Statistical Analysis Plan
ASPEN-COVID-19 Phase 2b

Study Title:	**Assessing Safety, Hospitalization and Efficacy of rNAPc2 in COVID-19 (ASPEN-COVID-19)**
Sponsor:	ARCA biopharma, Inc. 10170 Church Ranch Way, Suite 100 Westminster, Colorado 80021 Phone: 720-940-2100
Protocol Number	NAPc-201/301
ClinicalTrials.gov Identifier	NCT04655586
Study Drug:	rNAPc2
Comparator:	Heparin
Analysis Plan Version	3.0
Date:	February 24, 2022

Table of Contents

Approvals

Signatures		
Biostatistician (ARCA)	_____ Ian Carroll, PhD	_____ Date
Biostatistician (CPC)	_____ Karen Liu, MSc	_____ Date
Principal Investigator	_____ Marc Bonaca, MD, MPH	_____ Date
Chief Scientific Officer (ARCA)	_____ Michael Bristow, MD, PhD	_____ Date

Revision History

VERSION	DESCRIPTION OF CHANGE
1.0	Original
2.0	• Inserted language submitted in IND 150187, Sequence 0015 on 29 Nov 2021 containing Sponsor's response to the Division's review of Version 1.0 clarifying Wilcoxon Signed Rank will not be used in the primary efficacy endpoint analysis. • Added composite of clinical events as a secondary efficacy endpoint in Section 2.5.1. • Moved probability of discharge as assessed by ACTT from exploratory efficacy endpoint to secondary endpoint in Section 2.5.1 • Added multiplicity control description to Section 5.7 • Added Section 5.8 Transformations and Normality • Added Section 7.1.2 Secondary Analysis of the Primary Efficacy Endpoint • Added Section 7.2.3 Composite of Clinical Events • Added Section 10 Composite and Net Endpoints • Added Section 11 Subgroup and Sensitivity Analyses to provide additional description of how types of missingness will be addressed via multiple imputation. • Added Section 14.2 Description of Models and Methods which include descriptions of the Bayesian models, methods for multiple imputation, and methods for power analysis used in this study. • Added Section 12 Changes from Methods Planned in the Protocol to describe the key changes in either endpoint or analysis methodology.
3.0	• Corrected the censoring date for death in Section 7.2.2 to "day x+1, where x denotes the longest observation window among all subjects in the analysis" from "day of death". This brings censoring date for death into alignment with the NAPc-201 clinical protocol language. • Added clarifying language to Section 7.2.3 regarding pre-existing organ support of any modality as an additional covariate, representing a special case arising from the specifications of the analysis.

List of Abbreviations

Abbreviation	Term
ACTT	Adaptive COVID-19 Treatment Trial
AE	Adverse Event
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CEC	Clinical Endpoint Committee
CFR	Code of Federal Regulations
CI	Confidence Interval
CRF	Case Report Form
CSR	Clinical Study Report
CTMS	Clinical Trial Management System
DAPT	Dual Antiplatelet Therapy
DMP	Data Management Plan
DSMC	Data Safety Monitoring Committee
eGFR	Estimated Glomerular Filtration Rate
FDA	Food and Drug Administration
hsCRP	High Sensitivity C-reactive Protein
ICF	Informed Consent Form
IL-6	Interleukin-6
ITT	Intent-to-Treat
kg	Kilogram
LLN	Lower Limit of Normal
LMWH	low molecular weight heparin
MCAR	Missing Completely At Random
MedDRA	Medical Dictionary for Regulatory Activities
mg/dl	Milligram per deciliter
mITT	Modified Intent-to-Treat
mmHg	Millimeters of mercury

Abbreviation	Term
NSAIDS	Non-Steroidal Anti-Inflammatory Drugs
PCFS	Post-COVID Functional Status
PCR	Polymerase Chain Reaction
PT	Prothrombin Time
PTm	Preferred Term
PTT	Partial Thromboplastin Time
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Standard Deviation
SoC	Standard of Care
SOP	Standard Operating Procedure
TEAE	Treatment-Emergent Adverse Event
TFLs	Tables, Figures, and Listings
UFH	Unfractionated Heparin
ULN	Upper Limit of Normal
URC	Unblinded Review Committee
WHO DDE	World Health Organization Drug Dictionary (Enhanced)

1 Introduction

The ASPEN-COVID-19 ("ASPEN") protocol comprises sequential Phase 2b and Phase 3 randomized, multicenter, active comparator studies evaluating the hypothesis that rNAPc2, a novel, potent and highly selective tissue factor inhibitor with anticoagulant, anti-inflammatory and potential antiviral properties, reduces elevated D-dimer levels (Phase 2b primary endpoint) and shortens time to recovery (Phase 3 primary endpoint) compared to heparin in hospitalized patients with COVID-19 and elevated D-dimer levels. Study participants and Clinical Endpoint Committee (CEC) members assessing the clinical endpoints are blinded to treatment assignment. Analysis of Phase 2b data could lead to study discontinuation, adjustment of eligibility criteria or sample size, and will inform the rNAPc2 dose regimen to be studied in Phase 3.

This statistical analysis plan (SAP) describes all study related details and statistical analyses and summaries to be generated on data collected from the Phase 2b ASPEN-COVID-19 study.

The SAP is designed to assist in the analysis of clinical study data for final reporting of the Phase 2b study. This content includes, but is not limited to, statistical analysis and summary of all endpoints identified in the study protocol, heretofore referred to as 'the study' or 'study.'

Details of reporting for the purpose of the Data Safety Monitoring Committee (DSMC) and other committees are documented elsewhere. Details of planned interim analyses are documented in the DSMC Charter and DSMC Reporting Plan. Reporting of the Phase 3 study will be documented in a separate analysis plan.

2 Study Design and Plan

2.1 Design and Primary Objectives

The primary objective of the Phase 2b study is to identify an rNAPc2 dosing regimen with an acceptable safety profile that reduces D-dimer levels compared to standard of care (SoC) heparin regimens. Secondary and exploratory objectives are to compare the treatment groups in terms of biomarkers associated with outcomes including those related to coagulation and inflammation, as well as for clinical and functional outcomes.

2.2 Sample Size Determination

During Phase 2b, approximately 160 participants will be enrolled and randomized 1:1:2 to rNAPc2 dose regimen 1, rNAPc2 dose regimen 2, and heparin SoC, respectively. Assuming approximately 20% of participants may not be evaluable for any reason, 160 randomized participants will provide for approximately 130 evaluable participants. At a sample size of 160, the study will have adequate power across a range of clinically meaningful effect sizes, operationalized as proportional reductions in D-dimer (relative to heparin) from 20% to 40% in increments of 5% (see below). This sample size was selected, based on a statistical simulation using Wilcoxon rank sum and assuming a 20% missing/dropout rate, to ensure inferences for the primary endpoint analysis will be robust, given certain potential confounding factors such as missing data (random or non-random), heterogeneous application of heparin (low molecular weight and unfractionated; prophylactic and therapeutic), effects of subject-level covariates (e.g., comorbidities), and effects of site-level covariates (e.g., country).

Power, Primary Endpoint							
			Effect Size (Proportional Reduction)				
	Enrolled Subjects	*Evaluable* Subjects	*20%*	*25%*	*30%*	*35%*	*40%*
Sample Size	80	64	55.8%	67.4%	75.6%	81.0%	84.5%
	100	80	60.5%	72.7%	80.1%	84.9%	87.8%
	110	88	63.6%	76.2%	80.7%	86.6%	89.2%
	120	96	65.2%	78.0%	82.8%	86.9%	91.0%
	130	104	68.2%	79.3%	84.8%	88.7%	91.7%
	140	112	68.9%	80.7%	86.6%	89.3%	92.5%
	160	128	71.3%	83.1%	88.3%	91.3%	94.1%
	180	144	74.9%	84.5%	89.5%	92.5%	94.6%
	200	160	77.2%	86.5%	91.0%	93.8%	95.8%

*Significance via $\alpha = 0.05$
†Simulated with 20% Missing Data
‡Percentages as Averages Across Other Simulation Conditions

2.3 Participant Selection, Randomization, and Blinding/Unblinding

Approximately 160 participants 18 to 90 years of age with COVID-19 documented by a validated test such as polymerase chain reaction (PCR) within 1 week of hospitalization or screening and D-dimer above the upper limit of normal will be eligible for the Phase 2b study.

Participants will be randomly assigned 1:1:2 to receive rNAPc2 at one of 2 dose regimens, or a heparin regimen per site SoC (prophylactic or therapeutic). The higher rNAPc2 dose regimen will include a 7.5 µg/kg subcutaneous (SC) loading dose on Day 1 followed by 5µg/kg SC on Days 3 and 5; the lower dose regimen will be administered as a 5 µg/kg SC loading dose on Day 1 followed by 3µg/kg SC on Days 3 and 5. Prophylactic or therapeutic doses of low molecular weight heparin (LMWH) or unfractionated heparin (UFH) will be prescribed in the control group as per local SoC. Participants will be stratified by local laboratory D-dimer level at screening (≤ or > 2× the assay's upper limit of normal (ULN)).

The NAPc-201/301 study utilizes a modified Prospective Randomized Open-Label, Blinded Endpoint (PROBE) design. In a traditional PROBE design, the participant and investigator are aware of treatment assignment, but endpoints are classified by blinded adjudicators. In the ASPEN trial, the site investigator(s) and research coordinator(s), as well as the treating clinical team are aware of treatment

assignment, but the randomized participants will not be told their treatment assignment and all efforts will be made at the site to keep them blinded in order to minimize bias in safety and clinical endpoint reporting. In addition, adjudicators of clinical events will be blinded to treatment assignment as well as other information relevant to unblinding in order to minimize any potential bias in event categorization. To support the integrity of study results, blinding of study management personnel (Sponsor, Contract Research Organization/Academic Research Organization, Oversight Committees, etc.) will be managed at the study role level as described in the Blinding Plan. Additionally, neither the study site, Sponsor, CEC, nor study participants will have access to the central laboratory results during the conduct of the trial with the exception of clinical safety alerts. The investigator will not be provided with randomization codes. The study site will know the treatment assignment; however, participants should not be made aware of their treatment assignment.

2.4 Primary Endpoints

2.4.1 Efficacy

The primary efficacy endpoint is proportional change, represented as percent change, in D-dimer level from baseline to Day 8 or Day of Discharge if prior to Day 8.

Language pertaining to the primary efficacy endpoint in the ASPEN Protocol is operationalized as proportional change.

2.4.2 Safety

The primary safety endpoint is major or non-major clinically relevant bleeding within eight (8) days of randomization as compared to heparin.

2.5 Secondary Endpoints

2.5.1 Efficacy

Secondary efficacy endpoints are:

- Change in D-dimer level from baseline to 24 hrs post the institution of study medication following randomization (Study Day 2), and 48 hours after institution of study medication (Study Day 3)
- Probability of discharge as assessed by ACTT $\geq$ 6 by Day 8 and Day 30 post randomization
- Composite of death, thrombosis, organ support, or rehospitalization within 30 days adjudicated as to its relatedness to COVID-19
- Change in biomarkers associated with outcomes including those related to coagulation and inflammation from baseline through Day 8

2.5.2 Safety

Secondary safety endpoints are:

- Major or non-major clinically relevant bleeding with rNAPc2 vs. heparin
- Bleeding with higher vs. lower dose rNAPc2 through Day 30
- Other adverse events (AEs)

3 Data Set Specifications

3.1 Database Description

The following data sources will be used to create the final study tables, figures, and listings (TFLs).

- Protocol Deviation database which will store protocol deviations
- Clinical Endpoint Adjudication database which will store all adjudicated review fields and outcomes
- Central laboratory data from PPD
- Clinical database which will store all other data used for analysis

These databases are hosted by the following systems:

- Protocol Deviation database: DATATRAK UX CTMS
- Adjudication database: AG Mednet
- Central lab data from PPD transferred to CPC as SAS dataset
- Clinical database: DATATRAK UX EDC™

Separate data management and adjudication plans will describe the set-up and management of these databases including preparing and selecting data for study analysis. The adjudication, laboratory and clinical databases will be reconciled to ensure within- and cross-form consistency.

3.2 Creating Analysis Data Sets

To facilitate the analysis of study data, it will be necessary to extract and combine the relevant variables from the data sources described above. All analysis data sets will be generated using SAS software (SAS and all other SAS Institute Inc. products or service names are registered trademarks of SAS Institute Inc., Cary, NC, USA).

3.3 Generation and Validation of Data Analysis Programs and Output

Programming and output developed prior to the database lock will be produced and displayed using dummy treatment codes to protect the integrity of the study and minimize any unintended bias in accordance with the Study Blinding Plan.

3.4 Archiving Data and SAS Files

Following final database lock, the SAS® programs used to analyze data for the final Clinical Study Report (CSR) will be executed and archived.

4 Considerations and Methodology

The analysis of data for the study will be based on the objectives and endpoints outlined in the protocol. Final analysis will be produced on locked clinical study data using SAS version 9.4 or higher.

Minor edits and additions which do not affect the content or meaning of either the text or mock shells (e.g., typos, spacing, etc.) will not constitute the need for an amendment to this plan. Major changes to this reporting plan will require an amendment and will require approval. All changes which are made to the reporting plan from the initial approval will be tracked on a Programming Changes Log.

4.1 Reporting Timeframes

This reporting plan describes the required output for final reporting for the Phase 2b study and will occur when all participants have completed the Phase 2b study and those data have been cleaned and locked according to the Data Management Plan (DMP).

Interim reporting of efficacy and safety data to a DSMC, including planned interim analyses, will be described in the DSMC charter and a separate reporting plan.

4.2 Analysis Display Conventions

General conventions for statistical analyses are pre-specified. Departures from these general conventions may be given in the specific detailed sections of this SAP. When this occurs, the rules set forth in the specific section take precedence over the general convention.

- All tables will have the population sample size for each dosing group in the column heading.
- Continuous variables will be summarized using the standard display of n, mean, standard deviation (SD) medians, minimum, and maximum, first and third quartiles for participants with non-missing data.
- Ordinal or categorical variables will be summarized using the number and percentage of observed response in each category.
- Summary statistics for discrete variables will consist of the number and percent of responses in each category. The count and percentage of responses will be presented in the form of XX (XX.X%), where the percentage is in parentheses. If the count is "0" then the percentage may not be presented. If the percentage is "100", the decimal place may be dropped. In addition, the decimal place may also be dropped due to space constraints within a table where there is no impact on interpretation. Denominator for percent will be footnoted if not the population size in the column heading.
- Leading zeros will be used for decimal values less than one. For example, use "0.05" rather than ".05."
- *P* values will be presented as 4 decimals (e.g., "0.0054") or replacing the leading 0 with "< .0001" and "> .9999" if applicable.
- A participant's age will be calculated using birth date and informed consent date. Age will be reported in years and rounded to the lowest (floor) integer.
- Body Mass Index (BMI) calculation is based on the formula below:

$$\text{BMI (kg/m}^2\text{)} = \text{Weight (in kg)} / \text{Height}^2 \text{ (in meters)}$$

- All medical history and AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA) as described in the DMP.
- All medications will be coded using the World Health Organization Drug Dictionary [Enhanced] (WHO DDE) as described in the DMP.

4.3 Determination of Study Day

The day the participant is randomized will be considered study "Day 1." The day prior to randomization will be considered study "Day -1." There will be no Day 0. If the start day is missing the first day of the month will be used.

If the start year or stop year is missing it is assumed the medication was taken concomitantly with study drug.

4.4 Definition of Phases, Visit, or Windows

Visit labels will be used for reporting. For listings ordered by assessment date, if unscheduled visits are associated with a date of assessment, these will be included in the listing by date with the visit described as 'unscheduled.' Otherwise, if no date of assessment is associated with the unscheduled visit data, listings will be ordered by visit with unscheduled visits presented last. Unscheduled visits associated with discharge from hospital will be used in lieu of Day 8 in participants for whom there is no Day 8 assessment, and other unscheduled visits will not be reported in summaries.

Visits will be displayed as follows:

Visit Label	Short Descriptor if needed to accommodate space restrictions
Screen	SCRN
Baseline	BSLN
Day 2 (24h post-dose)	D2
Day 3	D3
Day 5	D5
Day 8	D8
Day 10	D10
Day 30	D30
Unscheduled	UNSCH

4.5 Display of Treatment Groups

Summaries will be displayed by treatment group. Summaries and listings based on the Safety population will be displayed by actual treatment received.

Treatment groups will be displayed as follows with respect to labels and ordering.

Label	Short Label if needed to accommodate space	Definition
All Heparin	All HEP	Prophylactic and therapeutic doses heparin groups combined
Heparin Prophylactic	HEP Pro	Initial Heparin Dosing Strategy (Prophylactic Heparin) as recorded by the investigator in the Case Report Form (CRF)
Heparin Therapeutic	HEP Ther	Initial Heparin Dosing Strategy (Therapeutic Heparin) as recorded by the investigator in the CRF
rNAPc2 Lower		Lower Dose rNAPc2 (5 µg/kg on Day 1 followed by 3µg/kg on Days 3 and 5)
rNAPc2 Higher		Higher Dose rNAPc2 (7.5 µg/kg on Day 1 followed by 5µg/kg on Days 3 and 5)
ALL rNAPc2		Lower and Higher rNAPc2 doses combined
Overall		All participants who receive any study treatment (Heparin or rNAPc2)

4.6 Protocol Deviations or Violations

All protocol deviations will be collected in the clinical trial management system (CTMS) database. Major deviations will be identified manually and programmatically prior to unblinding. A summary of major protocol deviations will be given by frequency tables showing the number of participants with major protocol deviations overall and by deviation. Protocol deviations will be listed for all screened participants.

5 Statistical Methods

5.1 Missing Data

The primary analysis will be performed using all complete data and, for subjects who do not have completed primary endpoint data, values obtained in the ± 48-hour window around Day 8 or Day of Discharge, as described in Section 7.1.1, will be used. If no values are present from the ± 48-hour window, Last Observation Carried Forward (LOCF) will be used. All values measured after the administration of the final dose will be eligible for LOCF.

5.1.1 Missing Values

In cases where data are missing but an entry is expected (e.g., required study procedure not performed) numeric values will appear in the database as follows:

- ‘Unable to obtain’ response will be recorded in the database as a ‘-97’.
- ‘Not Known’ response will be recorded in the database as a ‘-98’.

- 'Not Applicable' response will be recorded in the database as a '-99' value.

These will be translated to meaningful entries during listing and/or summarization of the data as appropriate.

5.1.2 Undetectable Values

If a measured value of D-dimer is reported from the central laboratory as below the detectible limit of the assay, the value will be entered as 1 unit below the detectible limit. For example, if the result is '< 120 ng/mL' the value will be entered as 119 ng/mL.

5.2 Analysis Populations and Scopes

All analyses are based on 2 elements:

- Analysis population, which specifies which participants will be included in an analysis; and
- Data scope, which specifies the time window within which data will be included in an analysis

Five analysis populations are defined for this study.

The **Intent-to-Treat (ITT) population** will include all participants who are randomized. Participants will be analyzed according to the treatment group allocated by randomization.

The **modified Intent-to-Treat (mITT) population** will include all participants who are randomized and received at least one dose of study drug. Participants will be analyzed according to the treatment group allocated by randomization.

The **Per Protocol population** will include all participants who are randomized, receive all three doses of study drug post randomization or receive SoC drug through day 7, have no major deviations from the protocol that affect the primary endpoint, and complete all pre-specified study assessments/activities. Participants will be analyzed according to the treatment actually received.

The **Safety population** will include all participants who are randomized and receive at least one dose or part of a dose of study treatment post-randomization. Participants will be analyzed according to the treatment actually received.

The **All Screened population** will be defined as all participants with any data entered into the database.

The analysis populations will be listed by participant.

The ITT, mITT and Per Protocol data scope includes outcomes and events observed from a given participant's date of randomization until either their final or Day 30 contact, determined by whichever occurs first. This will be applied mainly to the analyses of efficacy endpoints.

The ITT, mITT and Per Protocol data scope will also be used for the secondary safety endpoints of major or non-major clinically relevant bleeding through Day 30 and any bleeding through Day 30.

The on-treatment data scope will include all outcomes and events observed from randomization until 72 hours following permanent discontinuation of the study drug. This on-treatment data scope will be applied mainly to the analyses of safety endpoints.

5.3 Subgroups

For the primary efficacy endpoint, the treatment effects across the following subgroup factors will be examined:

- D-dimer randomization stratification factor (> vs. ≤ 2x local laboratory ULN)
- Treatment with an agent with demonstrated efficacy in patients with COVID-19, as described in Section 6.2.1 (yes, no)
- Background antiplatelet therapy at baseline, as described in Section 6.2.1
- rNAPc2 dose regimen
- Initial heparin dosing strategy (prophylactic vs. therapeutic; within regimen comparison only)
- Severity of disease (as defined in Section 14.1)
- Source of Laboratory Values used for primary endpoint determination (central vs. local)

5.4 Adjustments for Covariates

Where indicated analyses will include a covariate for study stratification 'D-dimer level at screening' (> or ≤ 2X the assay's ULN).

5.5 Multi-center Studies

Data from all centers will be pooled for analysis.

5.6 Pooling of Dose Groups

The primary analysis will involve a comparison of the pooled rNAPc2 treatment groups against the heparin treatment group. Because the optimal dose of rNAPc2 is unknown and its determination is a key objective of Phase 2b, the assumption that both doses will be effective and thus can be pooled for statistical analysis may not be correct. Therefore, each rNAPc2 dose regimen will also be analyzed as separate groups.

Although this variable is included as a pre-specified subgroup of interest, it is of higher inferential priority because it does not share the same characteristics of the other subgroups, attributable entirely to the fact that group membership is determined solely by the experimental manipulation (i.e., treatment assignment) rather than any subject- or clinically-derived characteristic.

5.7 Multiplicity Issues

Multiplicity control will be implemented for non-primary endpoints as described in Gou et al (2014), which is an improvement uniformly over the stepwise Hochberg procedure and in most cases over the Hommel procedure.

5.8 Transformations and Normality

In circumstances where univariate or multivariate normality is assumed but violated (where violations are determined via Shapiro-Wilk test), featured in the present study primarily in the Bayesian Linear Mixed Effects models, variable transformation will be performed in order to satisfy normality assumptions. Candidate transformations include square root, logarithmic, and reciprocal transformation for incremental severity of distributional skewness.

The expected circumstances where transformation may need to be applied are certain biomarkers like D-dimer.

6 Baseline Characteristics and Subject Disposition

6.1 Subject Disposition

Subject disposition will be summarized by treatment group in a table for all screened participants, and analysis population (ITT, mITT, Per Protocol) will be reflected within the table.

6.2 Demographics and Baseline Characteristics

Baseline demographic information and clinical characteristics including, but not limited to, sex, race, ethnicity, age, ACTT ordinal scale category, severity as identified by the WHO COVID severity definitions (Appendix 14.1), and study entry D-dimer value will be listed and summarized by treatment group for the ITT and Per Protocol populations and listed for all screened participants.

6.2.1 Baseline (pre-randomization) & Concomitant (post-randomization) Medications

Concomitant medications consist of any medication given during the observation window. Concomitant medications will be coded using WHODrug as described in the DMP. Medications will be listed and summarized by Anatomical Therapeutic Chemical (ATC) Class 2nd Level, medicinal product name, and treatment group. Selected relevant medication classes will be presented as shown below. Concomitant medications at screening will be listed by participant and summarized.

Selected Relevant Concomitant Medication Classes:

- Antiplatelet
    - Aspirin
    - P2Y12 inhibitor – e.g., clopidogrel, prasugrel, ticagrelor, ticlopidine
    - Other – e.g., dipyridamole; vorapaxar, cilostazol
- Anticoagulants
    - Warfarin
    - Factor Xa inhibitors (apixaban, rivaroxaban)
    - Unfractionated heparin
    - Low molecular weight heparin
- COVID therapies
    - Systemic (non-topical) corticosteroids
    - Remdesivir
    - Baracitinib
    - Tocilizumab
    - Convalescent and non-convalescent plasma
    - Monoclonal or polyclonal Abs directed against SARS-CoV-2
    - Anti-inflammatory agents
- NSAIDS

- Vasopressors
- Inotropes
- SARS-CoV-2 vaccine (Vaccine received yes or no; manufacturer, if known)

If medication start date and end date are blank, it is assumed the medication was taken concomitantly with study drug. If medication start date and end date are both prior to Day 1, the medication will be considered as taken 'prior' to study drug. If medication start date is prior to Day 1 and end date is blank, the medication will be considered 'ongoing' and as a concomitant medication.

6.2.2 Medical History

The incidence of medical/surgical history (including COVID-19 and cardiovascular specific medical history) will be listed and summarized by treatment group for the ITT and Per Protocol populations and listed for all screened participants.

6.2.3 Vital Signs

Vital Sign data will be listed and summarized by treatment group for the ITT and Per Protocol populations and listed for all screened participants.

6.2.4 Physical Examination

Vital Sign and Physical Examination data will be listed and summarized by treatment group for the ITT and Per Protocol populations and listed for all screened participants.

6.3 Treatment Exposure and Adherence

Treatment exposure and adherence will be summarized by treatment group for the ITT, Per Protocol and Safety population and listed.

7 Efficacy Analysis

The following efficacy analyses will be based on the ITT population and data scope. Accompanying analyses pertaining to mITT and Per Protocol population, subgroup and sensitivity are discussed in Section 11.

7.1 Primary Efficacy Endpoint

The primary efficacy endpoint of Phase 2b is proportional change in D-dimer level from baseline (Day 1 visit where randomization occurs and central laboratory samples are drawn) to Day 8, or Day of Discharge if prior to Day 8 between treatment groups. The primary analysis will involve a comparison of the pooled rNAPc2 treatment groups against the heparin treatment group.

7.1.1 Primary Analysis of the Primary Efficacy Endpoint

Proportional change, represented by percent change, will be calculated as 100 × (D-Dimer level at Day 8 or Day of Discharge if prior to Day 8 - D-Dimer level at baseline)/ D-Dimer level at baseline and will be based on data provided by the central laboratory, or local laboratory paired samples if the central laboratory values are not available. The pairing process will be operationalized as outlined:

For Day 1 (baseline) paired local-local or central-central laboratory values:

1. Use any lab D-dimer drawn during the current admission and up to randomization

2. Use the laboratory sample closest to randomization if there are multiple D-dimer measurements.
3. All samples obtained more than 1-hour post-initiation of study drug are ineligible.

For Day 8 or Day of Discharge paired laboratory values:

1. Use laboratory samples collected within ± 48 h of the Day 8 or Day of Discharge study visit (whichever occurs first).
2. If there are multiple laboratory samples collected within that window, use the lab value on the target day; if not available, use the lab value immediately after the target visit day.

A Wilcoxon rank sum test will be used to compare proportional change from baseline to Day 8, or Day of Discharge if prior to Day 8, between treatment groups. The between treatment group rank sum test is considered the primary analysis of the primary endpoint.

7.1.2 Secondary Analysis of the Primary Efficacy Endpoint

A within treatment analysis of the proportional change in D-dimer levels will be a secondary analysis, subject to multiplicity control as for secondary endpoints. Wilcoxon signed rank tests will be used to compare within treatment group baseline vs. Day 8 D-dimer values, or Day of Discharge if prior to Day 8, values within each treatment group (i.e., rNAPc2 pooled, Heparin), (i.e., rNAPc2 pooled vs. Heparin).

7.2 Secondary Efficacy Endpoints

Secondary efficacy outcomes are proportional change in D-dimer level from baseline to 24 hours post-dose (Day 2) and Day 3, probability of discharge accompanied by ACTT ≥ 6 by Day 8 and Day 30, the secondary efficacy clinical composite, and change in biomarkers associated with outcomes including those related to coagulation and inflammation from baseline to Day 8 (or Day of Discharge if prior to Day 8). Inflammatory biomarkers include interleukin-6 (IL-6) and high sensitivity C-reactive protein (hsCRP), while coagulation biomarkers include prothrombin time (PT), partial thromboplastin time (PTT), factor Xa, and tissue factor.

7.2.1 Analysis of Secondary Efficacy Endpoints

Secondary efficacy endpoints will be subject to the same analysis, same sensitivity analyses, and same subgroup analyses as the primary efficacy endpoints.

7.2.2 Time to Recovery

Time to recovery by Day 30, with Day 1 being the date of randomization. Recovery is defined as the first day a participant meets criteria for category 6, 7, or 8 of the ACTT ordinal scale. The analysis of time to recovery will be a comparison of the combined rNAPC2 dose groups vs. SoC heparin by a stratified log-rank test, where stratification is according to local laboratory screening D-dimer level ≤ or > 2× local laboratory ULN. Cumulative incidence by treatment group will be summarized by Kaplan-Meier estimates. Additionally, a treatment hazard ratio and corresponding 95% confidence interval will be estimated by a Cox proportional hazard model stratified by local laboratory D-dimer level at screening. In addition, the two rNAPc2 dose groups will be compared for time to recovery.

Participants who die within 30 days of randomization will be censored at the day x+1, where x denotes the longest observation window among all subjects in the analysis. In addition, participants surviving to their Day 30 visit who have not met the criteria for recovery by the time of their last ACTT scale assessment on or prior to the Day 30 visit will be censored on the day of their last assessment on or prior to the Day 30 visit. There will be no censoring prior to Day 30 due to study treatment discontinuation,

initiation of non-study treatment, or intermittently missed ACTT scale assessments. Patients lost to follow-up or otherwise of unknown status occurring prior to Day 30 will be censored based on, and at the day of, the last known status.

The frequency and percentage of participants with each recovery category (ACTT equal to 6, 7, or 8) will be summarized.

7.2.3 Composite of Clinical Events

The secondary efficacy clinical composite will be measured and compared between treatment groups. This composite will consist of death (all-cause), thrombotic events (Stroke, MI, PE, peripheral arterial and clinically relevant deep vein thrombosis), organ support as defined in Section 7.3.2, and rehospitalizations adjudicated as to its relatedness to COVID-19.

The primary formulation of the composite will limit the number of contributing events per component per patient to one, i.e., the aggregate will have values from 0 – 4 for each patient. The component-wise analysis will be performed using logistic regression, where the contribution of each component is necessarily a dichotomous variable. The aggregate will be analyzed using Poisson regression.

The secondary formulation will differ from the primary in that no limitation on the component-wise contributions will be imposed. Thus, component-wise analysis will also be Poisson regression except in the case of death or any data-driven circumstance where the only available values are binary.

Regression models will include treatment in the unadjusted formulation, and treatment, D-dimer stratum and pre-specified subgroups in the covariate-adjusted formulation. Pre-existing organ support of any modality is discussed below as an additional covariate, representing a special case arising from the specifications of the analysis.

Assumptions inherent to any of these models, e.g., whether data manifest in accordance with a Poisson distribution (where λ, the rate, is both mean and variance) will be tested. In the case of Poisson, testing significant deviations from assumptions will use Likelihood Ratio tests on Zero-inflated and/or Negative Binomial operationalizations of the dependent variable distribution (these are not mutually exclusive). Both Logistic and Poisson regression models may remove variables that introduce issues in the model estimation process, e.g., a regressor with no variance, cell sparsity, or sufficient multicollinearity to induce rank deficiency of the model-implied design matrix will flatten the likelihood function, destabilizing point estimates and inflating confidence intervals by orders of magnitude.

Events in these analyses must be treatment-emergent to be eligible for inclusion, and therefore subjects with pre-existing organ support will need adjustment in order to properly estimate effects related to initiation of new organ support. This will manifest as a binary covariate indicating pre-existing organ support of any modality.

7.2.4 Coagulation and Inflammation Biomarkers

The following laboratory parameters will be collected and analyzed by a central laboratory. Change from baseline to both Day 8 or discharge will be summarized for inflammatory biomarkers (IL-6, hsCRP) and coagulation biomarkers (PT, PTT, factor Xa and tissue factor). All data will be listed. Primary analyses will be by Wilcoxon-family tests as described above in the primary efficacy endpoint. For purposes of interpretation of outcomes, IL-6 and hsCRP are considered measures of systemic inflammation, PT, PTT and factor Xa are coagulation parameters, and tissue factor is the biologic target of rNAPc2. In addition,

in a subset of participants anti-phospholipid antibodies known to be generated in COVID-19 and inhibited by rNAPc2, are being measured at baseline, Day 8/discharge and on Day 30 at a specialty laboratory.

7.3 Exploratory Efficacy Endpoints

7.3.1 Composite of All-Cause Mortality and Thrombotic Events

Time to first occurrence of a composite of all-cause mortality and any adjudicated thrombotic event within 30 days of randomization and time to all-cause mortality within 30 days of randomization will be analyzed by the same methods as described for ACTT above, with the exception that deaths will be counted as an event. Participants without an adjudicated event by Day 30 will be censored on the study day of last contact or Day 30, whichever is earlier. There will be no censoring prior to Day 30 due to study treatment discontinuation or initiation of non-study treatment.

7.3.2 Healthcare Resource Utilization

Days in intensive care unit, on ventilator, vasopressors, renal replacement therapy, and circulatory support will be summarized with descriptive statistics.

Organ support free days will be compared between the rNAPc2 and heparin arms via Wilcoxon Rank Sum testing, as in analysis of the Primary Efficacy Endpoint. Organ support free days is defined as total number of days without any of the following: placement on invasive mechanical ventilation, intravenous inotrope or vasopressor administration, renal dialysis, mechanical circulatory support, or death will be documented for each patient, and the number of total as well as individual interventions per patient will be compared by treatment group. Subjects who die are assigned a score of -1 for organ support free days. In the event that a large number of ties exist (e.g., a substantial proportion of subjects have the entire 30 days free of organ support), then organ support free days will be categorized as 30, 0-29, and -1 representing no organ support needed, some organ support needed, and death, respectively. An equivalent scenario involving every day requiring organ support for many subjects would yield 1-30, 0, -1 as categories. This tabulation with be analyzed subsequently using an Exact Wilcoxon test, such that ordinality across categories is maintained. In a circumstance where only two of the three categories are observed, this becomes an exact binomial test.

7.3.3 Post-COVID Functional Status

Post-COVID Functional Status (PCFS) score at discharge and Day 30 will be summarized with descriptive statistics.

7.4 Significance

All statistical tests and confidence intervals will use a 2-sided analysis with a significance level of $\alpha = 0.05$.

8 Safety Analysis

The following safety analyses will be based on the Safety population and on-treatment data scope. Additional safety analyses are detailed in Section 11.4.

8.1 Primary Safety Endpoint

The primary safety endpoint is major or non-major clinically relevant bleeding through the earliest of Day 8 or two days after Day of Discharge if prior to Day 8.

The frequency and percentage of participants with major or non-major clinically relevant bleeding within 8 days of randomization as adjudicated by the CEC will be presented by each rNAPc2 dose and any heparin dose.

8.2 Secondary Safety Endpoints

8.2.1 Major or non-major clinically relevant bleeding through Day 30

The frequency and percentage of participants with major or non-major clinically relevant bleeding through Day 30 as adjudicated by the CEC will be presented by each rNAPc2 dose and heparin dose (pooled and via characteristics, i.e., Prophylactic vs. Therapeutic, Unfractionated vs Low Molecular Weight).

8.2.2 Any bleeding through Day 8 and through Day 30

The frequency and percentage of participants with any bleeding major, non-major clinically relevant, non-major and not clinically relevant as defined in the protocol) through Day 8 and through Day 30 will be presented by each rNAPc2 dose and heparin dose (pooled and via characteristics, i.e., Prophylactic vs Therapeutic, Unfractionated vs. Low Molecular Weight).

8.2.3 Other Adverse Events

Safety assessments will also include monitoring of serious adverse event(s) (SAE) and non-serious adverse event(s) (AE). Efficacy and safety endpoints that are adjudicated positively as study outcomes, including all bleeding events, will not be considered as AEs or SAEs.

All AEs will be collected from the time of randomization until the participant completes Day 30 or final contact, whichever is longer. All SAEs will be collected from the time of consent until the participant completes Day 30 or final contact, whichever is longer.

All reported AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

Participant counts will be summarized using frequencies and percentages. A participant will be counted only once per category of summarization (e.g., if a participant has multiple AEs, the participant will only be counted once in the overall ‘Any AE’ row and will be counted once in each row for the specific AE category reported). Where participants have more than one level of severity/relatedness recorded only the highest level of severity/relatedness will be summarized.

Participant counts will be summarized by system, organ class and preferred term (PTm) by treatment group for treatment-emergent AEs (TEAEs) and by seriousness, relatedness.

An AE identified by the investigator as ‘Related’ or ‘Potentially Related’ or where relationship is missing will be considered ‘related to study treatment.’

All AEs will be listed. Listings of SAEs, Related AEs, fatal and AEs of special interest (e.g., Overdose, misuse, or accidental exposure to study drug; medical error involving study drug), will also be listed. Additional details of SAEs and Deaths will also be listed.

Treatment Emergent is defined as any SAE/AE that has a start date during or after administration of the investigational product (i.e., randomized rNAPc2 or heparin treatment), post-randomization up to Day 30 or final contact.

Data will be displayed as collected in the clinical database. Adverse Events with missing start dates will be considered as treatment emergent.

8.3 Laboratory Data

A summary of laboratory (Central Laboratory) values (e.g., AST, ALT and bilirubin) outside normal limits (i.e., either above the ULN or below the lower limit of normal (LLN) where limits are defined), as well as a summary of any participant potentially meeting the criteria for Hy's law will be provided where Hy's Law criteria will be defined as: AST or ALT > 3× ULN and total bilirubin > 2× ULN at any time during the study.

9 Interim Analyses

Additional details of these analyses will be described in the DSMC Charter.

10 Composite and Net Endpoints

Composite endpoints will first be analyzed as an aggregate and then component-wise, as discussed in the context of modeling clinical events in Section 7.2.3. The primary approach to net outcomes will be to perform them in all randomized patients in the 30-day study period.

Composite/net clinical outcomes include the combination of irreversible harm and efficacy outcomes, including:

- All-cause mortality, thrombotic events, initiation of organ support, rehospitalization adjudicated as to its relatedness to COVID-19.

11 Subgroup and Sensitivity Analyses

The sensitivity analyses will be performed on the ITT, mITT, and Per Protocol populations.

11.1 Primary Efficacy Endpoint Sensitivity Analysis

As a sensitivity analysis, all types of missingness will be addressed via multiple imputation, with a target of 100 multiply imputed datasets. These analyses will be created for findings obtained initially with all available data, with subsequent analysis and pooling performed to evaluate if missing data produces bias in model point estimates or otherwise influences the inferential process.

Multiple imputation will be performed via PROC MI in SAS v9.4. The variables used in the imputation model are the six pre-specified sub-groups of interest, along with any relevant measurements for a subject that will better inform the imputation process (e.g., with missing Day 8 D-dimer, one would include baseline D-dimer in the model). The missing data mechanisms theoretically conferring unbiased estimation (given appropriate covariate selection in the imputation model) using this methodology are MCAR and MAR, although "modern" methodology (MI or FIML [full information maximum likelihood]) also demonstrates superior properties to listwise deletion et al in the MNAR case. Reference Appendices in Section 14.2 which contain descriptions of the Bayesian models, methods for multiple imputation, and methods for power analysis used in this study.

11.2 Primary Efficacy Endpoint Subgroup Analyses

In order to test for subgroup sensitivities, a Bayesian generalized linear mixed effects model will be constructed. This model will evaluate the influence of pre-specified subgroups of interest, D-dimer stratum, and potential site-level heterogeneity in patient outcomes.

A separate set of analyses for each rNAPc2 dosing instance will be constructed, such that the effect on D-dimer of each incremental dose of rNAP2 is estimated. These analyses will therefore be equivalent in

formulation to the Primary Efficacy Analysis as specified above, with estimates of instantaneous treatment effect at Days 3, 5, and 8 for subjects receiving 1, 2 or 3 doses of rNAPc2 or receiving SoC heparin through days 1, 3, and 7, respectively.

11.3 Secondary Efficacy Endpoint Subgroup and Sensitivity Analyses

Secondary efficacy endpoints will be subject to the same sensitivity and subgroup analyses as the primary efficacy endpoints.

The following laboratory parameters will be collected and analyzed by a central laboratory. Change from baseline to both Day 8 or discharge will be summarized for inflammatory biomarkers (IL-6, hsCRP) and coagulation biomarkers (PT, PTT, factor Xa and tissue factor). All data will be listed. Analyses will be by generalized linear mixed effects models as described above for sensitivity and subgroup in the primary efficacy endpoint. For purposes of interpretation of outcomes, IL-6 and hsCRP are considered measures of systemic inflammation, PT, PTT and factor Xa are coagulation parameters, and tissue factor is the biologic target of rNAPc2. In addition, in a subset of participants anti-phospholipid antibodies known to be generated in COVID-19 and inhibited by rNAPc2, are being measured at baseline, Day 8/discharge and on Day 30 at a specialty laboratory.

11.4 Additional Safety Analyses

Decrease in hemoglobin level reported by local laboratories will be monitored to be used as an indicator of potential bleeding risk. As a variable, decrease will be dichotomized such that a decrease of more than 3 g/dL will constitute an event, and any decrease of less than 3g/dL will constitute a non-event. To provide the broadest view of potential bleeding risk, heparin treated subjects will be considered together regardless of prophylactic or therapeutic regimens. The rNAPc2 subjects will be stratified with regard to regimen received. A Cochran Mantel Haenszel Test with stratification will be performed to examine the frequency of events and non-events across the two treatment groups.

12 Changes from Methods Planned in the Protocol

Wilcoxon Signed Rank will not be used in the analysis of the primary efficacy endpoint.

Probability of discharge as assessed by ACTT ≥ 6 by Day 8 and Day 30 post randomization is now a secondary efficacy endpoint.

Composite of death, thrombosis, organ support, or rehospitalization within 30 days is now a secondary efficacy endpoint.

Rehospitalizations and their relatedness to COVID-19 will be adjudicated.

13 References

Gou J, Tamhane AC, Xi D and Rom D. A class of improved hybrid Hochberg-Hommel type step-up multiple test procedures. Biometrika 2014;101(4):899–911. http://www.jstor.org/stable/43304695

14 Appendix

14.1 COVID Severity Definitions

Subgroup analyses according to baseline severity of COVID-19 illness will be performed. The primary categorization will be adapted from the WHO Scale for Clinical Improvement (mild vs. severe) from the COVID-19 Therapeutic Trial Synopsis (see below). In addition, patients may be categorized according to critical illness at baseline (based on use of inotropes or vasopressors, intubation, CPAP/BiPAP, Venturi, and non-rebreather oxygen supplementation) as well as the combination of both approaches, further subgrouping the WHO Hospitalized Severe into those with and without critical illness at baseline.

Severity	**Definition**
Hospitalized Mild	No oxygen therapy
Hospitalized Mild	Oxygen by mask or nasal prongs
Hospitalized Severe	Non-invasive ventilation (inc. high-flow O2)
Hospitalized Severe	Intubation and mechanical ventilation
Hospitalized Severe	Ventilation and additional organ support (vasopressors, renal replacement therapy, ECMO)

14.2 Descriptions of the Models and Methods Used

14.2.1 Bayesian Description

14.2.2 MI Description

14.2.3 Power Description