, and included in the model as a continuous variable. The significance of the interaction term will be presented. Additional subgroups may be examined in exploratory analyses based on observed results from the trial or information from external sources.

We will examine the distribution of baseline D-dimer in the entire mITT and in the randomized sample, and we will analyze the odds of the primary composite efficacy endpoint and major bleeding, irrespective of assigned treatment group by D-Dimer subgroup. Logistic regression models and associated ROC curves will be created for the primary efficacy endpoint and for major bleeding by continuous D-dimer level. LOESS curves for the logit of the primary efficacy endpoint and for major bleeding by continuous D-dimer level will also be examined.

3.9 ANALYSIS OF DURATION OF TREATMENT

Kaplan-Meier cumulative incidence curves will be created to assess the time to the first primary endpoint event and the time to the first safety event, irrespective of treatment assignment. Assuming that bleeding events occur at a fairly constant rate over time, we suggest that if ≥ 90% of the primary endpoint events occur in the first 21 days, then the DSMB will consider modifying the treatment arms such that the duration of therapy is shortened to 21 days. Curves stratified by treatment group will be examined before finalizing a recommendation.

4. INTERIM MONITORING PLAN FOR EFFICACY, FUTILITY AND SAFETY

4.1 OVERVIEW OF PLANNED INTERIM ANALYSIS

An independent data safety and monitoring board (DSMB) will review all interim analyses prepared by an unblinded statistician. The number of patients randomized, the number of randomized participants who initiated treatment and the primary, secondary and safety endpoints for the entire randomized and the entire mITT samples will be presented. Unadjusted risk of the defined primary, secondary and safety endpoints by assigned treatment group in the mITT sample will be examined on a monthly basis. For DSMB presentation, two versions of the primary, secondary and safety endpoint tables will be presented: one table will include the best information available which will include the adjudicated endpoints for events that have been classified by the CEC and self-reported endpoints for events that have not been classified by the CEC, and a second table that will include adjudicated endpoints only. A complete interim analysis of efficacy, futility, and safety will be conducted for each full DSMB review meeting which will occur approximately every 3 months.

A Bayesian analytic approach is proposed for the interim monitoring plan in order to utilize prior information when estimating the posterior probabilities in the sequential interim analyses.

Initially, the placebo group will serve as the “control group”; however, if the placebo arm is dropped and the trial continues, the aspirin arm will be designated as the control group for future treatment comparisons.

Decision rules have been established for efficacy based on the posterior probability that the active treatment regimen is beneficial as compared to placebo with respect to the primary composite efficacy endpoint. Decision rules were created such that the overall type I error approximates the pre-specified alpha=0.025 for a one-sided test. Assuming a non-informative prior distribution for each odds ratio at the first interim analysis, we will calculate the posterior probability that an active treatment is superior to placebo. We will update these posterior probabilities with new data at each subsequent interim analysis. If the posterior probability exceeds the pre-specified threshold for superiority at any of the interim analyses, the superior treatment will be declared efficacious and the other treatment may be dropped.

Decision rules have been developed for assessing futility of the active treatments. The posterior probability that each of the active treatments is inferior or equivalent to placebo with respect to the primary composite efficacy endpoint will be calculated assuming non-informative priors at the outset of the trial. When the posterior probability exceeds a specified threshold, futility will be established and the respective active therapy may be dropped from the trial.

Safety data will be presented and analyzed at each meeting, but no formal decision rules will be established a priori for the bleeding safety endpoints. Data will be presented so that the DSMB can evaluate the net risk benefit ratio for each treatment.

At each meeting, the DSMB will examine the rate of enrollment (and treatment initiation) in the trial as well as the overall risk of the primary endpoint. Based on this information, they may request a traditional futility analysis of conditional power to detect superiority. This involves a determination of the detectable risk ratio (or relative risk reduction) conditional on the observed data at that time under various assumptions regarding the future risk of endpoint events and the underlying treatment risk ratios.

4.2 FORMAL MONITORING OF SUPERIORITY BASED ON PRIMARY ENDPOINT

A logistic regression model will be created for the primary composite efficacy endpoint such that the effect of each active treatment group (relative to the placebo reference group) will be estimated adjusting for covariates (age, sex, race/ethnicity, D-dimer, and hsCRP, weight and calculated creatinine clearance) as specified in Section 3.4 based on the observed number of participants with events. The primary analyses for efficacy will be based on the odds ratios, comparing one treatment to another, derived from this model. One treatment is beneficial compared to another if the [Odds Ratio < 1.00] for the primary composite outcome. Assuming non-informative priors at the first look, we will calculate the posterior probabilities that the [Odds Ratio < 1.00] for each active treatment compared to placebo. If at any analysis time-point, the upper bound of the lower 99% credible interval for the odds ratio is less than 1.00, the active treatment arm will be considered superior.

The decision rule for declaring superiority based on the primary composite outcome is:

≥ 0.99 Posterior Probability that the OR (active vs placebo) < 1.00

The DSMB will use this information to make a recommendation to the NHLBI. The DSMB can recommend that the Outpatient COVID-19 trial should continue as proposed, that the control treatment arm should be dropped from the trial, that the trial protocol should be modified, or that the Outpatient COVID-19 trial should be terminated early. The final decision to stop trial rests with the NHLBI.

4.3 FORMAL MONITORING OF FUTILITY BASED ON PRIMARY ENDPOINT

Using the same logistic regression model that will be used for the primary analyses, we will determine the posterior probability that the active arm is equivalent or inferior to placebo adjusting for covariates (age, sex, race/ethnicity, D-dimer, and hsCRP, weight and calculated creatinine clearance) as specified in Section 3.4 based on the observed number of participants with events. Given that the trial was designed to have powered to detect a relative risk reduction of 33% with active treatment, futility will be defined for an active arm if the lower bound of the upper 95% credible interval for the odds ratio comparing the active arm to placebo is greater than 0.75.

The decision rule for declaring futility based on the primary composite outcome is:

$$\geq 0.95 \text{ Posterior Probability that the OR (active vs placebo)} > 0.75$$

This roughly corresponds to the having an estimated Odds Ratio that is 1.00 (or greater) and the two-sided 90% confidence interval extends from 0.75 to 1.33 (or greater).

The DSMB will use this information to determine its recommendation to NHLBI. The DSMB can recommend that the Outpatient COVID-19 trial should continue as proposed, that the futile active treatment arm should be dropped from the trial, that the trial protocol should be modified, or that the Outpatient COVID-19 trial should be terminated early. The NHLBI will make the final decision.

4.4 MONITORING SAFETY

A logistic regression model will be created for the major bleeding endpoint and for the composite safety endpoint (major bleeding, CRNMB and DIC) such that the effect of each active treatment group (relative to the placebo reference group) will be estimated and the odds ratios, comparing one treatment to another, will be derived from this model. We will not create explicit decision rules based on the bleeding posterior probability.

If safety issues arise, the DSMB will use their clinical judgement to assess the potential risks relative to the potential benefits for each active drug compared to control. The DSMB may also examine the safety and efficacy data in subgroups known to be high risk for bleeding such as those with older age and/or higher BMI. The DSMB will use the monitoring information to determine its recommendation to NHLBI. The DSMB can recommend that the Outpatient COVID-19 trial should continue as proposed, that one treatment arm may be dropped, that the trial protocol should be modified, or that the Outpatient COVID-19 trial should be terminated early for safety reasons.

4.5 STUDY STAGES AND INTERVENTIONS

The first Stage of this study has been determined and is outlined above. In Stage 1, there will be four intervention arms: (1) prophylactic anticoagulation with apixaban 2.5mg po bid; (2) therapeutic anticoagulation with apixaban 5.0mg op bid; (3) antiplatelet therapy with low dose aspirin 81mg po qd and (4) placebo. Subsequent Stages will incorporate recommendations from the DSMB.

The overarching plan for adaptive changes are as follows:

1. **If an active drug is found to be futile relative to placebo** (i.e. results indicate that an active arm is associated with a slightly reduced risk, no effect, or a greater risk of the primary outcome as compared with placebo): The futile active arm will be dropped, no new treatment arm will be added, and the trial will continue with the remaining treatment arms. The randomization scheme will be adjusted to include the 3 remaining arms with equal probabilities (i.e. 1:1:1), and the treatment comparisons among these arms will continue as designed.

2. **If an active drug is found to be superior to placebo**: We will declare a winner, and we will announce this finding. The placebo arm will be dropped. If the observed differences between the superior active arm and all of the other active arms are sizable (e.g. >20% relative reduction) but do not yet cross the decision boundary, the trial may be terminated based on a risk/benefit analysis by the DSMB. If the observed differences between the superior active arm and at least one of the other active arms is small, this would be announced, and the trial may continue with the “competitive arms” based on a risk/benefit analysis by the DSMB. The randomization scheme will be modified to assign each of the remaining treatment arms with an equal probability. The aspirin arm will become the reference arm for future statistical models.

3. **If a promising new drug is identified from external studies**: At the outset of this trial, we do not plan on adding any new treatment arms. However, if a promising candidate drug were to be identified in the next 6 months, we will consider adding an arm to the trial based on time and other pragmatic considerations. The randomization scheme and analytic approach would be modified to include an extra treatment arm.