# Statistical Analysis Plan

| | |
|---|---|
| TRIAL FULL TITLE | A Randomized, Double-Blind, Placebo-Controlled, Multicenter, Phase IIa/IIb Study to Evaluate the Safety, Tolerability, and Efficacy of NP-101 in Treating High- Risk Participants who have Tested Positive for Novel Coronavirus 2019 (BOSS-Covid-19-002) |
| SAP VERSION | 2.3 |
| SAP VERSION DATE | 04/29/2025 |
| TRIAL STATISTICIAN | Mehmet Kocak, Ph.D. |
| PRINCIPAL INVESTIGATORS | Enrique Villa, M.D. |
| SAP AUTHOR | Mehmet Kocak, Ph.D. |

## Contents

# 1 Abbreviations

| | |
|---|---|
| AE | Adverse Event |
| ANCOVA | Analysis of Covariance |
| BMI | Body Mass Index |
| CCRec | Complete Clinical Recovery |
| CCRes | Complete Clinical Resolution |
| CDC | Centers for Disease Control and Prevention |
| COVID-19 | Coronavirus Disease 19 |
| CRF | Case Report Form |
| DLT | Dose Limiting Toxicity |
| DSMB | Data and Safety Monitoring Board |
| eCRF | Electronic Case Report Forms |
| FAS | Full Analysis Set |
| FDA | Food and Drug Administration |
| ICH | International Conference on Harmonization |
| ITT | Intention-To-Treat |
| LSMEANS | Least-squares Means |
| LCVF | Last-Value-Carried-Forward |
| MAR | Missing at Random |
| MCAR | Missing Completely at Random |
| MED | Minimum Effective Dose |
| MI | Multiple Imputation |
| NMAR | Not-Missing at Random |
| PRO | Patient Reported Outcome |
| RT-PCR | Real-time Polymerease Chain Reaction |
| SAE | Serious Adverse Events |
| SAP | Statistical Analysis Plan |
| SARS-CoV | Severe acute respiratory syndrome-coronavirus-2 |
| SCR | Sustained Clinical Recovery |
| SCR-D5 | Day-5 Sustained Clinical Recovery |
| SOC | Standard of Care |
| WHO | World Health Organization |

# 2 Introduction

## 2.1 Purpose

This Statistical Analysis Plan (SAP) describes the planned analyses and reporting of BOSS-002 study, which is a randomized, double-blind, placebo-controlled study of NP-101 in treating high-risk Covid-19 patients on an outpatient basis.

The SAP also describes additional analyses not necessarily identified in this SAP and/or specified in the corresponding protocol which may be performed, should the investigating team or the study Data and Safety Monitoring Board (DSMB) deem it necessary. Any post-hoc, or unplanned analyses not identified in this SAP will be clearly identified in the respective Clinical Study Report (CSR).

The following documents were reviewed and used as guiding references in preparation of this SAP:

- Clinical Research Protocol: BOSS-002 Version 5.1 dated 22 April 2024
- ICH E9: Statistical Principles for Clinical Trials (September 1998)
- E9 Statistical Principles for Clinical Trials Guidance document by Food and Drug Administration (FDA)

The reader of this SAP is encouraged to read the entire clinical protocol for details on the intervention and conduct of this study, and the operational aspects of clinical assessments and their timings in the study.

Table 1. Summary of SAP Changes

| SAP Version/Date | Associated Clinical Protocol | Rationale | Specific Changes |
|---|---|---|---|
| 1.0 (01/25/2023) | BOSS-002 Ver. 4.0 | | SAP original version |
| 2.0 (03/22/2023) | BOSS-002 Ver. 5.0 & 5.1 | Protocol update | • Schedule of Events are updated,<br>• Protocol objectives are updated,<br>• Additional clarifications on the analysis approaches are added.<br>• Definition of adherence<br>• List of scheduled Data Monitoring Reviews |

| | | | |
|---|---|---|---|
| 2.1 (10/30/2024) | BOSS-002 Ver. 5.1 | Clarifications in Analysis Plans | • Definition of Sustained Clinical Recovery is clarified and the abbreviation of SCR-D5 is used throughout to mean Day-5 SCR, the primary efficacy measure.<br>• Clarifications on Phase-IIB study design are provided<br>• Details of the missing data imputation and sensitivity analysis approaches are provided<br>• Opt-in Long-Covid Expansion Portion of the Study exploratory objective was added. |
| 2.2 (01/20/2025) | BOSS-002 Ver. 5.1 | Clarifications in Analysis Plans | • Definition of patient evaluability for the interim analysis is clarified and updated<br>• Timing of the interim analysis is clarified<br>• Plans following the potential decisions from interim analysis are outlined. |

# 3 Study Design

## 3.1 Overview

This is a randomized, double-blind, placebo-controlled phase IIa/IIb study to assess the safety and efficacy of NP-101 versus placebo in the treatment of Covid-19 in high -risk participants in the outpatient setting (Please refer to Sections 5.1 and 5.2 in the protocol document for the inclusion/exclusion criteria). The Phase-IIa component is a dose-escalation study with three dose- specific cohorts and high risk Covid-19 participants will be randomly assigned in a 3:1 ratio to Cohort 1 (3 g), Cohort 2 (4.8 g), and Cohort 3 (6 g). Cohort I and 2 will be evaluated concurrently, followed by Cohort 3 at dose 6g, after which minimum effective dose (MED) will be determined by the parameters set by the SAP and approved by the Data and Safety Monitoring Board (DSMB). The phase-II component is a safety and efficacy study (Cohort 4) where high risk Covid- 19 participants will be randomly assigned either to take the MED of NP-101 or to take Placebo in a 1:1 ratio. The participants will receive up to 14 days of dosing.

In summary, the study will have the following four cohorts:

Cohort-1: Participants randomly assigned in a 3:1 ratio to dose 3g or placebo

Cohort-2: Participants randomly assigned in a 3:1 ratio to dose 4.8g or placebo

Cohort-3: Participants randomly assigned in a 3:1 ratio to dose 6g or placebo

Cohort-4: Participants randomly assigned in a 1:1 ratio to the MED or Placebo

### 3.2 Sample size

Phase-IIa:

Sample size for the dose-escalation phase of the trial will be 16 safety evaluable participants for each dose-specific cohort of 3 g, 4.8 g and 6 g total daily doses, where 20 participants will be randomized in a 3:1 ratio (active:placebo) resulting in 15 participants assigned to the NP-101 arm and 5 to the Placebo arm such that at least 16 safety evaluable participants (12 NP-101 and 4 Placebo) can be obtained based on an anticipated attrition rate of 20-25%.

Phase-IIb:

Phase IIb power calculation is based only on the second primary objective of the trial, Day-5 Sustained Clinical Recovery (SCR-D5). Based on the SCR-D5 rates estimated for high-risk participants treated in the prior BOSS-001 protocol Phase-II trial of TQ Formula (NP-101), a sample size of 200 participants randomized equally during the phase IIb part of the study to NP-101 and Placebo Arms will yield at least 80% power with a total of 2.5% final Type-1 error rate to detect a 20% improvement in SCR-D5 rate in high-risk Covid-19 participants treated at the MED of NP-101 compared to those treated with Placebo. This sample size is based on a group sequential design with one interim analysis when 100 evaluable participants are enrolled and completed their data for the primary objectives, using O'Brien Fleming Spending function with only 0.311% of the total 2.5% Type-1 error spent at the time of the interim analysis.

**The purpose of the interim analysis is** to determine whether or not to stop the trial early for any of the following: (1) if NP-101 does not provide sufficient evidence of activity or (2) has an unacceptable safety profile, or (3) if NP-101 has higher efficacy than anticipated based on the literature and the earlier Phase-II study (BOSS-001). Conducting the interim analysis is important because dependent upon the results, (1) we will not expose more patients to non-efficacious therapy, (2) we will not expose more patients to undue risks, or (3) in the event of higher efficacy than anticipated, NP-101 can potentially be made available to wider populations of high-risk outpatient Covid-19 patients in a more timely manner or (4) the trial might continue to n=308.

This prespecified interim analysis for early efficacy or futility will be carried out once 50% of the target enrollment of evaluable patients (n=100, at least 50 in each arm) in phase IIb have been followed through Day 28. The actual sample size for the interim analysis may be greater than 100 due to the random allocation of patients into the two arms and the fact that each arm needs at least 50 participants for the interim analysis according to the group sequential design. This analysis will be based on the primary efficacy and safety endpoints according to

prospectively set-stopping boundaries for efficacy and futility that will be reviewed by the Data and Safety Monitoring Board (DSMB).

Considering a conservative attrition expectation of around 20% to account for patient dropouts and rare missing data for the primary endpoint, the planned total accrual for the primary efficacy objective is up to 248 participants, equally split as up to 124 participants in each arm. With the same attrition expectation, the interim analysis will be carried out when at least 50 patients who were randomized and consumed at least one dose of the assigned drug within each arm (up to a total of 124 patients, up to 62 patients in each arm).

Definition of Day-5 Sustained Clinical Recovery (SCR-D5): the continuous alleviation or resolution of all COVID-19 PRO Symptom Survey scores to ≤1 ("none" or "mild"), starting on or before Day 5 after randomization and remaining so for at least 3 consecutive days without subsequent relapse, defined as not progressing to "moderate" or "severe" for at least two consecutive days or without hospitalization or death up to Day 28 from randomization.

Definition of Evaluability for Primary Objectives: This double-blind, placebo-controlled, two-arm clinical trial is designed to enroll up to 248 patients during the phase IIB part of the study, equally randomized between the two arms of the study. An evaluable patient is defined as a patient who has received at least one dose of the intervention drug or placebo and has been followed up to Day 28 from Randomization. Although all patients regardless of whether or not they receive the assigned medicine will be included in the primary analyses based on the intent-to- treat principle, evaluability- and adherence-based analyses will also be carried out as part of the sensitivity analyses.

Here are some endpoint combinations justifying our sample size for this study:

**Table 9.2.1.** Group Sequential Design scenarios with varying effect sizes

| Response Rate in Placebo | Response Rate in NP-101 | Minimum Sample Size Required | Sample Size with 20% Attrition | Sample size for the Interim Analysis | Stop Early for Futility (Z-Test) | Stop Early for Efficacy (Z-Test) | Final Decision Boundary (Z-Test) |
|---|---|---|---|---|---|---|---|
| 0.20 | 0.40 | 168 | 208 | 104 | 0.688 | 2.737 | 1.935 |
| 0.25 | 0.45 | 182 | 224 | 112 | 0.71 | 2.722 | 1.935 |
| 0.30 | 0.50 | 192 | 240 | 120 | 0.693 | 2.736 | 1.935 |
| 0.35 | 0.55 | 198 | 248 | 124 | 0.712 | 2.723 | 1.935 |
| 0.40 | 0.60 | 200 | 248 | 124 | 0.695 | 2.736 | 1.935 |
| 0.45 | 0.65 | 198 | 248 | 124 | 0.712 | 2.723 | 1.935 |

The early stopping probabilities for a set of selected response combinations are provided below. At the time of the interim analysis, we will take a conservative approach and use the minimum of the early stopping for futility boundaries listed among the scenarios above (i.e., 0.688) to stop early for futility, and will use the maximum of the Early Stopping for Efficacy boundaries listed among the scenarios above (i.e., 2.737) to stop early for efficacy. At the targeted effect size, the

early stopping probability for efficacy is 24%, for futility is 9%.

**Table 9.2.2.** Early Stopping probabilities with varying effect sizes

| Response Rate Difference (Active-Placebo) | Early Stopping Probability for Efficacy | Early Stopping Probability for Futility | Total Early Stopping Probability |
|---|---|---|---|
| 0 | 0.003 | 0.757 | 0.760 |
| 0.1 | 0.043 | 0.374 | 0.417 |
| 0.2 | 0.243 | 0.090 | 0.333 |
| 0.3 | 0.626 | 0.009 | 0.635 |

Participants will be assigned to one of the two study arms in a 1:1 ratio based on a block-randomization scheme. No stratifications for randomization will be used. A independent unblinded statistician will carry out the interim analysis as the study biostatistician will remain blinded to the study data until after the final database lock.

### 3.3 Randomization, Blinding, and Patient Selection

During Phase-IIa, high-risk Covid-19 participants will be randomized to three dose-specific cohorts in a 3:1 ratio to receive NP-101 or Placebo respectively for 14 consecutive days:

- Cohort-1: Study participants will receive 3 g total daily dose (three 600 mg capsules in AM, two 600 mg capsules in PM ),
- Cohort-2: Study participants will receive 4.8 g total daily dose (4 capsules BID)
- Cohort-3: Study participants will receive 6 g total daily dose (5 capsules BID)

Study participants will receive either NP-101 at the minimum effective dose (MED) based on the safety and early efficacy assessments during Phase-IIa by the DSMB or receive placebo for 14 consecutive days.

The study is double blinded; all participants, investigators, and study personnel involved in the conduct of the study, including the primary biostatistician, will be blinded to treatment assignment. The third party data management team and the head of regulatory affairs remain unblinded, along with the independent, unblinded statistician forf the study who will generate analysis reports and provide them to the DSMB for review and decision making in accordance with the SAP after Phase-IIa is completed and at the time of the interim analysis of Phase-IIb. The final analyses will be conducted after the data lock for both Phase IIa and IIb. The opt-in opt-out long-covid expansion data will start after the interim analysis and the data will be housed in a separate database and will be analyzed at a later time.

Sponsor and PIs will remain blinded to study treatment. Randomization codes will be assigned through the electronic system.

Participants will complete the PRO Symptom Survey, a daily questionnaire on Covid-19 symptoms, which will be used to determine the primary endpoint for the study. For this reason, it

is important that participants remain blinded to treatment allocation.

**Unblinding**

If unblinding is required, the principal investigator can unblind through the electronic system, 24 hours, 7 days a week. Treatment unblinding is discouraged if knowledge of the treatment assignment will not materially change the planned management of a medical emergency.

Unblinding should be discussed in advance with the medical monitor if possible. If the investigator is not able to discuss treatment unblinding in advance, then he/she must notify the medical monitor as soon as possible about the unblinding incident without revealing the participant's treatment assignment. The reasons, dates, times, and the affected participants of all these unblinding incidences will be recorded and captured in the study database.

## 4 Study Calendar and Data Collection

The schedule of the study visits and data to be collected are shown below.

Schedule of Activities (SOA) (Quoted from the protocol, Section 1.3)

| **Procedures** | **Screen Visit** | **Baseline[d] Visit** | **Study Visit 2[f]** | **Study Visit 3[g]** | **Study Visit 4[f]** | **Study Visit 5[g]** | **Study Visit 6[f]** | **Final Study Visit[f]** | **Data Collection via APP only** | **Safety Follow Up Phone Call** |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Day -3 to -1** | **Visit 1, Day 1** | **Day 4 +/-1 day** | **Day 7 +/- 1 day** | **Day 10 +/-1 day** | **Day 14 +/- 1 day** | **Day 21 +/- 1 day** | **Day 30 +/- 1 day** | **Days 37 +/- 1 day** | **Day 45 +/- 1 day** |
| Informed consent | X | | | | | | | | | |
| Demographics | X | | | | | | | | | |
| Medical history, including Covid vaccination (type and dates) | X | | | | | | | | | |
| Inclusion/Exclusion Criteria Evaluation | X | X | | | | | | | | |
| Nasal swab: RT-PCR sample for quantitative viral load analysis | | X[c] | | X | | X | | | | |
| Blood Collection (including sample for immune monitoring) | | X | | X | | X | | | | |
| Physical exam (complete at screening and baseline, focused on day 7 and day 14) | X | X | | X | | X | | | | |
| Vital signs | X | X | | X | | X | | | | |
| Pulse Oximetry | X | X | | X | | X | | | | |
| Height | X | | | | | | | | | |
| Weight | X | | | | | | | | | |
| Hematology | X | | X[e] | X | X[e] | X | X[e] | X[e] | | |
| Serum chemistry [a] | X | | X[e] | X | X[e] | X | X[e] | X[e] | | |
| Best Supportive Care | X | X | X | X | X | X | X | X | | |
| Pregnancy test [b] | X | | | | | X | | | | |
| Randomization and dispense IP, first dose taken at visit | | X | | | | | | | | |
| Drug Accountability | | X | | X | | X | X if not returned at Day 14 | | | |
| Concomitant medication Review, including medications for Covid symptoms | X | X | | X ..... X | | | | | | |
| Adverse event review/evaluation | X | X | X | X | X | X | X | X | | X |
| Long-Covid Symptoms | | X[g] | | | | X[g] | | X[g] | | X[g] |
| Patient Reported Outcomes (PRO Symptom Survey and Conjunctivitis Questionnaire* Completed daily by patient | X+ | X | X | X | X | X | x | X | X | |
| Complete Case Report Forms (CRFs) | X | | X ..... X | | | | | | | X |

a: Albumin, alkaline phosphatase, total bilirubin, direct bilirubin, bicarbonate, BUN, calcium, chloride, creatinine, glucose, LDH, phosphorus, potassium, total protein, AST, ALT, sodium, CBC/diff, (Optional: LDH, CRP, prothrombin time, activated partial prothrombin time, and D-dimer)
b: Serum or urine pregnancy test (women of childbearing potential).
c: Covid-19 quantitative viral load
d: Screening and Baseline can be combined
e: If clinically indicated at the discretion of the investigator
f: Visits at Day 4 (visit 2), Day 10 (visit 4), and Day 21 and Day 30 (final visit) will be telehealth and can be changed to in-person at the discretion of the investigator g: In the event of a complete lockdown due to COVID, or if patients can't make it to clinic, e.g. due to sickness, visits on Day 7(Visit 3) and Day 14 (Visit 5) could also be replaced by home health visit for sample collection and telehealth visit for clinical evaluation.

g. Long-Covid symptoms will be collected at Baseline, Day 14, Day 30 and Day 45 through the app.
*Administered everyday
+ PRO Symptom Survey at Screen will be completed on site to determine if patient meets inclusion criterion

## 5 Objectives and Endpoints

**Overall Hypothesis:** NP-101 is safe, tolerable, and superior to placebo as assessed by the percentage of participants who have reached sustained clinical recovery on Day 5 after randomization

**Primary Safety and Efficacy Objectives**

| Objectives/Purpose | Endpoints/Outcome Measures | Justification for Endpoints |
|---|---|---|
| **Primary Safety Objective (for phase IIa)** | **Primary Safety Endpoint (for phase IIa)** | |
| 1. To identify the Minimum Effective Dose (MED) of NP-101 in high-risk Covid-19 patients treated on an outpatient basis. | 1. Number of Dose Limiting Toxicities (DLT) in NP-101 arm at each dose level compared to placebo and the safety threshold. | Identifying a MED is important for describe the safety profile of NP-101 and it will be used in the Phase-IIb portion of the study. |
| **Primary Safety Objectives (for phases IIa and IIb)** | **Primary Safety Endpoints (for phases IIa and IIb)** | |
| 1. To evaluate the safety and tolerability of NP-101 compared to placebo. | 1. Number of overall adverse events, related adverse reactions, adverse events leading to discontinuation of study drug and hospitalizations, or deaths reported in participants taking the MED of NP-101 versus participants taking placebo. All AEs/SAEs will be captured throughout the study as per schedule of activities. As part of the primary safety assessment, treatment discontinuations will be compared between participants taking the MED of NP-101 versus participants taking placebo | Comparing the number of AEs and SAEs and any relationship to IP is important for determining the safety profile of NP-101. We hypothesize that NP-101 will not have significantly higher frequency of AEs, SAEs and treatment discontinuation incidences. All AEs/SAEs and treatment discontinuation will be captured throughout the study as per schedule of activities. |

**Primary Efficacy Objective/IIb**

| Primary Efficacy Objectives (Only Phase IIb MED) | Primary Efficacy Endpoints (Only Phase IIb MED) | Justification for Endpoints |
|---|---|---|
| 2. To compare the rate of Sustained Clinical Recovery on Day 5 (SCR-D5) after randomization in participants taking the MED of NP-101 as compared to placebo.<br><br>* Day-5 **Sustained Clinical Recovery (SCR-D5)** is defined as the continuous alleviation or resolution of all COVID-19 PRO Symptom Survey scores to ≤1 ("none" or "mild"), starting on or before Day 5 after randomization and remaining so for at least 3 consecutive days and without subsequent relapse, defined as symptoms progressing to "moderate" or "severe" for at least two consecutive days, hospitalization, or death until Day-28 from randomization. | 2. Measurement of the difference of SCR-D5 rates in participants taking the MED of NP-101 as compared to placebo. | To assess the efficacy of NP-101 in treating COVID- 19 high risk participants treated in an outpatient setting in terms of percentage of participants recovering from moderate to severe symptoms by Day 5 compared to placebo. The choice of Day 5 Sustained Clinical Recovery was mainly based on our previous Phase-II study of NP-101 (BOSS-001), that showed that the greatest difference was seen around Day 5 from randomization and also based on the CDC's 5-day isolation requirement. |

### 5.1 Secondary Objectives

| Secondary Objectives | Secondary Endpoints | Justification for Endpoints |
|---|---|---|
| 1. To compare the change in viral load and cycle threshold profiles from Randomization to Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo. | 1. Measurement of the difference of the average viral load and cycle threshold measures from Randomization to Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo | Faster decline in viral load and faster incline in cycle threshold are hypothesized to be indications for faster recovery from the illness and less infectivity [109, 110] |
| 2. To compare the percentage of RT-PCR negative/undetectable (i.e., viral clearance) on Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo. | 2. Measurement of the difference in the percentage of negative/undetectable RT-PCR (i.e., viral clearance) on Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo. | Treatment with the MED of NP-101 given orally on outpatient basis is hypothesized to result in higher percentages of viral clearance by RT-PCR. Faster decline of viral load is hypothesized to lead to faster recovery from the illness and |

| Secondary Objectives | Secondary Endpoints | Justification for Endpoints |
|---|---|---|
| | | less infectivity on Day 5 and Day 14 [109, 110] |
| 3. To compare the distribution of days to Sustained Clinical Recovery (SCR), days to Complete Clinical Recovery (SCRec), and days to Complete Clinical Resolution (SCRes), between participants with taking the MED of NP-101 as compared to placebo.<br>***The time to sustained clinical recovery (SCR) is defined as*** the number of days from randomization to the first of 3 consecutive days of alleviation or resolution of all symptoms to Mild or None.<br>**First Day of Complete Clinical Recovery (CCRec) is defined as** the first day from randomization when all symptoms scored as moderate or severe at study entry are scored as mild or absent AND all symptoms scored mild or absent at study entry are scored as absent and remain so afterwards until Day 30.<br>**First Day of Complete Clinical Resolution (CCRes) is defined as** the first day from randomization when all symptoms are scored as absent and remain so afterwards until Day 30. | 3. Estimate of the difference in days to SCR, SCRec, and SCRes survivorship functions in participants taking the MED of NP-101 as compared to placebo. The time to sustained clinical recovery is defined as the number of days from randomization to the first of 3 consecutive days of alleviation or resolution of all symptoms to Mild or None. | We hypothesize that the distribution of days to Sustained Clinical Recovery, days to Complete Clinical Recovery, and days to Complete Clinical Resolution will be significantly shorter than that of placebo. |
| 4. To compare the change in effector immune cells, specifically CD4+ and CD8+ T lymphocytes, from Randomization to Day 7 and Day 14 in in participants taking the MED of NP-101 as compared to placebo. | 4. Measurement of the difference of the average level of effector immune cells, specifically CD4+ and CD8 T-cells, from Randomization to Day 7 and 14 in participants taking the MED of NP-101 versus those taking | We hypothesize that participants treated with NP-101 will have significantly higher CD4+ and CD8+ T-cell counts on average compared to placebo, which will be a sign of stronger immunity build- up against the infection. |

| Secondary Objectives | Secondary Endpoints | Justification for Endpoints |
|---|---|---|
| | placebo. | |
| 5. To compare the rate of disease progression measured by the worsening by the Covid-19 symptoms from “none” or “mild” to “moderate” or “severe” and remaining at that level for at least two days before alleviation or resolution, or the experience of hospitalization or death due to Covid-19 symptom worsening in participants taking the MED of NP-101 versus placebo. | 5. Measurement of the difference in the percentage of disease progression measured by the worsening by the Covid-19 symptoms from “none” or “mild” to “moderate” or “severe” and remaining at that level for at least two days before alleviation or resolution, or the experience of hospitalization or death due to Covid-19 symptom worsening in participants taking the MED of NP-101 versus placebo. | We hypothesize that treatment with NP-101 will lower the likelihood of disease progression and hospitalization. |
| 6. To compare the rate of Long-Covid symptoms in participants taking the MED of NP-101 versus placebo.<br>A patient will be classified as having experienced symptoms indicative of Long-Covid if they reported any new or ongoing mild, moderate, or severe Long-Covid symptoms captured on the daily symptom questionnaire, on any of the following days:<br>Day 30 ± 2 days, Day 37 ± 2 days, or Day 45 ± 2 days”. | 6. Measurement of the difference in the percentage of Long-Covid symptoms in participants taking the MED of NP-101 versus placebo. | We hypothesize that participants treated with NP-101 will experience lower rate of Long-Covid symptoms compared to placebo. |

## 5.2 Exploratory objectives

| Exploratory Objectives (Phase IIa only) | Exploratory Endpoints (Phase IIa only) | Justification for Endpoints |
|---|---|---|
| 1. To assess the early indication of efficacy and its dose-dependency in order to guide the determination of the Minimum Effective Dose (MED) of NP-101 | 1. Estimate of the Day-5 Sustained Clinical Recovery (SCR-D5) by dose, the average of the CD4+ and CD8+ expression change from baseline to Day-14, and Long-Covid rate on the final study visit (Day-30) in a blinded fashion. | As the study aims to identify a Minimum Effective Dose (MED), these early efficacy indicators need to be estimated by dose level in a blinded fashion and shared with the DSMB. If DSMB requests, an unblinded statistician will produce the early efficacy indicators for NP-101 arm alone as well |

| Exploratory Objectives (Phase IIb only) | Exploratory Endpoints (Phase IIb only) | Justification for Endpoints |
|---|---|---|
| 1. To compare the change in inflammatory cytokines and coagulation factors from Randomization to Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo. | 1. Measurement of the difference between the average level of inflammatory cytokines and coagulation factors, from Randomization to Day 7 and 14 in participants taking the MED of NP-101 versus those taking placebo. | The inflammatory cytokines and immunological markers are of importance because of their correlation with disease severity in COVID- 19. |
| 2. To construct a risk prediction model for Sustained Clinical Recovery based on known risk factors as well as other patient and disease characteristics. | 2. Significance assessment of each risk factor individually in predicting the Sustained Clinical Recovery and its adjusted significance controlling for the effects of other risk factors in the model | Identifying which risk factor is more clinically and statistically important may help guide in generating closer follow-up of such participants. |
| 3. To compare the proportions of patients developing any feature of conjunctivitis between NP-101 and placebo and to compare the time to conjunctivitis and time to resolution among those who developed conjunctivitis. | 3. Measurement of the difference in the percentage of conjunctivitis symptoms in participants taking the MED of NP-101 versus placebo and the difference in time to onset of conjunctivitis and time to resolution of it. | We hypothesize that participants treated with NP-101 will have significantly lower rate of conjunctivitis symptoms (which are evolving with new COVID-19 variants) compared to placebo, and longer time to onset and shorter time to resolution of such symptoms with treatment with NP-101 compared to placebo |

| | | |
|---|---|---|
| 4. To compare the rate of Long-Covid symptoms per the questionnaire in participants taking the MED of NP-101 versus placebo in **the Opt-in Long-Covid Expansion Portion of the Study** | 4. Measurement of the difference in the percentage of Long-Covid symptoms in participants taking the MED of NP-101 versus placebo. | We hypothesize that participants treated with NP-101 will experience lower rate of Long-Covid symptoms compared to placebo. We will do our due diligence in following the CDC guidelines/definitions of long-covid at the time of the analyses. |

# 6 Statistical Methods

## 6.1 Study Assessments and Procedures

### Covid-19 Quantitative Viral Load (RT-PCR)

Participants will have an Real-time Polymerease Chain Reaction (RT-PCR) sample taken at baseline (day 1 of randomization), day 7 and day 14 to measure quantitative viral load and viral load clearance. These samples will be sent to a central laboratory for analysis.

### Patient Reported Outcomes (PRO)

Participants will complete the Patient Reported Outcomes (PRO) Symptom Survey (Appendix 2) electronically daily for the entirety of the study, from randomization to final study visit. The responses will be used to measure severity and change in Covid-19 symptoms throughout the study.

**Day 0:** Most current symptom assessment before randomization. In most cases, it is expected to be on the day of randomization, right before randomization to assess the eligibility

**Day 1:** Symptom assessment within the first day after receiving the first dose of the medication.

**Day X:** Symptom assessment after X days after receiving the first dose of the medication. Thus, **Day 5** will be the assessment on Day 5 from the therapy start date, which is ideally the day of randomization

### Blood samples Collection (including Blood sample for immune monitoring)

Participants will have blood collected for immune monitoring. This testing will explore pharmacokinetics and inflammatory cytokines, coagulation factors and effector immune cells in samples. These samples will be sent to a central laboratory for analysis.

## 6.2 General Statistical Methods

Categorical variables will be presented as frequencies and percentages with column and row percentages in 2x2 or RxC tables as appropriate. Association between two nominal variables will be investigated using Fisher's Exact test. Cochran Armitage trend-test may be

included among the possible tests to investigate the trend between one ordinal factor and one binary factor. Association between two ordinal variables may be investigated through Non-zero Correlation test.

Continuous variables will be presented as mean and standard deviation as well as five-number summary (namely, minimum, first quartile (Q1), median, third quartile (Q3), and maximum). Range and inter- quartile range may also be explicitly provided as needed. Distribution of a continuous variable among the levels of a binary or polynomial factor will be investigated graphically through boxplots and through Wilcoxon-Mann-Whitney test and/or Kruskal-Wallis tests. Association between two continuous variables will be described through scatter plots and Pearson's and Spearman's correlation coefficients as appropriate.
When needed, normality assumption of a given continuous variable will be tested using Shapiro-Wilk Normality test.

In all comparisons and tests for the secondary and exploratory aims, a two-sided p-value will be used and a Type-1 error rate of 0.05 will be used to indicate significance. Regarding addressing the multiplicity correction, as there is a single estimand for the primary efficacy objective, no multiplicity correction is needed, and conclusions will be based on the decision boundaries provided as part of the statistical design presented above in Section 3.2 Sample Size. Similarly, no multiplicity correction will be employed for statistical tests addressing the secondary and exploratory objectives as the results from these objectives must be considered within the hypotheses generating context for future trials.

The primary statistical hypotheses are:

1. **Dose Escalation**: Considering the nature of NP-101, Dose Level 6g is expected to be safe and to be declared as the minimum effective dose (MED) for the Phase IIb trial.
2. **Safety**: Safety and Tolerability of treatment with NP-101 at the MED given orally on an outpatient basis (study drug) will be comparable with taking placebo in participants with COVID-19 infection treated in the outpatient setting.
3. **Efficacy**: Treatment at the MED given orally on outpatient basis (study drug) will statistically significantly increase the Day 5 Sustained Clinical Recovery rate compared to placebo in high-risk participants with COVID-19 infection treated in the outpatient setting.

**Day-5 Sustained Clinical Recovery (SCR-D5)** is defined as the continuous alleviation or resolution of all COVID-19 PRO Symptom Survey scores to ≤1 ("none" or "mild"), starting on or before Day 5 after randomization and remaining so for at least 3 consecutive days and without subsequent relapse, defined as progressing to "moderate" or "severe" for at least two

consecutive days, hospitalization, or death until Day-28 from randomization.

The secondary statistical hypotheses are as follows:

1. Study drug will significantly differently reduce quantitative viral load from Randomization to Day 7 and Day 14 compared to placebo in participants with COVID-19 infection treated in the outpatient setting.
2. Study drug will significantly differently increase the percentage of negative/undetectable RT-PCR viral load at Day 7 and Day 14 in participants taking the MED of NP-101 versus participants taking placebo in participants with COVID-19 infection
3. Study drug will significantly differently reduce the days to Sustained Clinical Recovery, days to Complete Clinical Recovery, and days to Complete Clinical Resolution compared to placebo in participants with COVID-19 infection
4. Study drug will significantly differently increase the CD4+ and CD8 T-cell concentrations on Day 7 and/or Day 14 compared to placebo in participants with COVID-19 infection
5. Study drug will significantly reduce the disease progression defined as the worsening of the Covid-19 symptoms on Day 5 from “none” or “mild” to “moderate” or “severe” and remaining at that level for at least two days before alleviation or resolution, or the experience of hospitalization or death due to Covid-19 symptom worsening in participants taking the MED of NP-101 versus placebo
6. Study drug will reduce the rate of Long-Covid symptoms compared to placebo.

- Primary Safety and Efficacy Endpoint(s) (Phase IIa and IIb):
    1. **Dose Finding (Phase IIa)**: Number of Dose Limiting Toxicities (DLT) in NP-101 arm at each dose level compared to placebo and the safety threshold.
    2. **Safety:** Number of overall adverse events, related adverse reactions, adverse events leading to discontinuation of study drug and hospitalizations, or deaths reported in participants taking the MED of NP-101 versus participants taking placebo. All AEs/SAEs will be captured throughout the study as per schedule of activities. As part of the primary safety assessment, treatment discontinuations will be compared between participants taking the MED of NP-101 versus participants taking placebo
    3. **Efficacy:** Measurement of the difference of Sustained Clinical Recovery rates on Day 5 in participants taking the MED of NP-101 as compared to placebo..

- Secondary Efficacy Endpoint(s):
    1. Measurement of the difference of the average viral load and cycle threshold measures from Randomization to Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo.

2. Measurement of the difference in the percentage of negative/undetectable RT-PCR (i.e., viral clearance) on Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo.
3. Estimate of the difference in days to Sustained Clinical Recovery, days to Complete Clinical Recovery, and days to Complete Clinical Resolution survivorship functions in participants taking the MED of NP-101 as compared to placebo. The time to sustained clinical recovery is defined as the number of days from randomization to the first of 3 consecutive days of alleviation or resolution of all symptoms to Mild or None.
4. Measurement of the difference of the average level of effector immune cells, specifically CD4+ and CD8 T-cells, from Randomization to Day 7 and Day 14 in participants with the MED of NP- 101 versus those taking placebo
5. Measurement of the difference in the percentage of disease progression on Day 5 measured by the worsening of the Covid-19 symptoms from “none” or “mild” to “moderate” or “severe” and remaining at that level for at least two days before alleviation or resolution, or the experience of hospitalization or death due to Covid-19 symptom worsening in participants taking the MED of NP-101 versus placebo
6. Measurement of the difference in the percentage of patients experiencing Long-Covid symptoms in participants taking the MED of NP-101 versus placebo.

### 6.3 Data and Safety Monitoring Board (DSMB) Reviews

The first DSMB review will take place after all patients in Dose 3.0 g daily (Cohort 1) and Dose 4.8 g daily (Cohort 2) complete their treatment and follow-up periods to assess the available safety data to decide whether or not it is safe to start enrollment at the highest dose level of 6 g daily (Cohort 3).

If DSMB recommends dose escalation to Dose 6 g daily, then the second DSMB review will take place after all patients in Cohort 3 complete their treatment and follow-up periods to assess the available safety as well as the early indication efficacy data through sustained clinical recovery, Long-Covid rate, and immune response evaluation to identify the Minimum Effective Dose (MED) of NP-101 in high-risk Covid-19 patients treated on an outpatient basis and to recommend the start of the Phase-IIb component of the trial.

The third DSMB review will take place at the time of the only interim analysis designed as part of the study after at least 100 evaluable patients are treated at the MED and their follow-up periods are completed to assess the available data and recommend whether or not the

enrollment continue based on whether or not the efficacy and futility boundaries are crossed by the accumulated efficacy data.

The DSMB may have other reviews and assessments according to the DSMB charter to ensure patient safety.

### 6.4 Analysis Sets

The following are the main analysis datasets defined in Section 9.3 of the study protocol:

- Full Analysis Set (FAS): All randomized participants regardless of whether they receive any study intervention. This corresponds to Intent-to-treat (ITT) analysis without any modification.
- Full Analysis Set for the subset of randomized participants whose COVID-19 symptoms started within 3 days of randomization (sub-FAS): All participants who were randomized within 3 days of Covid-19 symptom onset regardless of whether they receive any study intervention.
- Safety Analysis (SA) Dataset: All randomized participants who receive at least one dose of study intervention.
- Safety Analysis with Covid-19 related concomitant medications (SA-ConMed) Analysis Dataset: These safety analysis datasets may have multiple versions depending on the combination of Covid-19 related concomitant medications that the patients used along with NP-101 from enrollment date to off-study date. Due to randomization, concomitant medication types and dosages are expected to be distributed similarly between the two study arms.
- Modified Intent-to-Treat-1(mITT1): All randomized participants regardless of whether they receive any study intervention, who had sufficient self-reported symptom data so that the sustained clinical recovery on Day 5 can be assessed.
- Modified Intent-to-Treat-2(mITT2): All participants who were randomized within 3 days of Covid-19 symptom onset regardless of whether they receive any study intervention, who had sufficient self-reported symptom data so that the sustained clinical recovery on Day 5 can be assessed.
- Per-Protocol (PP) Analysis Dataset: The Per-Protocol Analysis Set (PP) is defined as all participants in the FAS who completed all visits up to and including Day 14, and were compliant with study medication at ≥ 85% adherence and do not have any major protocol violation. Major protocol violations will be identified prior to breaking the blind.

The level of adherence to the protocol therapy will be based on the final pill-count and the reported missed doses. Every participant have a 14 days of treatment (28 dosing-times considering the morning and evening doses). A study participant will be considered “at least 85% adherent” if s/he receives his/her doses in at least 24 of these 28 dosing-times when reported missed-doses are considered and confirmed by the final pill-count.

- Per-Protocol (PP) Analysis with Covid-19 related concomitant medicines (PP-ConMed) Analysis Datasets: These analysis datasets may have multiple versions depending on the combination of Covid-19 related concomitant medications that the patients used along with NP-101 from enrollment date to off-study date.

## 6.5 Statistical Analyses

### 6.5.1 Demographic and Baseline Characteristics

Demographic and baseline characteristics data will be summarized with descriptive statistics. Continuous variables will be summarized with n, mean, standard deviation, minimum, maximum and median. Frequency counts and percentage of subjects within each category will be provided for categorical data.

### 6.5.2 Primary Analyses

**Primary Analyses for Phase-IIa:**

Primary Estimand for Phase-IIa is the number of Dose Limiting Toxicities (DLT) in NP-101 arm compared to Placebo and the safety threshold. Based on the number of DLTs observed and early indication of efficacy data, a Minimum Effective Dose (MED) of NP-101 will be determined.

Phase-IIa portion of this study is a randomized dose-escalation trial, where the enrollment for the initial two dose levels, namely, 3g and 4.8g daily dose, will be started simultaneously based on safety data in literature as well as our previous randomized Phase-II study utilizing 3g daily dose of NP-101.

**Definition of Dose Limiting Toxicity (DLT):** A Dose Limiting Toxicity (DLT) is defined as any Grade 3+ adverse events that occurred within the first 21 days from randomization and are found to be attributable to the study regimen and not resolved through symptomatic treatments within 48 hours.

**DLT Observation Period:** In this study, DLT observation period will be the initial 21 days from randomization.

**Evaluability for Safety Assessment:** Any patient who received at least one dose of the assigned study medication will be considered evaluable for the NP-101 safety assessment.

At each dose level, up to 20 participants will be randomly assigned in a 3:1 ratio to NP-101 or Placebo in a double-blind fashion, resulting in up to 15 participants being treated with NP-101 and up to 5 patients being treated with Placebo. Number of patients with DLTs will be counted dynamically without unblinding the random assignments on a continuous basis. At the same time, the study database will be also counting the number of evaluable patients with DLTs in NP-101 arm ($N_{DLT,3g,\ NP101}$, $N_{DLT,4.8g,\ NP101}$, $N_{DLT,6g,\ NP101}$, respectively for dose levels 3g, 4.8g. and 6g, which are back-end variables only visible to the study monitor) dynamically and independently at each dose level being investigated.

A dose level with at most two DLTs will be considered safe, and a dose-escalation or MED decision will be recommended by the study Data and Safety Monitoring Board (DSMB) based on the accumulated safety, Long-Covid, and early indication of efficacy data.

- If 3 g is found to be NOT safe, dose de-escalation to 2.4 g will occur.
- If Dose Level 3g is found to be safe and Dose Level 4.8g is found to be NOT safe, then Dose Level 3g will be considered to be MED and enrollment in Phase-IIb trial will be initiated automatically at Dose Level 3g.
- If both Dose Levels 3g and 4.8g are found to be safe, enrollment to Dose Level 6g will be initiated automatically.
- If both Dose Level 3g and 4.8g are found to be safe while Dose Level 6g was found to be NOT safe, then Dose Level 4.8g will be considered to be MED and enrollment in Phase-IIb trial will be initiated automatically at Dose Level 4.8g.
- If all dose levels are found to be safe, then Dose Level 6.0 will be considered to be MED and enrollment in Phase-IIb trial will be initiated automatically at Dose Level 6g.
- In the very unlikely event that Dose 4.8g is found to be safe while Dose 3g is found to be UNSAFE, the DSMB will more closely investigate the attribution of these DLTs to NP-101 considering detailed patient characteristics; this may require protocol amendment to add further safety measures to address these safety concerns. The DSMB will provide their final recommendation of dose escalation, de-scalation, or declaring a dose-level to be the MED to the study team based on their assessment and this decision will be manually coded in the

study database to manage the future enrollment.

The above DLT monitoring rules will be employed regardless of patient evaluability for the primary endpoints; however, to be included in the safety monitoring, it will be required that the patient received at least one dose of the study regimen and attribution of a given adverse event to the study regiment is clearly established as 'at least possibly attributable' to the study regimen. In this attribution assessment, the treating physician will consider the known and expected adverse events of all concomitant medications to truly establish the attribution of a given AE to NP-101.

The study protocol database will be built on a flexible data-capture and decision-making platform. The following detailed decision-making algorithm will be utilized to generate seamless DLT monitoring and safety summary reports for the review of the medical monitor (unblinded to the random assignments) as well as the study team without unblinding.

As part of the Phase-IIa safety assessment, frequencies of overall adverse events, related adverse reactions, adverse events leading to discontinuation of study drug and hospitalizations, or deaths reported in participants taking the Selected Dose of NP-101 versus participants taking placebo will be tabulated and compared for each dose level investigated.

**Decision Algorithm for Phase-IIa Dose Escalation to Guide the Database Systems:**

Dose Limiting Toxicity (DLT) observation period is defined to be the first 21 days from randomization. For each patient, the database will be capturing two critical pieces of information: one indicating whether or not the patient is evaluable for Dose Limiting Toxicity (DLT) evaluation, and the second indicating whether or not the patient had any DLTs.

The database will be also counting the number of evaluable patients with DLTs in the NP-101 arm ($N_{DLT,3g,\ NP101}$, $N_{DLT,4.8g,\ NP101}$, $N_{DLT,6g,\ NP101}$, respectively for dose levels 3g, 4.8g. and 6g, which are back-end variables only visible to the medical monitor) dynamically and independently at each dose level being investigated; anyone who is supposed to be blinded to the randomization allocation will be kept blinded to this back-end dynamic calculations in the database.

**Important Safety Alert:** If this dynamic counting of the number of patients having DLTs reach beyond 3 or more DLTs, then the database will immediately alert the investigator, the medical monitor, and PhamacoVigilence (PV) team and the enrollment to the study will be put on hold for new patient enrollment until the PV team, who is unblinded to the treatment allocation, review and confirm these reported DLTs and recommends that the enrollment may continue. The PV committee will generate a report describing such incidences, assuring the blinded nature of the study as much as possible.

1. Once all enrolled patients at **Dose Level 3g** complete their DLT observation period and their DLT evaluability and DLT status are entered into the database, compare the number of NP-101 patients experiencing DLTs (call it $N_{DLT,3g,\ NP101}$) against the following decision boundaries:
    a. If $N_{DLT,3g,\ NP101}<3$, then declare that Dose Level 3g is safe
    b. If $N_{DLT,3g,NP101}\geq 3$, then declare that Dose Level 3g is **NOT** safe and the study team will decide whether or not a lower dose level (Dose Level 2.4g) can be investigated.
2. Once all enrolled patients at **Dose Level 4.8g** complete their DLT observation period and their DLT evaluability and DLT status are entered into the database, compare the number of NP-101 patients experiencing DLTs (call it $N_{DLT,4.8g,\ NP101}$) against the following decision boundaries:
    a. If $N_{DLT,4.8g,\ NP101}<3$, then declare that Dose Level 4.8g is safe and ***patient enrollment to Dose Level 6g will be initiated***.
    b. If $N_{DLT,4.8g,NP101}\geq 3$, then declare that Dose Level 4.8g is **NOT** safe.
        i. If Dose Level 3g was declared to be safe at Step-1 above, then ***Dose Level 3g will be declared the Minimum Effective Dose (MED)*** (also to be referred to as 'the Selected Dose'). In this case, Dose Level 6g will not be investigated.
        ii. If Dose Level 3g was declared to be NOT safe at Step-1 above, then and the study team will decide whether or not a lower dose level (Dose Level 2.4g) can be investigated.
3. Once all enrolled patients at **Dose Level 6g** complete their DLT observation period and their DLT evaluability and DLT status are entered into the database, compare the number of NP-101 patients experiencing DLTs (call it $N_{DLT,6g,\ NP101}$) against the following decision boundaries:
    a. If $N_{DLT,6g,NP101}<3$, then ***declare that Dose Level 6g is safe.*** At this point, the DSMB will review the entire safety data including the Long-Covid data and early indication of efficacy data to determine which of the three dose levels, namely, 3.0, 4.8, and 6.0, be recommended as the Minimum Effective Dose (MED)
    b. If $N_{DLT,6g,NP101}\geq 3$, then declare that Dose Level 6g is **NOT** safe. At this point, the DSMB will review the entire safety data including the Long-Covid data and early indication of efficacy data to determine which of the two safe dose levels, namely, 3.0 and 4.8, be recommended as the Minimum Effective Dose (MED)

Dose-escalation Decision flow chart is given in Figure-1 below:

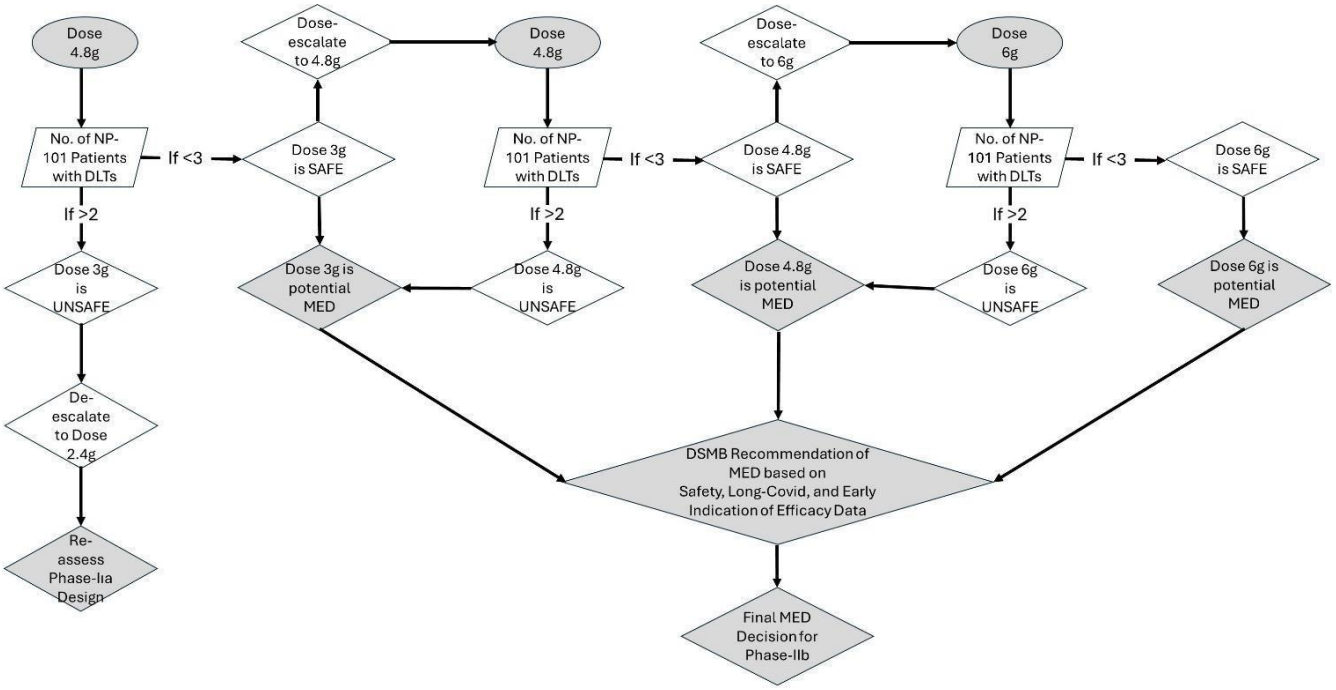


Figure-9.4.2: Decision flow-chart for Phase-IIa Dose Escalation Trial. MED: Minimum Effective Dose, DLT: Dose Limiting Toxicity

Once the safe doses are determined according to the above algorithm, then the DSMB will convene to assess the entire safety and early indication of efficacy data to determine which dose is recommended to the Minimum Effective Dose (MED). To help with this decision-making process, a detailed report for the exploratory objective for the Phase-IIa component of the study will be generated and provided to the DSMB. This report will include the estimate of the Day-5 Sustained Clinical Recovery (SCR-D5) by dose and the average of the CD4+ and CD8+ expression change from baseline to Day-14 in a blinded fashion. If DSMB requests, an unblinded statistician will produce an unblinded report specific to early efficacy indicators for NP-101 arm alone as well.

The above DLT monitoring rules will be employed regardless of patient evaluability for the primary endpoints; however, to be included in the safety monitoring, it will be required that the patient received at least one dose of the study regimen and attribution of a given adverse event to the study regiment is clearly established as ‘at least possibly attributable to the study

regimen.

As part of the Phase-IIa safety assessment, frequencies of overall adverse events, related adverse reactions, adverse events leading to discontinuation of study drug and hospitalizations, or deaths reported in participants taking the Selected Dose of NP-101 versus participants taking placebo will be tabulated and compared for each dose level investigated. In generating and comparing these AE tables between the two study arms, additional tabular dimensions regarding the Covid-19 specific concomitant medications as part of the Standard of Care (SOC) can be added.

**Primary analyses for Phase-IIb:**

**Primary Safety Analyses**

Primary Estimand-1 for the safety objective is the difference of frequencies of overall adverse events (AE), related adverse reactions, adverse events leading to discontinuation of study drug and hospitalizations, or deaths reported in participants taking the Selected Dose of NP-101 versus participants taking placebo regardless of whether they actually receive their assigned therapy (ITT) and assuming that their hospitalization and treatment discontinuation rates will be similar. Possible intercurrent events for this estimand can be, though expected to be quite rare, hospitalization (i.e., transitioning from outpatient status to inpatient status) and treatment discontinuation, both of which may lead to potential missing data or change in the primary measures. Every effort will be put forward to capture any adverse events from Day 1 to Day-21 as well as on the final Safety call on Day 45 from randomization, especially during possible intercurrent events. Therefore, hospitalization-free, while-on-study and per-protocol treatment versions of the above estimand will also be considered within each population definitions above. A representative comparison table is presented below:

Table 6.1. Representative Table Structure for AE comparison between the two study arms

| | | Grade-1 | | Grade-2 | | Grade-3 | | Grade-4 | | Grade-5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Adverse Events | Study Arm | N | % | N | % | N | % | N | % | N | % |
| AE-1 | NP-101 | | | | | | | | | | |
| | Placebo | | | | | | | | | | |

Similar tables will be generated for Serious Adverse Events (SAE), hospitalizations, and deaths regardless of attribution to the study treatment as well as based on ‘at least possible attribution’ to the study treatment. Similar tables will also be added based on the classes of

concomitant medications, especially based on Covid-19 specific medications as part of SOC.

**Primary Efficacy Analyses (Phase IIb only)**

Primary estimand-2 aims to assess the efficacy of the proposed therapy and is defined as the difference of Sustained Clinical Recovery rates on Day 5 in participants taking the Selected Dose of NP-101 as compared to placebo regardless of whether they actually receive their assigned therapy (ITT), and assuming that their hospitalization and treatment discontinuation rates will be similar. The primary efficacy analyses will be conducted based on Phase-IIb participants only.

Possible intercurrent events for this estimand can be, though expected to be quite rare, hospitalization (i.e., transitioning from outpatient status to inpatient status) and treatment discontinuation, both of which may lead to potential missing data or change in the primary measures. The study participants will be reminded to submit their symptom questionnaires even during and after these intercurrent events, if any, to avoid missing data as much as practically possible. As the patient-reported symptom outcomes are captured through eCRFs online, such missing data is expected to be limited. Hospitalization-free, while-on-study and per-protocol treatment versions of the above estimand will also be considered as defined above.

Sustained Clinical Recovery rate will be estimated at Day 5 as the ratio of those who achieve improvement of all symptoms to Mild and None and remained so for at least three days to all participants treated in each arm. 95% Clopper-Pearson Confidence Intervals will be constructed for Sustained Clinical Recovery Rate for each arm, Wald Asymptotic Conference Interval will be constructed for the Sustained Clinical Recovery Rate difference between the two arms, and Z-test will be used to assess the significance of the difference between the two arms. Although all analyses will be conducted based on the intent-to-treat context, sensitivity analyses based on the adherence to the protocol regimen and based on different missing data imputation approaches will also be carried out for the primary as well as secondary objectives.
Primary efficacy analyses will also include a dose*response analyses utilizing the data from Phase-IIa dose escalation trial alone, where 95% Clopper-Pearson Confidence Intervals will be constructed for Sustained Clinical Recovery Rate for each dose level of 3g, 4.8g and 6g, and Z-test will be used to assess the significance of the difference between each of the three dose levels (expecting about 15 patients at each level) and the combined placebo arm from the Phase-IIa trial (expecting about 15 patients).

The PRO Symptom Survey asks the following specific question: “It was easy to report my symptoms using this survey”, to which ‘Agree and ‘Disagree’ are the two response

options.

The responses to this question will be analyzed to evaluate the survey after the Phase-IIa component of the protocol is completed. Based on the satisfaction level, the language of the questionnaire might be improved before its use by the Phase-IIb participants.

**Algorithm for Sustained Clinical Recovery Assessment**

**Day-5 Sustained Clinical Recovery is defined as** the continuous alleviation or resolution of all COVID-19 PRO Symptom Survey scores to ≤1 ("none" or "mild"), starting on or before Day 5 after randomization and remaining so for at least 3 consecutive days and without subsequent relapse, defined as not progressing to "moderate" or "severe" for at least two consecutive days. The Day-5 Sustained Clinical Recovery (SCR-D5) assessment will utilize all symptoms data from randomization to Day 7.

These SCR assessments will be conducted beyond Day-5 as well until Day-28 for the secondary objectives. As the sustained clinical recovery assessment requires two additional days of symptoms data, the last day that it can be assessed is Day 28. Therefore, the SCR status of each patient will be captured under the following variables in the study database.

- Whether a given patient achieves a sustained clinical recovery by Day 5 (SCR5)
- Days to SCR by Day-5 (SCR_Days5)
- Whether a given patient achieves a sustained clinical recovery by Day-28 (SCR28)
- Days to SCR by Day 28 (SCR_Days28)

**For Day-5 Sustained Clinical Recovery (SCR-D5) assessment, the following iterative algorithm will be used in the study database and independently confirmed outside the database:**

**Step-1:**

**For each patient and starting from Day 1 from Randomization,**

Check if patient has any progression for two consecutive days, hospitalization or death, or if they withdraw between Day-8 and Day-28 of randomization. If so, the patient will be declared not to achieve Day-5 SCR. If not, then check all the symptom scores on **Day 1** are <=1 (≤"mild" or ≤"2 times" or "My sense of smell is THE SAME AS usual" or "My sense of taste is THE SAME AS usual")

1. If YES, check if all the symptom scores on **Day 2** are <=1
    a. If YES, check if all the symptom scores on **Day 3** are <=1
        i. If YES, then SCR5=YES, SCR_Days5=1, and go to Final Step
        ii. If NO, go to Step-2
    b. If NO, go to Step-2
2. If NO, go to Step-2

**Step-2:**

Check if all the symptom scores on **Day 2** are <=1 (≤"mild" or ≤"2 times" or "My sense of smell is THE SAME AS usual" or "My sense of taste is THE SAME AS usual")

1. If YES, check if all the symptom scores on **Day 3** are <=1
    a. If YES, check if all the symptom scores on **Day 4** are <=1
        i. If YES, then SCR5=YES, SCR_Days5=2, and go to Final Step
        ii. If NO, go to Step-3
    b. If NO, go to Step-3
2. If NO, go to Step-3

**Step-3:**

Check if all the symptom scores on **Day 3** are <=1 (≤"mild" or ≤"2 times" or "My sense of smell is THE SAME AS usual" or "My sense of taste is THE SAME AS usual")

1. If YES, check if all the symptom scores on **Day 4** are <=1
    a. If YES, check if all the symptom scores on **Day 5** are <=1
        i. If YES, then SCR5=YES, SCR_Days5=3, and go to Final Step
        ii. If NO, go to Step-4
    b. If NO, go to Step-4
2. If NO, go to Step-4

**Step-4:**

Check if all the symptom scores on **Day 4** are <=1 (≤"mild" or ≤"2 times" or "My sense of smell is THE SAME AS usual" or "My sense of taste is THE SAME AS usual")

1. If YES, check if all the symptom scores on **Day 5** are <=1
    a. If YES, check if all the symptom scores on **Day 6** are <=1
        i. If YES, then SCR5=YES, SCR_Days5=4, and go to Final Step
        ii. If NO, go to Step-5
    b. If NO, go to Step-5
2. If NO, go to Step-4

**Step-5:**

Check if all the symptom scores on **Day 5** are <=1 (≤"mild" or ≤"2 times" or "My sense of smell is THE SAME AS usual" or "My sense of taste is THE SAME AS usual")

1. If YES, check if all the symptom scores on **Day 6** are <=1
    a. If YES, check if all the symptom scores on **Day 7** are <=1
        i. If YES, then SCR5=YES, SCR_Days5=4, and go to Final Step
        ii. If NO, then SCR5=NO, SCR_Days5=., and go to Final Step
    b. If NO, then SCR5=NO, SCR_Days5=., and go to Final Step
2. If NO, then SCR5=NO, SCR_Days5=., and go to Final Step

**Final Step:** Day-5 SCR Evaluation of the current patient is complete

Here is the decision flowchart for this algorithm.

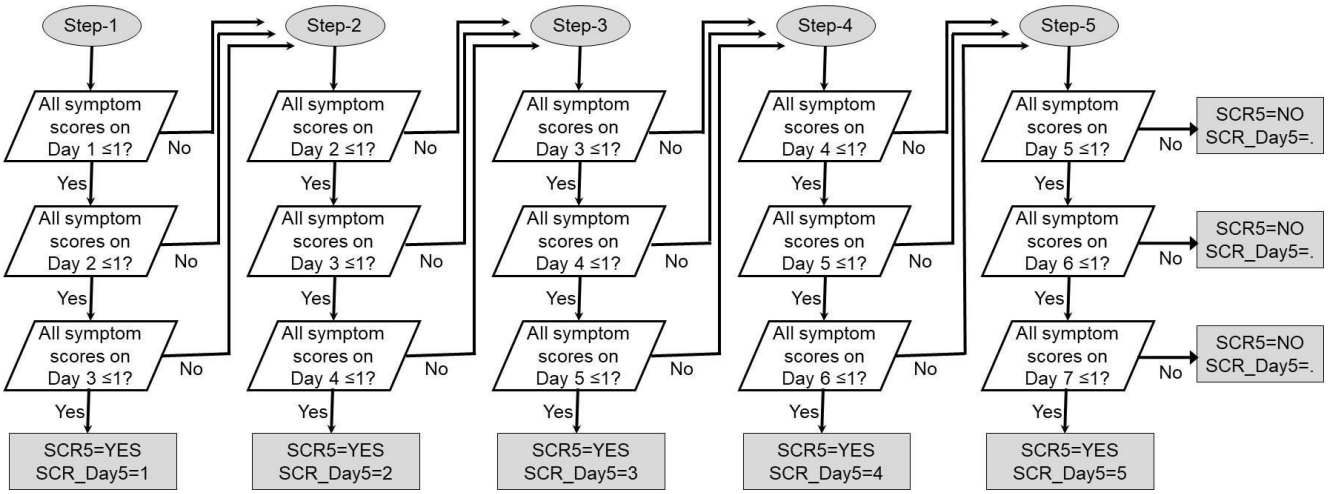


Figure 2. Algorithm for Day-5 Sustained Clinical Recovery (SCR-D5) Assessment (Day 1 here is the day after randomization, and Day 5 is five days after Randomization).

An expanded algorithm up to Day 28 will be also generated for the secondary objectives using the logic and flow as in the above algorithm as the daily symptom survey will be implemented until Day 30.

Primary Efficacy objective has a planned interim analysis to be conducted using the accumulated data when at least 50 patients at each study arm are randomized and receive at least one dose of the assigned dose in the Phase IIb portion of the study. Due to the random allocation of patients, the actual sample size for the interim analysis can be higher than 100. The details of this planned interim analysis are provided in Section 8.6 of this SAP.

**Missing Data Handling Strategy for the Primary Objective**

We do not expect excessive missing data for the primary objective data considering the safeguards in place in the protocol as described in Section 4.1 (sub-section Patient Reported Outcomes and Conjunctivitis Questionnaire). For the primary efficacy endpoint, all missing values within 28 days from randomization will be conservatively considered to be at least moderate level symptoms. When appropriate, Last-Value-Carried-Forward (LVCF) imputation approach will be utilized for the missing symptom values especially in the symptom questionnaire as the patient-reported symptoms will be captured from randomization through Day 45 and intermittent missing values can be imputed via the LVCF approach. The LVCF approach makes clinical sense and is a more conservative approach as the symptoms of

Covid-19 are expected to improve over time. As a sensitivity analysis, we also plan to employ Next-Value-Carried-Backward (NVCB) approach as another missing data imputation approach to account for the effect of late-developing symptoms in a more conservative manner.

For the primary efficacy objective, symptoms reported up to Day 5 will be critical to assess 'clinical recovery' and the symptoms reported from Day 5 to Day 7 are critical to assess the 'sustained' clinical recovery. As this is a study with a short span of enrollment and treatment windows (14-day therapy, 45-day symptom assessment), we do not anticipate excessive missing values in symptoms especially within the first 7 days where the participants are expected to be more engaged with the study; however, missing data structure will still be analyzed as to whether or not the missingness is Missing Completely at Random (MCAR), or Missing at Random (MAR), or Not-Missing at Random (NMAR). In the event of MCAR and MAR, sensitivity analyses will be carried out in the final analysis by generating 100 randomly imputed data using the MI (Multiple Imputation) procedure in SAS ® Version 9.4 based on the following steps and sample SAS code:

- The ordinal nature of the symptoms will be retained. The symptoms codes will be modified to have a more unified scale such that 0 represents the non-existence of the symptom, 1 represents "mild" symptoms, 2 represents "moderate" symptoms, and 3 represents "severe" symptoms. For the smell and taste symptoms, 1 will represent "some loss" in smell or taste and 2 will represent "compete loss" of smell or taste.
- The symptom data will be expressed horizontally as Symptom1, Symptom2, …, Symptom14. This structure will ensure that the correlation among symptoms will be also taken into account when multiply-imputed datasets are being produced.
- If a patient has both the screening and baseline symptoms, their baseline symptoms will be used as part of the imputation process, so that each patient has a single set of and most current pre-screening symptom data.
- All symptom data including Day-29, Day30, Day-37, and Day-45 symptoms as well will be utilized.
- Days from Symptom Onset to Randomization, Days from Randomization, Age, Gender, Race, Obesity Status, Diabetes Status, and smoking status will be also added to the MI procedure as predictors for more accurate imputation through increasing the plausibility of Missing at Random (MAR) assumption. Age was categorized as <30, 30-39, 40-49, 50-59, and 60+ to help with the convergence of MI estimations. As the symptoms are also a potential products of annual seasons, a season variable is also added to represent winter, spring, summer, and fall to help with the symptom imputation.
- 100 multiply-imputed datasets will be produced using a code similar to the following:

```
proc mi data=boss2_symptomdata out=mi_output nimpute=100 seed=101;
```

```
class season agecat gender hr_Diabetes hr_obesity IE_SMOKER race ss_nose--
ss_taste;
var season agecat daysfromenrollment days_from_symptom_onset  gender
hr_Diabetes hr_obesity IE_SMOKER race ss_nose--ss_taste;
fcs logistic(season agecat gender hr_Diabetes hr_obesity IE_SMOKER race
ss_nose--ss_taste);
run;
```

Sensitivity analyses will also include hospitalization-free, while-on-study, and per- protocol treatment cohorts as described above.

.

### 6.5.3 Secondary Analyses

None of the following secondary endpoints depend on the findings of the primary endpoint.

**Secondary Objective-1:**

The first secondary objective is “To compare the change in viral load and cycle threshold profiles from Randomization to Day 7 and Day 14 in participants taking the Selected Dose of NP-101 as compared to placebo”

Secondary estimand-1 is defined as the difference of the average viral load (VL) and cycle threshold (CT) measures from Randomization (Day 0) to Day 7 and Day 14 in participants taking the Selected Dose of NP-101 as compared to placebo regardless of whether they actually receive their assigned therapy (ITT) and assuming that their hospitalization and treatment discontinuation rates will be similar. The viral load and cycle threshold will be measured at Baseline (Day of Randomization, i.e., Day 0), on Day 7, and on Day 14; therefore, the outcome variable of interest will be a longitudinal ratio-scale measure. When RT-PCR test is negative, no viral load is reported. Therefore, in such cases, viral load will be imputed with a value of zero. The change of viral load and cycle threshold from Day 0 to Day 7 and Day 14, respectively will be computed and named as *Change_VL_Day7*, *Change_CT_Day7*, *Change_VL_Day14*, and *Change_CT_Day14*. Then, distribution of VL and CT at each visit (namely, Day 0, Day 7, and Day 14) as well as their changes from Day 0 will be summarized and compared between NP-101 and Placebo arms using Wilcoxon-Mann-Whitney test. Although the baseline distributions of viral load and cycle threshold are expected to be similar due to randomization, Analysis of Covariance (ANCOVA) models will also be constructed for the log (viral load) and cycle threshold according for the baseline (Day 0) measurements in a model similar to the following, where Day 0 Assessment and Study Arm interactions will be assessed as well:

$$Change\_CT\_Day7 = Day0\_CT + Study\_Arm + Day0\_CT * Study\_Arm$$

To compare the change over time in VL and CT, Random Coefficients Models will be utilized where VL and CT will be the dependent variables and study arm and time (i.e., days from randomization) will be the main predictors similar to the below representative model:

*Log(VL)=Time (in actual days)+Study_Arm+ Time*Study_Arm*

In these models, rather than using the visit timing such as Day 7 and Day 14, we will utilize the actual day of assessment from randomization if they slightly deviate from the protocol specified visit date.

The above mentioned intercurrent events, namely, hospitalization (i.e., transitioning from outpatient status to inpatient status) and treatment discontinuation, are applicable for Secondary estimand-1 as well. Therefore, hospitalization-free, while-on-study and per-protocol treatment versions of the above estimand will also be considered within each analysis set defined above.

**Secondary Objective-2:**

Secondary estimand-2 is defined as the measurement of the difference in the percentage of negative/undetectable RT-PCR (i.e., viral clearance) on Day 7 and Day 14 in participants taking the Selected Dose of NP-101 as compared to placebo regardless of whether or not they actually receive their assigned therapy (ITT) and assuming that their hospitalization and treatment discontinuation rates will be similar. To address this estimand, we will tabulate the results as in the representative table below:

Table 6.1. Representative Table Structure for AE comparison between the two study arms

| | | Day 7 | | Day 14 | |
|---|---|---|---|---|---|
| RT-PCR Results | Study Arm | N | % | N | % |
| Negative | NP-101 | | | | |
| | Placebo | | | | |
| Positive | NP-101 | | | | |
| | Placebo | | | | |

Then, the difference of RT-PCR negative/undetectable rates between the two study arms will be estimated along with 95% Wald Asymptotic Conference Interval, and Z-test will be used to assess the significance of the difference between the two arms separately on Day 7 and Day 14 visits.

The above mentioned intercurrent events are applicable for Secondary estimand-2 as well. Therefore, hospitalization-free, while-on-study and per-protocol treatment versions of the above estimand will also be considered within each analysis set defined above.

**Secondary Objective-3:**

Secondary estimand-3 will be the difference of the survivorship functions modeling the days of Sustained Clinical Recovery (SCR), days to Complete Clinical Recovery (CCRec), and days to Complete Clinical Resolution (CCRes) survival in participants taking the Selected Dose of NP-101 as compared to placebo. Distribution of the days to Sustained Clinical Recovery for the treatment and placebo arms will be estimated using the Kaplan-Meier method and compared using the Log-rank test between the two study arms. If the proportional hazards assumption is not satisfied, Wilcoxon-Gehan test will be used. Recovery Free Survivorship estimates along with their 95% confidence intervals will be provided for each day from Day 1 from randomization to Day 28.

For the survivorship estimates for SCR, events will be defined as participants that attain Sustained Clinical Recovery while on study until Day 28 and days to event will be calculated as the days from randomization (Day-0) to the first of the three consecutive days of Clinical Recovery, that is, the first day when all symptoms are resolved to at most “mild” or “none” and remained so for at least two additional days for a total of three consecutive days. Participants who do not attain CRS during the follow-up window will be censored on Day 28 or on the day when they are taken off study before Day 28 due to other reasons.

For the survivorship estimates for Complete Clinical Recovery (CCRec), events will be defined as participants that attain CCRec while on study until Day 30 and days to CCRec will be calculated as the number of days from randomization to the first day when all symptoms scored as moderate or severe at study entry are scored as mild or absent AND all symptoms scored mild or absent at study entry are scored as absent and remain so afterwards until Day 30. Patients who do not attain CCRec will be censored on Day 30 or on the day when they are taken off study before Day 30 due to other reasons.

For the survivorship estimates for Complete Clinical Resolution (CCRes), events will be defined as participants that attain CCRes while on study until Day 30 and days to CCRes will be calculated as the number of days from randomization to the first day when all symptoms are scored as absent and remain so afterwards until Day 30. Patients who do not attain CCRes will be censored on Day 30 or on the day when they are taken off study before Day 30 due to other reasons.

Survival distributions for each of the endpoint above will be estimated using Kaplan-Meier method and compared between the two study arms using Log-rank test. In the event that other factors of interest would be needed to be adjusted for, then Cox Proportional Hazards Models will be employed.

Although all analyses will be conducted based on the intent-to-treat context, sensitivity analyses based on the adherence to the protocol regimen will also be carried out with respect to

the analysis data sets defined above.

**Secondary Objective-4:**

Secondary estimand-4 is the difference between the average level of effector immune cells, specifically CD4+ and CD8 T-cells, from Baseline (day of randomization) to Day 7 and Day 14 in participants taking the selected dose of NP-101 versus those taking placebo. The following variables will be computed:

*Change_CD4_Day7*, *Change_CD8_Day7*, *Change_CD4_Day14*, and *Change_CD8_Day14*

These markers at each visit (namely, baseline, Day 7, and Day 14) as well as their change from Baseline to Day 7 and Baseline to Day 14 will be compared for between the arms using Wilcoxon-Mann-Whitney test. Although the baseline distributions of CD4+ and CD8+ markers are expected to be similar due to randomization, Analysis of Covariance (ANCOVA) models will also be established in a model structure similar to the following, where Baseline Assessment and Study Arm interactions will also be assessed:

*Change_CD4_Day7=Baseline_CD4+Study_Arm+ Baseline_CD4*Study_Arm*

To compare the change over time in these markers, Random Coefficients Models will be utilized where each of these markers will be the dependent variable and study arm and time (i.e., days from randomization) will be the main predictors similar to the below representative model:

*CD4=Time (in actual days)+Study_Arm+ Time*Study_Arm*

In these models, rather than using the visit timing such as Day 7 and Day 14, we will utilize the actual day of assessment from randomization if they slightly deviate from the protocol specified visit date. A similar model will be constructed for CD8+ as well.

**Secondary Objective-5:**

The secondary estimand-5 is the difference in the percentage of disease progression on

Day 5 measured by the worsening of the Covid-19 symptoms from “none” or “mild” to “moderate” or “severe” and remaining at that level for at least two days before alleviation or resolution, or the experience of hospitalization or death due to Covid-19 symptom worsening in participants taking the Selected Dose of NP-101 versus placebo. Such data will be summarized in a table like the following:

Table 3. Symptom Progression and hospitalization summary

| | | NP-101 | | Placebo | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Symptom Progression | Yes | | | | |
| | No | | | | |
| Hospitalization | Yes | | | | |
| | No | | | | |

To address this estimand, the disease progression rate difference between the two study arms will be estimated along with 95% Wald Asymptotic Conference Interval, and Z-test will be used to assess the significance of the difference between the two arms on Day 5. These analyses will be conducted under each analysis datasets defined above.

**Secondary Objective-6:**

Each Phase-IIb patient will be assessed to determine whether or not his/her symptom burden qualifies as having experienced Long-Covid symptoms. To do that, a patient will be classified as having experienced Long-Covid if they reported any mild, moderate, or severe Long-Covid symptoms as defined and listed by Center for Disease Control and Prevention (CDC), as new or ongoing symptoms per questionnaire conducted on any of the following days: Day 30 ± 2 days, Day 37 ± 2 days, or Day 45 ± 2 days”. The Long-Covid symptom assessment will be summaries as follows:

Table 4. Long-Covid Symptom summary

| | | NP-101 | | Placebo | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Long-Covid Symptoms | Yes | | | | |
| | No | | | | |

The secondary estimand-6 is the difference in the percentage of patients experiencing Long-Covid symptoms in participants taking the MED of NP-101 versus placebo. To address this estimand, the Long-Covid symptom rates difference between the two study arms will be

estimated along with 95% Wald Asymptotic Conference Interval, and Z-test will be used to assess the significance of the difference between the two arms. These analyses will be conducted under each analysis datasets defined above.

### 6.5.4 Other Analyses

The study includes exploratory objectives regarding the effect of NP-101 on inflammatory cytokines and coagulation factors at Days 0, 7, and 14 in participants, and a plan to construct a risk prediction model for Sustained Clinical Recovery based on known risk factors as well as other patient and disease characteristics in participants.

Analysis plan for Exploratory Aim-1:

First exploratory objective aims to estimate and compare the change in inflammatory cytokines and coagulation factors from Baseline to Day 7 and Day 14 in participants taking the MED of NP-101 as compared to placebo.

To do that, descriptive and graphical statistics will be used to present the distribution of inflammatory cytokines and coagulation factors on Baseline, Day 7 and 14 within each study arm. The change in these markers from Baseline to Day 7 and Day 14, respectively, will be computed and named as *change_marker_day7, change_marker_day14*. The distribution of these markers at each visit (namely, Baseline, Day 7, and Day 14) as well as their change on study (namely, change from baseline to Day 7 and change from Day0 to Day 14) will be compared for differential distribution between the NP-101 and Placebo arms using Wilcoxon-Mann-Whitney test. ANCOVA models like the following will also be constructed:

$$change\_marker\_day7 = marker\_day0 + study\_arm + marker\_day0 * study\_arm$$

To compare the change over time in these markers, Random Coefficients Models will be utilized where each of these markers will be the dependent variable and study arm and time will be main predictors in models like the following:

$$marker = time + study\_arm + time * study\_arm$$

As these analyses are exploratory to generate hypotheses, no multiplicity correction will be applied at this point.

Analysis plan for Exploratory Aim-2:

Second exploratory objective aims at constructing a risk prediction model for Day-5 Sustained Clinical Recovery (SCR-D5) based on known risk factors as well as other patient and disease characteristics.

To do that, we first plan to utilize multivariable Logistic Regression approach, where Day-5 Sustained Clinical Recovery (SCR-D5) will be the primary outcome variable, and patient characteristics such as study arm, gender, ethnicity, age, body mass index (BMI) and clinical characteristics such as comorbidities, baseline viral load or baseline cycle threshold, baseline CD4+ and CD8+ level, baseline inflammatory cytokines and coagulation factors, vaccination status, days from Symptom Onset to Randomization, etc. will be the predictors. If needed, we will utilize stepwise variable selection approach to refine the final model by also including significant interaction terms especially with Study Arm. Some of these significant interactions may lead to further subset analyses and modeling.

Secondly, we plan to utilize a similar modeling strategy targeting the time to SCR-D5 through Cox Proportional Hazards Models, where time to SCR-D5 will be the primary dependent variable along with SCR-D5 status being the censoring variable, and the final predictive model will identify from among the potential predictors mentioned above a set of variables and interactions especially with Study Arm that are significantly associated with time to SCR-D5 in a context of right- censored data.

Analysis plan for Exploratory Aim-3:

Third exploratory objective aims at comparing the conjunctivitis rates between the two arms of the study. A patient will be classified as having experienced conjunctivitis if they reported any mild, moderate, or severe conjunctivitis symptoms listed in Appendix 2C anytime while on study until Day 30.

Table 5. Conjunctivitis Symptom summary

| | | NP-101 | | Placebo | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Conjunctivitis Symptoms | Yes | | | | |
| | No | | | | |

To address this exploratory aim, the conjunctivitis symptom rates difference between the two study arms will be estimated along with 95% Wald Asymptotic Conference Interval, and Z-test will be used to assess the significance of the difference between the two arms. These analyses will be conducted under each analysis datasets defined above.

Analysis plan for Exploratory Aim-4:

The forth exploratory objective aims at comparing the long-Covid rates between the two

arms in the Opt-In Expansion Component of the study. The Long-Covid symptom rates will be reported as in Table-6 below.

Table 6. Long-Covid Symptom summary

| | | NP-101 | | Placebo | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| Long-Covid Symptoms | Yes | | | | |
| | No | | | | |

To address this exploratory aim, the Long- Covid symptom rates difference between the two study arms will be estimated along with 95% Wald Asymptotic Conference Interval, and Z- test will be used to assess the significance of the difference between the two arms. These analyses will be conducted under each analysis datasets defined above.

### 6.5.5 Covariate and Subgroup Analyses

The exploratory objectives of the study already include such analyses. However, additional covariate and subgroup analyses will be presented when applicable and supported by the available data as additional exploratory analyses to guide similar prospective studies.

### 6.5.6 Safety Analyses

Safety analyses were provided in details as part of the Primary Objectives of Phase-IIa and Phase-IIb in Section 6.5.2. above.

## 7 Definitions for Clinical Evaluations

### 7.1 Follow-Up Time Points

The treatment time for each participant is 14 days and the follow-up time is for 7 days (up to Day 21 both as DLT observation period and up to 45 days for symptom assessment) or until ongoing adverse events are resolved up to 45 days as part of safety follow-up.

### 7.2 Points of Enrollment

Subjects will be considered enrolled in the study when they have been informed of all aspects of the study, had ample time to review the Informed Consent document, signed the Informed Consent document, satisfied the Inclusion/Exclusion Criteria, and randomized.

### 7.3 Definitions for Clinical Evaluation

Clinical evaluation will be done using the protocol specified Patient Reported Outcome (PRO) questionnaire for high risk Covid-19 patients treated in an outpatient setting. The participants will fill out the questionnaire at the time of Screening, Enrollment (i.e., Randomization), and from Day-1 to Day-30, Day 37 and Day 45 post randomization each day using their study app.

## 8 General Considerations

If a subject is not able to complete the study after being enrolled, the Investigator, or designee will document the last contact with the subject. All attempts to contact the subject will be documented within the EDC. Missed or out-of-range visits will be considered protocol deviations.

### 8.1 Missing Data and Outliers:

Missing data strategies are provided in Section 6.5.2 above.

The data quality checks will be an ongoing effort for the study and potential outliers will be detected and resolved as much as possible while the study is ongoing, without unblinding, with close collaboration with the database management team, the enrolling sites, and the affiliated labs. In the analysis phase, outliers based on their face values will be evaluated based on whether or not they are still outliers and influential values in given statistical models as such outliers may be perfectly in line with the overall association captured by such models built and do not necessarily violate modeling assumptions. If they are found to be influential, then possible data transformation approaches such as log-transformation will be considered.

### 8.2 Protocol Violations and Deviations

Missed visits and out of range visits will be considered protocol deviations while such participants will still be deemed evaluable for the primary objective. Investigators will not deviate from the clinical protocol except to deliver emergency care or to eliminate an immediate hazard to the subject. All study deviations from the clinical protocol will be captured in the study database. All deviations with the potential to affect subject safety, rights, or well-being will also be reported as required to the IRB and to sponsor.

### 8.3 Disposition of Subjects and Withdrawals

All subjects who provide informed consent and are enrolled will be accounted for in this study. Subjects maintain the right to discontinue their participation in the study at any point. If a subject exercises this right, the Investigator will make every effort to ascertain and document the reason for withdrawal.

### 8.4 Sensitivity Analysis

Sensitivity analyses based on different data imputation approaches were described in Sections 6.5.2 and 8.1 above.

### 8.5 Changes to Planned Analysis after Blind Review of Data and Unblinding Consideration (if Applicable)

No changes to the planned analyses are expected. While the study is enrolling and patients are still on study, unblinding of the randomization status of a study participant will only occur if it is necessary for the patient safety and such occurrences are not expected to affect the statistical analysis plans. If the planned analyses have to be changed for any unforeseen reasons, then this SAP will be amended with reasons provided and submitted to FDA.

### 8.6 Interim Analysis and Data Monitoring

There is one interim analysis whose decision rule will be built into the study database using the accumulated data of the initial at least 100 evaluable patients of the Phase IIb portion of the study (at least 50 patients who were randomized and who consumed at least one dose of the assigned drug within each arm). Here, evaluability of a patient will be defined as being randomized and consuming at least one dose of the assigned drug. When at least 50 patients in each arm are enrolled, randomized, and confirmed to have consumed at least one dose of the study medication and completed at least 28 days on study, the enrollment will be paused for the interim analysis. Due to random allocation without block randomization, the actual sample size for the interim analysis can likely be greater than 100 participants.

The following decision-making algorithm will be built-in to the study database and the computations will be done seamlessly and the resulting report will be submitted to the PV Committee as well as the study team and the sponsor:

Once the Sustained Clinical Recovery evaluation is done for each patient, we carry out the following computations to produce the value of the test statistics for the interim analysis:

- Number of patients treated in the in the NP-101 arm: $N_{NP101}$
- Number of patients treated in the in the Placebo arm: $N_{Placebo}$
- Number of patients attaining SCR-D5 in the in the NP-101 arm: $X_{NP101}$
- Number of patients attaining SCR-D5 in the in the Placebo arm: $X_{Placebo}$
  - Proportion of Patients attaining SCR-D5 in the NP-101 arm: $\hat{p}_{NP101} = \frac{X_{NP101}}{N_{NP101}}$
  - Proportion of Patients attaining SCR-D5 in the Placebo arm: $p_{Placebo} = \frac{X_{Placebo}}{N_{Placebo}}$

Using the above statistics. Compute the test statistic for the proportion difference as follows:

$$Z = \frac{\hat{p}_{NP101} - \hat{p}_{Placebo}}{\sqrt{\frac{\hat{p}_{NP101}(1-\hat{p}_{NP101})}{N_{NP101}} + \frac{\hat{p}_{Placebo}(1-\hat{p}_{Placebo})}{N_{Placebo}}}}$$

The decision boundary for the interim analysis is presented in Table 9.2.1 above in Section 3.2.

Among the selected parameterization, the minimum value to stop early for FUTILITY is 0.688 and maximum value to stop for EFFICACY is 2.737.

Compare the Z-value computed above with these boundary values:

- If Z≤0.688, then recommend ***STOPPING EARLY for FUTILITY***
- If Z≥2.737, then recommend ***STOPPING EARLY for EFFICACY***
- If 0.688<Z<2.737, then recommend that the ***PATIENT ENROLLMENT CONTINUE*** to reach the targeted enrollment size up to 248 patients.
- If the study DSMB recommends early stopping for futility or early stopping for efficacy based on the accumulated safety and efficacy data and based on the evidence relative to the above decision boundaries, the sponsor may decide to stop the trial early: Only if the study is stopped, once all data cleaning procedures are completed and the study database is permanently locked, the study statistician will be unblinded and carry out analyses to address all secondary and exploratory objectives according to the analyses sets in Section 6.4.

The above recommendation will be evaluated by the DSMB along with accumulated other safety data as well and by the study team to finalize the decision regarding early stopping for futility, early stopping for efficacy, or continued enrollment.

### 8.7 Timing with Respect to Database Locking

The data quality checks will be an ongoing effort while the study is still ongoing. For the interim analysis, once at least 50 patients in each arm are enrolled and randomized and confirmed to have consumed at least one dose of the study medication and followed for at least 28 days, then the enrollment will be halted and the data will be soft-locked for the interim analysis.

Following the interim analysis, if the decision is to continue to complete the enrollment up to 248 patients, then when all patients are enrolled, randomized and complete their symptom assessments up to Day-45, a complete data- extraction will be requested for a efficacy data quality control and confirmation. Once the efficacy data is deemed to be complete and accurate, the efficacy data will be soft-locked and another complete data extraction will be requested to generate a preliminary efficacy report. As part of this preliminary efficacy analysis, all the members of the study team will remain blinded except the unblinded statistician, the data manager and the head of regulatory affairs. The preliminary safety report will not contain any individual level data that may potentially reveal the randomization status. Upon completion of the 45-day safety follow-

up of the last enrolled patient, there will be a final data quality assessment finalizing the safety parameters, generating final set of queries for the study sites to correct/modify/update the existing data as needed. Once all data quality check queries are adequately addressed, the study database will be ordered to be locked permanently and the entire study team will be unblinded to patient randomization.

### 8.8 Analysis Software

All analysis will be performed using SAS® Software version 9.4 or later. When appropriate, R-package will be used as well.

### 8.9 List of Tables, Figures and Listings

Demographic characteristics table, AE and SAE tables, tables of Covid-19 test results both from the rapid test or RT-PCR at screening, the distribution statistics table for the viral load and cycle threshold assessments, any other secondary objective measure tables. PRO symptom assessments will be shown graphically as time series plot with a reference line at 1.5 separating Level-1 (‘mild’) and Level-2 (“moderate”) symptoms where the y-axis will show the entire symptom scale from ‘None (Level-0)’ to ‘Severe (Level-3)’.

## 9 References

### 9.1 Applicable Regulatory Guidelines:

- ICH E9: Statistical Principles for Clinical Trials (September 1998)

### 9.2 Other applicable Documents

Not Applicable