Statistical Analysis Plan

Sponsor Name:	**Nohla Therapeutics**
Protocol Number and Title:	**2603.00 Multi-center, open-label randomized study of single or double myeloablative cord blood transplantation with or without infusion of off–the-shelf ex vivo expanded cryopreserved cord blood progenitor cells in patients with hematologic malignancies**
Original Protocol Version, Date:	**Version 1, 05SEP2012**
Protocol Amendment Number, Date	**Amendment 1, 02NOV2012; Amendment 2, 09MAY2013; Amendment 3, 16SEP2013; Amendment 4, 18DEC2013; Amendment 5, 13JAN2014; Amendment 6, 13MAY2014; Amendment 7, 31JUL2014; Amendment 8, 26JAN2015; Amendment 9, 19MAR2015; Amendment 10, 13APR2015; Amendment 11, 11JUN2015; Amendment 12, 07OCT2015; Amendment 13, 07DEC2015; … Amendment 22, 23AUG2017; Amendment 23, 20SEP2017; Amendment 24, 24APR2018; Amendment 25, 24AUG2018**
Syneos Health Project Code:	**1011511**
SAP Version:	**Final Version 3.0**
SAP Version Date:	**26FEB2019**

Revision History

Version #	Date (dd-mmm-yyyy)	Document Owner	Revision Summary
1	11MAY2018	Ted A. Gooley	Initial Release Version
1.1	04JUN2018	K Kwarteng	Update version 0.1 to full clinical study report SAP.
2	24JUL2018	K Kwarteng	Update based on sponsor comments
3	26FEB2019	T Marmon	Update to remove safety and efficacy infection endpoints based on blinded review of the data, as reliable analyses for these infection endpoints were deemed impossible, and the study was terminated.

I confirm that I have reviewed this document and agree with the content.

APPROVALS	
Nohla Therapeutics	
	26 FEB 2019
Tonya K. Marmon, DrPH Vice President, Biometrics	**Date** (dd-mmm-yyyy)
	26 Feb 2019
Sarah Noonberg, MD, PhD Chief Medical Officer	**Date** (dd-mmm-yyyy)

TABLE OF CONTENTS

1. GLOSSARY OF ABBREVIATIONS

Abbreviation	Description
AE	Adverse Event
ALL	Acute Lymphocytic Leukemia
AML	Acute Myeloid Leukemia
ANC	Absolute Neutrophil Count
ATC	Anatomical Therapeutic Chemical
BAL	Biphenotypic Acute Leukemia
BM	Bone Marrow
CB	Cord Blood
CBT	Cord Blood Transplant
CI	Confidence Interval
CML	Chronic Myeloid Leukemia
CNS	Central Nervous System
CR	Complete Response
CR1	First Complete Response
CR2	Second Complete Response
CRF	Case Report Form
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DLCO	Carbon Monoxide Diffusing Capacity (Lung Function)
DSMB	Data Safety Monitoring Board
ECOG	Eastern Cooperative Oncology Group
EXP	Experimental Arm
FAB	French-American-British Classification System
FAS	Full Analysis Set
FHCRC	Fred Hutchinson Cancer Research Center
GRFS	Graft-versus-Host Disease-free, Relapse-Free Survival
GVHD	Graft-versus-Host Disease
HCT	Hematopoietic Cell Transplant
HCT-CI	Hematopoietic Cell Transplantation–Specific Comorbidity Index

Abbreviation	Description
HLA	Human leukocyte Antigen
ICH	International Council for Harmonization
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IPSS	International Prognostic Scoring System
IRB	Institutional Review Board
Max	Maximum
MedDRA	Medical Dictionary for Regulatory Activities
MDS	Myelodysplastic Syndrome
Min	Minimum
mITT	Modified Intent to Treat
MLL	Mixed-Lineage Leukemia
N/A	Not Applicable
NIH	National Institutes of Health
NRM	Non-relapse Mortality
PCR	Polymerase Chain Reaction
PI	Principal Investigator
PPS	Per-Protocol Set
RAEB	Refractory Anemia with Excess Blasts
RAEBt	Refractory Anemia with Excess Blasts in Transformation
REFR	Refractory
REL	Relapse
Rh	Rhesus
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SI	Standard International System of Units
SOC	Standard of Care
SOP	Standard Operating Procedure
SS	Safety Set
TBI	Total-body Irradiation
TCR	T-cell Receptor

Abbreviation	Description
TEAE	Treatment Emergent Adverse Event
TKI	Tyrosine Kinase Inhibitors
TLF	Table, Listing and Figure
TREC	T-cell Receptor Excision Circles
TRM	Transplant-related Mortality
WHO	World Health Organization

2. PURPOSE

The purpose of this Statistical Analysis Plan (SAP) is to ensure that the data listings, summary tables and figures which will be produced to support the completion of the Clinical Study Report (CSR), and the statistical methodologies that will be used, are complete and appropriate to allow valid conclusions regarding the study objectives according to International Council for Harmonization (ICH) E9[1], ICH E3[2] and ICH E6[3]. This analysis plan is based on the latest version of both the protocol and case report forms. Any ad hoc or unplanned analyses performed and not identified in this SAP will be documented in the CSR.

2.1. RESPONSIBILITIES

Syneos Health will perform the statistical analyses and is responsible for the production and quality control of all tables, listings and figures (TLFs) under the direction of Nohla Therapeutics.

2.2. TIMING OF ANALYSES

The analyses for this study will be performed as follows:

- An interim analysis was performed when 80 subjects were enrolled and followed sufficiently long to assess their time to engraftment. An O'Brien-Fleming boundary was used at this interim analysis in order to maintain most of our type I error rate of 0.05 at the final analysis. A two-sided significance level of 0.0054 was used at this interim analysis, consistent with an O'Brien-Fleming boundary.

- A primary analysis will be performed after all subjects have completed 100 days follow-up post- Hematopoietic Cell Transplant (HCT). The clinical data cut date will take place 100 days beyond HCT for the last subject randomized. The analysis of the primary endpoint of ANC engraftment and the key secondary endpoints will be conducted at the time of the database lock for the primary analysis, ensuring complete capture of a minimum of 100 days and maximum of 2 years follow-up for each subject.

- Additional ad hoc analyses will be performed as needed. Any subsequent analyses of the primary and key secondary endpoints would be viewed as supportive analyses.

- A final supportive efficacy analysis is planned after all subjects have 2 years of follow-up or have discontinued the study.

3. INFORMATION FROM THE STUDY PROTOCOL

3.1. STUDY OBJECTIVES

3.1.1. Primary Objective

Compare the time to neutrophil engraftment (first of two consecutive days of an absolute neutrophil count [ANC] ≥500/μL) in subjects receiving a standard-of-care myeloablative cord–blood transplant (CBT) augmented with an off-the-shelf pre-expanded and cryopreserved cord-blood product to those who do not receive the product.

3.1.2. Secondary and Exploratory Objectives

Provide initial data on clinical and economic benefit, such as time to platelet engraftment, duration of initial hospitalization, transplant-related mortality (TRM), and death without engraftment. The kinetics of immune system recovery will also be evaluated in both arms.

3.2. BRIEF DESCRIPTION OF STUDY DESIGN

This is a Phase II multi-center, open-label, randomized study to compare outcomes following use of single or double myeloablative cord blood transplantation with or without infusion of off–the-shelf ex vivo expanded cryopreserved cord blood progenitor cells in subjects with hematologic malignancies.

3.2.1. Background

Cord blood (CB) is an effective and widely used source of stem cells for patients undergoing hematopoietic stem cell transplantation for hematologic malignancies. Most patients now receive two units of CB derived from different donors to better ensure provision of adequate stem cell numbers for reliable donor engraftment. However, with this modest two-fold increase provided by the second CB unit, the time to donor engraftment is still relatively delayed, averaging more than three weeks to achieve adequate numbers of myeloid cells. This leaves patients susceptible to infection and associated morbidity and mortality. Thus, although CB offers some distinct advantages over conventional stem cell sources as it is readily available with no donor attrition and allows reduced stringency in human leukocyte antigen (HLA) matching without an increase in graft-versus-host disease (GVHD) thereby increasing the donor pool especially for minority or mixed ethnicity patients, the increased risks of delayed engraftment or graft failure and early transplant related mortality that is associated with the low cell doses in a CB graft remains a barrier to the more wide spread use of this stem cell source. To address this barrier, research has investigated the role of the Notch signaling pathway in regulating ex vivo expansion of hematopoietic stem/progenitor cells with the goal of generating increased numbers of progenitor cells capable of rapid repopulation in vivo to improve the kinetics of hematopoietic recovery following CBT. This expansion is studied in this randomized Phase II clinical trial.

3.2.2. Planned Study Duration

A total of 160 subjects were enrolled into the study across 8 participating centers. Subjects enrolled to the study will be followed for 2 years post-HCT.

3.3. SUBJECT SELECTION

All subjects who sign the informed consent form, and have met all eligibility criteria, as confirmed by the Principal Investigator (PI) of the coordinating center, will be enrolled.

3.3.1. Inclusion Criteria

Ability to understand and the willingness to sign a written informed-consent document

1. Subject age
   - High-dose total-body irradiation (TBI) regimen: 6 months to ≤ 45 years
   - Middle-intensity TBI regimen: 6 months to ≤ 65 years

2. Diseases

A. Acute Myeloid Leukemia (AML), including Biphenotypic Acute Leukemia (BAL) or Mixed-Lineage Leukemia (MLL)
   - All subjects must have AML that is considered best treated by stem cell transplant by the referring physician and the attending transplant physician.
   - All subjects must be in complete response (CR) as defined by < 5% blasts by morphology/flow cytometry in a representative bone marrow sample with cellularity ≥ 15% for age.
   - Subjects in which adequate marrow/biopsy specimens cannot be obtained to determine remission status by morphologic assessment, but have fulfilled criteria of remission by flow cytometry, recovery of peripheral blood counts with no circulating blasts, and/or normal cytogenetics (if applicable) may still be eligible. Reasonable attempts must be made to obtain an adequate specimen for morphologic assessment, including possible repeat procedures. These subjects must be discussed with the Principal Investigator prior to enrollment.

B. Acute Lymphocytic Leukemia (ALL), including BAL or MLL
   - High-risk first complete response (CR1) [for example, but not limited to: t(9;22), t(1;19), t(4;11) or other MLL rearrangements, hyplodiploid] or HR as defined by referring institution treatment protocol, greater than 1 cycle to obtain CR; CR2 or greater.
   - All subjects must be in CR as defined by <5% blasts by morphology/flow cytometry in a representative bone marrow sample with cellularity ≥ 15% for age.

- Subjects in which adequate marrow/biopsy specimens cannot be obtained to determine remission status by morphologic assessment, but have fulfilled criteria of remission by flow cytometry, recovery of peripheral blood counts with no circulating blasts, and/or normal cytogenetics (if applicable) may still be eligible. Reasonable attempts must be made to obtain an adequate specimen for morphologic assessment, including possible repeat procedures. These subjects must be discussed with the Principal Investigator prior to enrollment.

C. Chronic myeloid leukemia (CML) excluding refractory blast crisis. To be eligible in first chronic phase (CP1) subject must have failed or be intolerant to tyrosine kinase inhibitor therapy.

D. Myelodysplasia (MDS) International Prognostic Scoring System (IPSS) Int-2 or High risk (i.e., refractory anemia with excess blast [RAEB], refractory anemia with excess blast in transformation [RAEBt]) or refractory anemia with severe pancytopenia or high-risk cytogenetics. Blasts must be < 10% by a representative bone marrow aspirate morphology.

3. Organ Function and Performance Status Criteria

A. Performance status score
- Karnofsky (≥ 16 years old) ≥ 70 or Eastern Cooperative Oncology Group (ECOG) 0-1
- Lansky (<16 years old) ≥ 60

B. Renal Function
- Adults: Calculated creatinine clearance must be > 60 mL and serum creatinine ≤ 2 mg/dL.
- Children (<18 years old): Calculated creatinine clearance must be > 60 mL/min

C. Hepatic Function
- Total serum bilirubin must be <3mg/dL unless the elevation is thought to be due to Gilbert's disease or hemolysis.
- Transaminases must be < 3x the upper limit of normal per reference values of referring institution

D. Pulmonary function
- Carbon Monoxide Diffusing Capacity (Lung Function), DLCO corrected >60% normal
- For pediatric subjects unable to perform pulmonary function tests, O2 saturation >92% on room air
- May not be on supplemental oxygen.

E. Cardiac function
- Left ventricular ejection fraction >45% OR
- Shortening fraction > 26%

3.3.2. Exclusion Criteria

A. Uncontrolled viral or bacterial infection at the time of study enrollment

B. Active or recent (prior 6 month) invasive fungal infection without ID consult and approval

C. History of HIV infection

D. Pregnant or breastfeeding

E. Prior myeloablative transplant containing full dose TBI (greater than 8 Gy)

F. Central nervous system (CNS) leukemic involvement not clearing with intrathecal chemotherapy and/or cranial radiation prior to initiation of conditioning. Diagnostic lumbar puncture is to be performed as per Section 9.7 of the study protocol.

G. Subjects ≥ 45 years; comorbidity score of 5 or higher.

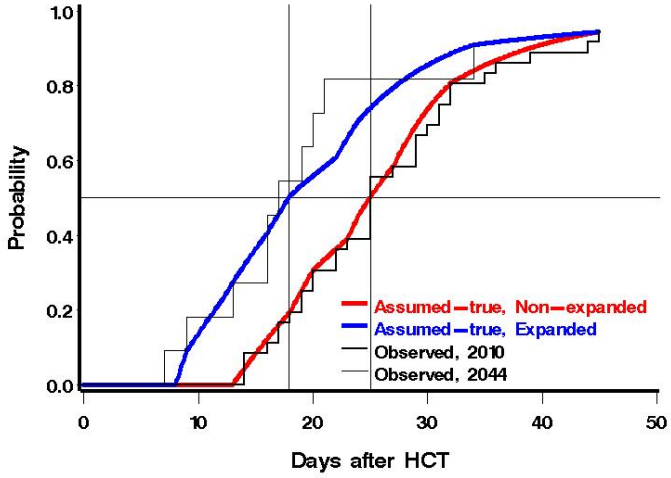
3.4. DETERMINATION OF SAMPLE SIZE

The clinical trial will randomize approximately 80 subjects per arm for a total of 160 subjects. The primary endpoint in this clinical trial, and hence the endpoint upon which the sample size/power are based, is ANC engraftment. This endpoint is a time-to-event outcome, with the occurrence of the event being the first of 2 consecutive days on which ANC≥500. Subjects who die without engraftment are included in the analysis as non-events at infinity.

To derive sample size, assumed-true distributions for engraftment were derived from data from previous studies for each of the arms, assuming piecewise exponential distributions for each arm. Each assumed-true distribution consisted of 9 windows, with the "hazard" of engraftment assumed to be constant within a window, but different across the windows. In general, the hazards of engraftment within each window were chosen so that the probability of engraftment at the end of the window matched the observed probability of engraftment at the corresponding time in the appropriate treatment group. The choice of 9 windows was somewhat arbitrary, but many choices were examined until the "fit" of the assumed-true distribution appeared to match the observed data. The observed data and the fit to the observed data, i.e., the assumed-true distributions, are shown in the figure below, which displays the observed time to engraftment for subjects who had received non-expanded CB (via protocol 2010) at FHCRC.

A piece-wise exponential curve was fit to these data, this fitted curve (in red) serving as the assumed-true time-to-engraftment distribution for subjects who receive a single or double CBT without ex vivo expanded off-the-shelf CB progenitor cells. The median time to engraftment associated with this assumed-true distribution is 25 days (the observed median

time to engraftment in this group is also 25 days). Shown in blue is the assumed-true time-to-engraftment distribution for the group to receive ex vivo expanded cells, the true median time being 17.9 days. This distribution is believed to be achievable, as it is consistent with what has been observed in previous studies using an expanded product at FHCRC.



The log-rank test was chosen, as this not only considers when subjects engraft, but time that subjects have been followed without engraftment. With 80 subjects per arm and the above assumed-true time-to-engraftment distributions, there will be approximately 87% power to observe a statistically significant difference (at the two-sided level of 0.05) in engraftment rates (power estimated from 5,000 simulations). Treating deaths without engraftment as non-events as infinity (rather than censoring) provides approximately 83% power.

3.5. TREATMENT ASSIGNMENT AND BLINDING

Approximately 160 subjects who meet inclusion/exclusion criteria will be enrolled and randomized in the study over multiple investigative sites. At the time of randomization, eligible subjects will be randomized into one of the following two treatment groups in a 1:1 ratio:

Experimental Arm: This arm will undergo Umbilical cord transplant (CBT) which consists of unmanipulated unit(s) [umbilical cord blood #1 or #2] followed by infusion of ex vivo expanded product at a minimum of 4 hours after completion of last unmanipulated unit.
Standard of Care (SOC) Arm: This arm will only undergo CBT which consists of unmanipulated unit(s) [umbilical cord blood #1 or #2].

The randomization will be done using a permuted block design with random block sizes and stratified on the number of cord blood units (1 vs. 2) to ensure balance in the number of subjects receiving 1 and 2 unexpanded CB units in each arm.

A central randomization list will be used to assign subjects to treatment.

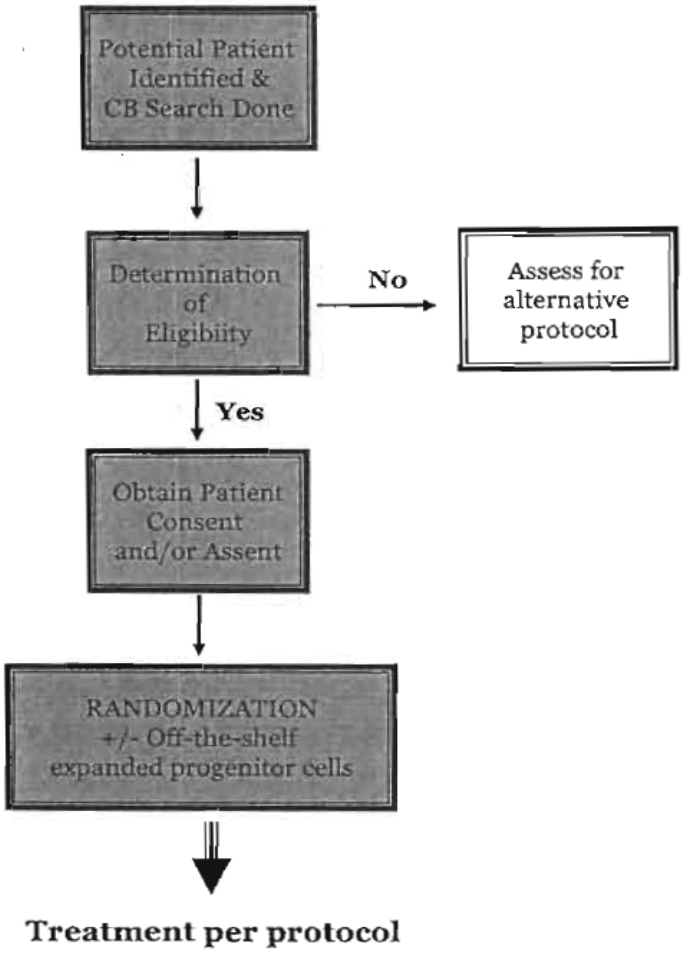
This is an open-label study because it is not possible to blind investigators or study subjects. The study endpoints are largely objective and clinically determined, but to reduce the potential for bias and maintain the study integrity Syneos Health will not provide Nohla Therapeutics with any analyses of aggregate data by treatment group. For all

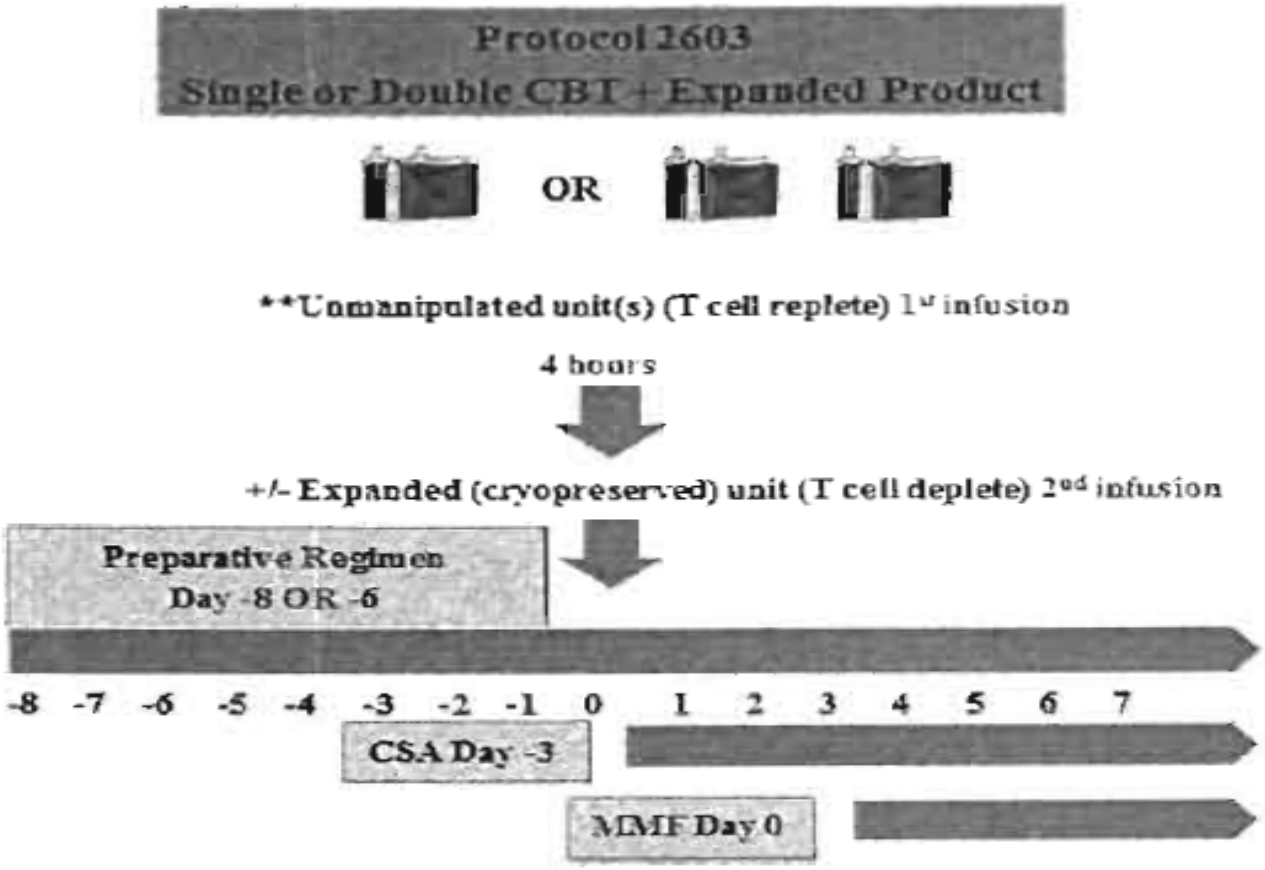
statistical deliverables such as SDTM/ADaM, analysis outputs for dry-runs, prior to the completion of the primary analysis, dummy treatment assignments will be used.

3.6. ADMINISTRATION OF STUDY TREATMENT

The unmanipulated units and the ex vivo expanded unit will be selected for CBT according to protocol Sections 6.1 and 6.2. The treatment schema at pre- and post-transplants are provided below.

Pre-transplant Determination of Eligibility and Randomization





Post-transplant analysis of donor engraftment: peripheral blood flow sorted for CD3, CD33, CD56, CD14, and CD19 on Days 7, 14, 21, 28, 42, 56, 80, 180, 365

****Single or double cord blood donor infusion based on donor selection criteria**

3.7. STUDY PROCEDURES AND FLOWCHART

Details of the visit structure and any time windows applied to these visits according to the schedule of visits in the protocol are provided below.

Study Evaluations	Prior to Conditioning	Days Post-Transplant											Long Term Follow-up
		Day 0	Daily to Engraftment	Weekly post Engraftment to Day 100	7 +/- 2 days	14 +/- 3 days	21 +/- 3 days	28 +/- 3 days	42 +/- 7 days	56 +/- 7 days	80 +/- 7 days	100 +/- 7 days	Day 180 +/- 30 days , 1 yr, 2 yr
Informed Consent	X												
Medical History, Physical Exam	X	X	X	X									X[4]
Karnofsky/ECOG/Lansky	X										D80-100		X[4]
PFT, MUGA/Echo, CT	X			PRN									PRN
CBC w/differential[8]	X	X	X	X									X[4]
Blood Chemistry	X	X	X	PRN								X	X[4]
Panel Reactive Antibody	X	Once at time of count recovery per investigator discretion											
Viral screening including CMV PCR	X												PRN
CMV Surveillance by PCR	X	As clinically indicated or per institutional guidelines for cord blood recipients											PRN
Urinalysis	X	X[7]											PRN
Infusion Toxicity		X											
Disease restaging- BM aspirate (+/- biopsy as indicated) within 30 days of conditioning	X							X			X		At 1 year[4]
Chimerism – Bone Marrow								X			X		X-At 1 yr then PRN[4]
Chimerism – Peripheral Blood					X	X	X[1]	X[1]	X[1]	X	X[2]	X[2]	X[4]
Immunophenotyping evaluation[6]	X							X		X		X	X[3,4]
TCR sample collection[6]	X							X		X		X	X[3,4]
TREC sample collection[6]								X		X		X	X[3,4]
IgA, IgG, IgM levels	X							X		X	X		X[3,4]
Lymphocyte Panel (TCSNK)								X		X		X	X[3,4]
GVHD Evaluation				X							X	X	X[4]
HCT-CI (for patients > 45 y/o)	X												
FHCRC only - Graft vs. Graft[5]	X[5] Investigator Discretion See Section 10.10												

Notes: [1] Day 21 and 42 to be done only if previous chimerism did not show 95% engraftment from a single cord blood unit. Day 28 may be skipped if the patient had Day 21 chimerism and it showed at least 95% donor engraftment from a single cord blood unit. [2]Day 80 may be performed at Day 100 at collaborating centers. Day 100 PB chimerisms do not need to be completed on Days 80 and 100 for FHCRC participants.

[3]Test done at Day 180, 1 and 2 year. [4] Every effort will be made to complete the 1 and 2 year evaluations as close to these dates as possible, taking into consideration the patient's circumstances. [5] See Section 10.10 for details of FHCRC-only research samples [6] Ship at time of collection. [7]Expanded product only. [8]CBC/differentials +/-4 hour window allowed

4. ENDPOINTS

4.1. PRIMARY EFFICACY ENDPOINT

Time to engraftment, defined as the first of 2 consecutive days with ANC ≥ 500/µL.

4.2. SECONDARY EFFICACY ENDPOINTS

1. Platelet engraftment, defined as the first of 7 consecutive days with platelets ≥ 20,000/µL without transfusions
2. Severe (grades III-IV) acute graft-versus-host disease (GVHD)
3. Overall Survival
4. Non-relapse mortality (hereafter referred to as transplant-related mortality, TRM)
5. Chronic GVHD defined by the 2014 NIH criteria

4.3. EXPLORATORY EFFICACY ENDPOINTS

1. In vivo persistence of the ex vivo expanded cord blood product
2. Duration of initial hospitalization
3. Number of intensive care unit (ICU)-free days within the first 100 days post HCT
4. Relapse/recurrence of the underlying malignancy, defined as follows:
    a. For subjects with a pre-transplant diagnosis of AML or ALL, relapse is defined as the presence of ≥ 5% leukemic blasts in the marrow or peripheral blood by morphology and/or flow cytometry. Extramedullary disease that requires initiation of induction chemotherapy, such as presence of leukemic blasts in the CNS or testicles, will also be considered relapsed disease. Similarly, subjects will also be considered to have relapsed disease if at any time point post-transplant they require initiation of chemotherapy to treat the underlying malignancy, even with < 5% blasts in the bone marrow or peripheral blood.
    b. For subjects with a pre-transplant diagnosis of Myelodysplastic Syndrome (MDS), recurrent/relapsed disease is defined as the presence of abnormal findings in the peripheral blood or bone marrow sufficient to diagnose MDS per French American System (FAB) or World Health Organization (WHO) classifications and requiring systemic therapy, or the presence of >10% leukemic blasts in the peripheral blood or bone marrow. The use of single agent chemotherapy (such as Azacitidine) as prophylaxis against disease recurrence will not be an indication of relapsed disease.
    c. For subjects with pre-transplant diagnosis of CML, relapsed disease is defined as leukemic blast crisis and/or accelerated phase with ≥ 5% leukemic blasts in the bone marrow or peripheral blood. Subjects with evidence of molecular or

cytogenetic disease alone is not considered disease relapse unless the disease is refractory to Tyrosine Kinase Inhibitors (TKIs).

5. Disease-free survival
6. GVHD-free, relapse-free survival (GRFS), with GVHD defined as either severe acute or chronic GVHD as defined by the 2014 National Institutes of Health (NIH) criteria
7. Common Terminology Criteria for Adverse Events (CTCAE) Grade ≥ 3 infusional toxicity, defined as grade 3+ events within the first 24 hours after completion of infusion of expanded cells
8. Graft failure: Primary and secondary
9. Immune reconstitution: kinetics of immune system recovery as measured by T and B cell subsets, T-cell Receptor Excision Circles (TREC), and T-cell Receptor (TCR) sequencing
10. Time to platelet engraftment of 50,000/µL, defined as the first of at least 2 separate days with platelets ≥50,000/µL that span a period of at least 7 days without intervening platelet transfusion.

4.4. SAFETY ENDPOINTS

Safety assessments will include incidences of adverse events (AEs), serious adverse events (SAEs), performance status, and laboratory values results (complete blood count, blood chemistry, immunoglobulins, and lymphocyte panel).

5. ANALYSIS SETS

5.1. SCREENED SET

The Screened Set will include all subjects screened. This set will be used for the summaries of subject disposition.

5.2. SAFETY SET

The Safety Set (SS) will be defined as all randomized subjects who undergo HCT. Subjects will be analyzed according to treatment they actually received. The SS will be used for all analyses of safety endpoints.

5.3. INTENT-TO-TREAT OR FULL ANALYSIS SET

The Intent-to-Treat Set or Full Analysis Set (FAS) will be defined as all randomized subjects regardless of whether they receive assigned treatment. Subjects will be analyzed according to randomized treatment. The FAS will be used for supportive analyses of primary, secondary and exploratory endpoints and for the presentation of subjects in all subject listings.

5.4. MODIFIED INTENT-TO-TREAT SET

The Modified Intent-to-Treat (mITT) Set will be defined as all randomized subjects who undergo HCT and received assigned treatment per randomization arm. Subjects will be analyzed according to randomized treatment. The mITT Set will be used as the primary analysis set for all analyses of primary, secondary and exploratory endpoints.

5.5. PER PROTOCOL SET

The Per Protocol Set (PPS) will include all subjects who are randomized, underwent HCT, have sufficient data to assess the primary efficacy endpoint, and who have no important protocol deviations (IPDs). Subjects will be analyzed according to treatment they actually received.

The important protocol deviations will be identified and defined prior to the primary analysis data lock and primary analysis. Any additional protocol deviations between primary analysis and final analysis will be identified and defined prior to the final data lock and analysis (after all subjects have 2 years of follow-up). The Nohla Therapeutics study team will identify and approve all IPDs categories and determinations. The IPD specifications will be defined prior to the database lock for primary analysis. The protocol deviations SAS datasets SDTM.DV and ADAM.ADDV will then be generated.

6. GENERAL ASPECTS FOR STATISTICAL ANALYSES

6.1. GENERAL METHODS

This SAP provides the formal specification of principal analyses to be conducted in this study. In the event of any discrepancies between this document and the study protocol, interpretation of results should be based on the specifications in the SAP. The protocol would be amended only if there would be changes to a principal feature of trial design, conduct or analyses.

Summaries and statistical analyses of study data will be performed by Syneos Health. Validated results will be reviewed by Consultant Biostatistician and clinicians and biostatistician at Nohla Therapeutics. All analyses will be performed by treatment group [Experimental Arm (EXP), Standard of Care Arm (SOC)]. Subject listings of all data from the case report forms (CRFs) as well as relevant derived variables will be presented.

All analyses will be implemented using SAS Version 9.4 or more recent version. The latest MedDRA version at the time of statistical analyses would be used to code medical history and adverse events datasets. Prior and concomitant medications will be coded using the latest version of the World Health Organization Drug Dictionary (WHO-DD) at the time of analysis.

Unless otherwise specified, continuous variables will be summarized by presenting the number of non-missing observations, mean, standard deviation, median, minimum and maximum, and distribution according to a clinically relevant discretization. Categorical variables will be summarized by presenting the number of subjects and percentage for each category.

Unless otherwise noted, all formal analyses of the primary and key secondary endpoints will be conducted in a manner to preserve the experimental two-sided 0.05 error rate, and two-sided 95% confidence intervals (CI) will be calculated. The interim analysis used 0.0054 alpha, so if the primary endpoint is statistically significant at the two-sided alpha level of 0.0492 (per the O'Brien-Fleming boundary), then formal comparisons will be conducted for the first 2 of the 5 secondary endpoints – platelet engraftment and severe acute GVHD. Using a Bonferroni adjustment for multiplicity, statistical significance for each of the first 2 secondary endpoints analyses will be based on an alpha of 0.0246 (=0.0492/2). The other 3 secondary endpoints (overall survival, TRM, and chronic GVHD) which may require longer follow-up, will be treated similar to exploratory endpoints, with reporting of nominal p-values for descriptive purposes only.

If major imbalances (due to chance alone) are identified across treatment groups for factors known to influence outcome, supportive analyses will be conducted adjusting for these imbalances using appropriate regression techniques (linear for continuous, logistic for binary, and Cox for time-to-event outcomes). Sensitivity analyses will be conducted in an exploratory manner to examine the consistency of effects across various subgroups as

defined in Section 6.5. When parametric analyses are conducted, in the event of highly skewed data, sensitivity analyses will be conducted using, for example, a logarithmic transformation to reduce skewness.

P-values will be presented for the pre-specified analyses of primary, secondary, and exploratory endpoints conducted as specified in Sections 9.7, 9.8, and 9.9, respectively. P-values will be rounded to 3 significant digits.

All efficacy analyses will be conducted from date of HCT, with the mITT population as the primary population. Timing of assessments will be stated relative to date of HCT (e.g., 100 days of follow-up following HCT). Sensitivity analyses will be conducted for engraftment endpoints from date of randomization using the full analysis population.

6.2. KEY DEFINITIONS

6.2.1. Baseline

Baseline is the last measurement taken prior to or on the day of HCT, unless otherwise specified.

6.2.2. Protocol Day

Per the study protocol, Protocol Day 0 is regarded as the date of Hematopoietic Cell Transplant (HCT).

6.2.3. Study Day

The reference date is the day of HCT, considered as Protocol Day 0. Study day will be calculated as (date of assessment – date of HCT + 1) if assessment date is on or after the date of HCT; or date of assessment – date of HCT, if assessment date is prior to the date of HCT. Therefore, Protocol Day 0 is equal to Study Day 1.

6.2.4. Completed Study

Study completion for the primary analysis is met if one of the following apply:

- A subject has completed the two-year follow up visit
- A subject remains in follow-up (i.e., has not withdrawn) at the time of the clinical cut date
- A subject has died prior to the clinical cut date

Study completion for the final analysis is met if one of the following apply:

- A subject has completed the two-year follow up visit
- A subject has died prior to the clinical cut date

6.3. PARTIAL DATES

Data points that appear to be spurious will be investigated and will not be excluded from the listings. Influential cases will be handled in an appropriate statistical manner.

All dates recorded on the Adverse Events (AEs) and prior/concomitant medication CRF should be complete. However, if there are records for which complete start or stop dates could not be obtained following best effort attempts for query resolution, partial or incomplete start/end dates in the AEs or prior/concomitant medication datasets will be imputed to help determine treatment-emergent adverse events (TEAE) and (prior or concomitant) medications.

For incomplete start dates, use the following imputation rules:

- Missing day - Impute the 1st of the month unless month is same as month of HCT then impute using the day of HCT.
- Missing day and month – impute 1st January unless year is the same as year of HCT then impute using the month of HCT for missing month, and the first day of the month of HCT for missing day.
- Completely missing – Impute HCT date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the start date.

When imputing a start date, ensure that the new imputed date is sensible i.e. is prior to the end date of the AE or medication.

For incomplete concomitant data end dates only (not applicable to AE datasets), use the following imputation rules:

- Missing day - Impute the last day of the month
- Missing day and month – Impute 31st December and reconcile with the discontinuation date or death date. If the imputed date with 31st December is greater than the discontinuation date or date of death, then use the discontinuation date. For subjects who have not discontinued, check if the imputed date with 31st December is greater than their last available visit date, if so then use their last available visit date.
- Completely Missing – Determine if the medication is still ongoing before imputing a date and when the medication started in relation to date of HCT. If the ongoing flag is not missing, then assume that the medication is still being taken (i.e. do not impute a date). If the medication has stopped and start date is prior to HCT date then impute the HCT date, if the medication started on or after HCT date then impute a date that is after the last available visit date.
- Should any of the previous stop dates created come before a start date, either a complete date or an imputed one, use the start date instead of the date that would otherwise be created.

6.4. VISIT WINDOWS

Details of the planned protocol-specified visits and visit windows are given in the protocol and in Section 3.7.

No derivation of analysis windows are planned for the study regarding data collected outside the protocol-specified visit windows. Study data will be analyzed according to the scheduled visits reported.

Where assessments are carried out more than once at a visit or within the same window, data and timings of all assessments would be reported in the data listings, including assessments that may have occurred from unscheduled visits. However, for data analysis in the tables, the closest value to the protocol-specified visit will be used. If there is a tie in determining the closest value, then the data value just prior to the protocol-specified visit will be used.

6.5. EXAMINATION OF SUBGROUPS

Subgroups will be defined only if there are a sufficient number of subjects in each category of the subgroup (e.g., >5 subjects per treatment group).

Subgroup analyses will be descriptively conducted to assess treatment effects and include, but are not limited to, the following:

- Age groups of < 18 versus 18-65 years;
- Weight (below the median for the FAS versus above the median for the FAS)
- Receipt of single versus double CBT;
- FHCRC versus non-FHCRC sites;
- Conditioning regimen (High intensity: fludarabine, cyclophosphamide, total body irradiation versus Mid Intensity: fludarabine, cyclophosphamide, total body irradiation, Thiotepa, Other); and
- Underlying disease (AML, ALL, CML, MDS, Other).

The primary and key secondary efficacy endpoints will be analyzed for the subgroups using the mITT Set. Forest plots will be presented. In addition, select safety analyses will also be analyzed by subgroup using the SS as specified in Section 10. Subgroup characteristics will be summarized by treatment group. During the CSR preparation, additional subgroup analyses may be performed. Those analyses will be considered as ad hoc analyses and will be presented and discussed in the CSR.

7. STUDY POPULATION

7.1. SUBJECT DISPOSITION

The following subject data will be presented:

- Number of subjects screened, signed informed consent (overall only);
- Number of subjects screened failed and reasons for screen failure (overall only);
- Number and percentage of subjects randomized in each analysis set (SS, FAS, mITT, PPS), by treatment group and overall;
- Number and percentage of subjects completed or discontinued study, by primary reason for study discontinuation for FAS and mITT Set, by treatment group; and
- Number and percentage of subjects excluded from PPS by reason for exclusion defined in Section 5.5, by treatment group for FAS.

A subject listing for early discontinuations and reasons for discontinuation will be provided for the FAS.

7.2. PROTOCOL DEVIATIONS

Protocol deviations will be collected by study monitors. Sufficient detail on the types of protocol deviations may include but are not limited to

- Subject did not meet all of the inclusion or exclusion criteria
- Subject used excluded medications during the study
- No post-HCT efficacy data
- Any known major protocol non-compliance

Important protocol deviations that could potentially affect the efficacy or safety conclusions of the study will be identified prior to database lock. Important protocol deviations will be summarized by deviation category and treatment group using the FAS and mITT Set.

All protocol deviations will be presented in a data listing.

In addition, prior to database lock, a data listing will be prepared for sponsor review in order to determine the subjects to exclude from the PPS, as well as categorize each protocol deviation criteria as not important or important.

8. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATIONS

8.1. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Age (years) will be calculated as the number of years between the date of birth and the date of randomization.

Age at randomization = floor (randomization date – date of birth +1)/365.25 and truncated to complete years.

If height is recorded in inches (in), then height will be summarized in centimeters (cm) using the following conversion:

Height (cm) = Height (in) x 2.54

If weight is recorded in pounds (lb), then weight will be summarized in kilograms (kg) using the following conversion:

Weight (kg) = Weight (lb) x 0.4536

All demographic and baseline characteristics will be summarized by treatment group using descriptive statistics for the FAS, mITT Set, and SS. No formal hypothesis testing will be performed to compare differences between treatment groups.

8.1.1. Recipient Demographic

The following demographic variables for the recipient will be summarized:

- Age (years) at time of randomization, sex, ethnicity, race, height (cm), and weight (kg).

8.1.2. Donor Demographic

The following demographic variables for the donor will be summarized:

- Cord blood storage (years) at time of transplantation as determined by donor age, and donor sex.

8.1.3. Significant Pre-Transplant Medical History

The number and percentage of subjects with previous and active medical or surgical conditions at screening will be summarized and presented by MedDRA system organ class and preferred term.

8.1.4. Significant Pre-Transplant Treatment/Disease History

The following will be summarized:

- Number and percentage of subjects with each type of previous therapy.
- Result/Outcome for each previous therapy [Complete Response (CR), Partial Response (PR), Stable Disease (SD), Progressive Disease (PD), Relapse (REL) and Refractory (REFR)].
- Duration of time since last treatment prior to randomization (months).

8.1.5. Disease Characteristics at Diagnosis

The following disease characteristics as recorded on the CRF will be summarized:

- Diagnosis (AML, ALL, CML, MDS, other)
- Duration of disease prior to randomization (months) defined as (randomization date – date of diagnosis)/(365.25/12)
- Cytology Results
- Morphology Results
- Flow
- Disease specific characteristics for AML, ALL, CML, MDS, other

8.1.6. Transplant-Related Information

The following transplant-related information will be summarized:

- Number and percentage of subjects who underwent previous transplants
- Recipient viral serology status [Cytomegalovirus , Herpes Simplex Virus, Varicella Zoster Virus, antibody titer-A [Immunoglobulin G (IgG), Immunoglobulin M (IgM)], antibody titer-B (IgG, IgM)
- Recipient Viral Reactivations: Virus, reactivation site, first Polymerase Chain Reaction (PCR) copy # (unit), max PCR copy #

- ABO compatibility for each recipient/donor pair (A-A, A-<specify per data>, B-B, B-<specify per data>, AB-AB, AB-<specify per data>, and O-O, O-<specify per data>)
- Recipient and Donor Rhesus (Rh)
- Recipient Panel reactive antibody [class I (%), class II (%)]
- HLA Subject/Donor and [HLA match (A, B, DR), Locus of mismatch (C if available), and locus of mismatch (DQB1 if available) for Unit #1 and Unit #2 respectively]
- Hematopoietic Cell Transplantation –Specific Comorbidity Index (HCT-CI)
- Abnormal donor findings, (Yes/No)
- Donor red cell depleted, (Yes/No)

8.1.7. Performance Status

The following will be summarized:

- Type of performance status assessment (ECOG, Lansky, Karnofsky)
- Results

8.1.8. Cardiovascular and/or Pulmonary Function Tests (PFT)

The following pre-transplant cardiovascular and/or PFT results will be summarized:

- PFT Tests (DLCO and Oxygen saturation)
- Echocardiogram/Multigated Acquisition (Ejection fraction and fractional shortening)

8.1.9. Conditioning Regimen

The following conditioning regimen variables will be summarized:

- Treatment (Fludarabine, Cyclophosphamide, TBI, Thiotepa, Other)
- Dose (reported as total dose) and units

8.1.10. Disease Evaluation at Time of Transplant

The following variables will be summarized:

- Weight (kg)
- Disease status at transplant [CR1, CR2, >= CR3, PR, CP1 (CML only), CP2 s/p AP (CML only), CP2 s/p BC (CML only), CP>=3 (CML only), AP1 (CML only), AP2 s/p BC (CML only), AP >=3 (CML only), CR (CML only), Other
- Disease Response status at transplant (complete remission, partial remission, stable disease, progressive disease)
- Minimal residual disease (Yes/No/Unknown/Not applicable)
- Cytology results
- Morphology
- Flow
- FISH
- Molecular characterization
- Imaging scan type [Cat scan (CT), X-ray, Magnetic Resonance Imaging (MRI)] and results

8.1.11. Donor Cell Panel

The following variables will be summarized:

- Test results [CD34 (×10^6), CD34kg (×10^6), TNC (×10^7), TNCkg (×10^7), TVOL], viability (%) for Unit #1, Unit #2, and expanded product.

8.2. MEDICAL HISTORY AND CONCOMITANT DISEASES

Significant pre-transplant medical history will be collected. The number and percent of subjects with each medical history condition will be summarized by MedDRA system organ class and preferred term using the mITT Set by treatment group. All medical history data will be listed.

8.3. MEDICATIONS

Medications received and stopped before HCT day will be considered as prior medications. Medications that stopped or are ongoing after HCT through to the end of study or early termination will be considered as concomitant medications. The medications will be summarized for each treatment group by therapeutic subgroup (ATC 2nd level) and preferred WHO-DD name for the mITT Set. Subjects taking the same medication multiple times will be counted once per medication.

Handling missing or incomplete start or stop date in defining concomitant medications is described in Section 6.3.

All medication data will be listed, sorted by investigative site treatment group, subject number, start and stop date. Information listed will include medication, indication, total daily dose, dose units, frequency and route of administration.

Descriptive statistics of the number of antibiotic-free days within the first 100 days for subjects will also be provided.

9. EFFICACY

All efficacy analyses will be conducted using the mITT Set as the primary analysis. If the total FAS or PPS differ from the mITT Set by more than 10 subjects, efficacy analyses will be repeated using these analysis sets for relevant primary, key secondary, other secondary efficacy endpoint analyses, and select exploratory efficacy endpoints. Analyses using the FAS will use the date of randomization, instead of date of HCT, as the reference date. All efficacy endpoints, recorded and derived, will be presented in data listings. The analyses of primary and key secondary endpoints will be repeated for the subgroups as specified in Section 6.6.

9.1. ADJUSTMENTS FOR COVARIATES

N/A.

9.2. HANDLING OF DROPOUTS OR MISSING DATA

No imputation will be done for missing efficacy data. Observed data will be used for analyses.

9.3. INTERIM ANALYSES AND DATA MONITORING

9.3.1. Interim Analysis

For the primary endpoint of time to engraftment (ANC ≥ 500), an interim analysis for superiority was conducted when 80 subjects have been enrolled and followed sufficiently long to assess their time to engraftment. An O'Brien-Fleming boundary was used for the interim analysis in order to maintain most of our type I error rate of 0.05 at the final analysis. In particular, if there is a statistically significant difference in favor of the experimental regimen at the two-sided significance level of 0.0054, the boundary was to be considered crossed at the interim analysis. At the primary analysis, analysis of this endpoint will be conducted at a two-sided significance level of 0.0492 (per the O'Brien-Fleming boundary).

To enable the trial to provide more reliable evidence with respect to other outcomes, stopping for superiority at the interim analysis as defined above was to occur only if the O'Brien-Fleming boundary also is crossed for the endpoint of overall survival, but at the overall two-sided 0.01 level of significance.

9.3.2. Safety Monitoring Committee

A separate Data Safety Monitoring Board (DSMB) will receive deliverables based on actual treatment assignments. Additional details on the role of the DSMB are provided in a separate charter.

9.4. MULTICENTER STUDIES

The randomization is not stratified by site. Likewise, analyses of efficacy data will not be stratified by study site. The number and percentage of subjects randomized by study site will be summarized by treatment group and for all subjects.

9.5. USE OF AN "EFFICACY SUBSET" OF SUBJECTS

N/A

9.6. MULTIPLE COMPARISONS/MULTIPLICITY

The interim analysis used 0.0054 alpha, so for the final analysis of the primary and key secondary efficacy analyses, the overall 2-sided level of significance will be alpha=0.0492. The hypothesis testing of key secondary endpoints will be conducted by splitting the 0.0492 alpha equally among the 3 key secondary endpoints (0.0164), provided the primary efficacy endpoint comparison is statistically significant at an alpha level 0.0492. If this comparison is not statistically significant, then the comparison of key secondary efficacy endpoints will be considered nominal, descriptive and exploratory. This procedure controls the study-wise type I error.

9.7. PRIMARY EFFICACY ANALYSIS

The primary endpoint of ANC engraftment is assessed daily until a subject engrafts, rejects their graft, or reaches day 55 following HCT, whichever comes first. Per convention in HCT, ANC engraftment is defined as the first of 2 consecutive days in which neutrophil count ≥500/μL.

The analysis of the primary endpoint (ANC engraftment) will be conducted by treating engraftment as a time-to-event outcome from date of HCT. A stratified log-rank test will be used for the primary analysis of this endpoint, with stratification on number of cord-blood units (1 vs. 2), reflecting the randomization stratification variable. Subjects who die without engraftment will be included in the analysis as non-events at infinity (rather than censoring). Subjects who reject the graft or reach day 55 following HCT without engraftment will be censored at day 55.

Kaplan-Meier (KM) estimates of the distribution of the time-to-event will be tabulated and graphed by treatment group. The tabulation will include the KM estimate of the medians, 25^{th} and 75^{th} quartiles, and corresponding 95% CIs, if estimable. The tabular and graphical summaries will include the at-risk counts for every visit. The number and percent of subjects censored and with events will be presented. The hazard ratio and 95% CI will be determined based on a Cox regression model stratified by randomization stratification variable to estimate the magnitude of the effect.

The primary analysis will be conducted in days, from date of HCT, but a sensitivity analysis will be conducted from date of randomization. A sensitivity analysis will also be done by censoring deaths without engraftment at time of death.

9.8. SECONDARY EFFICACY ANALYSES

The secondary endpoints include:

1. Platelet engraftment, defined as the first of 7 consecutive days with platelets ≥ 20,000/μL without transfusions
2. Severe (grades 3 or higher) acute graft-versus-host disease (GVHD)
3. Overall Survival
4. Transplant-related mortality(TRM)
5. Chronic GVHD defined by the 2014 NIH criteria

The first 2 secondary endpoints listed above (platelet engraftment, severe acute GVHD) will be formally compared between treatment groups and are therefore considered "key" secondary endpoints as described below in Section 9.5.1. The remaining 3 secondary endpoints will be considered "other" secondary endpoints as described below in Section 9.6.

9.8.1. Key Secondary

The key secondary endpoints include:

1. Platelet engraftment, defined as the first of 7 consecutive days with platelets ≥ 20,000/μL without transfusions
2. Severe (grades 3 or higher) acute graft-versus-host disease (GVHD)

The endpoints listed above (platelet engraftment, severe acute GVHD) will be formally compared between treatment groups, provided the primary endpoint of ANC engraftment is deemed to be statistically significantly different between treatment groups (at the two-sided significance level of 0.0492) as noted above in Section 6.1. In such case, we shall conduct formal analyses where the two-sided significance level of 0.0492 will be shared equally among these three secondary endpoints that will undergo formal comparison. Therefore, the two-sided level of significance for these formal comparisons will be 0.0246. The key secondary endpoints of platelet engraftment and severe acute GVHD will be conducted from time of HCT, with a sensitivity analysis conducted from date of randomization.

9.8.1.1. Key Secondary Efficacy Endpoint 1: Platelet Engraftment

The secondary endpoint platelet engraftment is assessed and analyzed in the same manner as the primary endpoint of ANC engraftment (Section 9.7). Platelet engraftment, per convention in HCT, is defined as the first of seven consecutive days in which platelets ≥20,000/μL in the absence of platelet transfusions. Subjects who reject the graft or do not achieve engraftment by day 100 will be censored at that date.

9.8.1.2. Key Secondary Efficacy Endpoint 2: Severe (grade 3 or 4) Acute GVHD

GVHD is defined according to the National Institutes of Health (NIH) dictionary of cancer terms as a disease caused when cells from a donated stem cell graft attack the normal tissue of the transplant subject. Symptoms include jaundice, skin rash or blisters, a dry mouth, or dry eyes.

To reduce potential bias, a single central reader, Dr. Paul Martin will grade all subjects for acute GVHD. Every effort will be made to shield Dr. Martin from information that may contain reference to treatment assignment. The grading is based on data collected within the first 100 days following HCT and is assessed per the acute GVHD assessment outlined in protocol Appendix C. Dr. Martin has been the sole grader of acute GVHD at FHCRC for more than 20 years.

Severe (grades 3 or 4) acute GVHD will be analyzed using the Cochran-Mantel-Haenszel test (stratified on number of cord-blood units) to compare the proportion of subjects with severe GVHD between treatment groups. The number and percentage, with associated two-sided exact (Clopper-Pearson) 95% CIs, of subjects with and without severe GVHD will be presented by treatment group and an odds ratio calculated (SOC arm/experimental arm).

9.8.2. Other Secondary Efficacy Analyses

The other secondary endpoints include:

1. Overall Survival
2. Transplant-related mortality(TRM)
3. Chronic GVHD defined by the 2014 NIH criteria

Other secondary endpoints will be compared in a manner consistent with exploratory endpoints as they may require longer follow-up time. Nominal p-values will be provided for descriptive and exploratory purposes only. The endpoints of overall survival, TRM, and chronic GVHD will be conducted in days using date of HCT, with minimum follow-up of 100 days and maximum follow-up of 2 years after HCT. Sensitivity analyses will be conducted using date of randomization and the FAS.

9.8.2.1. Other Secondary Efficacy Endpoint 1: Overall Survival

Overall survival will be captured until the end of follow-up for each subject, defined as 2 years after HCT. The analysis of overall survival will occur at the clinical cut date (100 days after HCT for the last randomized subject) if at least 30 total deaths have occurred by this point. Otherwise, the analysis of overall survival will occur after the clinical cut date at the point once the 30th death occurs. Subjects alive at the time of analysis will be censored at the date last known to be alive (up to two years). Subjects lost to follow-up prior to two years will be censored at their last study date known to be alive. Overall survival will be compared between groups using the same analysis methods specified for the primary endpoint (Section 9.7).

9.8.2.2. Other Secondary Efficacy Endpoint 2: Transplant-Related Mortality (TRM), Previously Referred to as Non-relapse Mortality (NRM)

TRM is defined as the occurrence of death without prior relapse. Relapse is defined in Section 4.3 under the 4th exploratory endpoint. Subjects who relapse will be censored at the date of first relapse. TRM will be analyzed using the same analysis methods specified for the primary endpoint (Section 9.7).

9.8.2.3. Other Secondary Endpoint 3: Chronic GVHD

The cause-specific hazards of chronic GVHD, defined according to the 2014 NIH criteria, will be compared between treatment groups using the stratified log-rank test, stratified for number of cord-blood units (1 vs. 2), reflecting the randomization stratification variable. Chronic GVHD will be captured for each subject until end of follow-up, defined as 2 years after HCT. Deaths without GVHD will be censored at time of death. In addition, the cumulative incidence functions of chronic GVHD will be compared using Gray's test[4, 5, 6], with death without chronic GVHD regarded as a competing-risk event[5].

9.9. EXPLORATORY EFFICACY ANALYSES

The exploratory endpoints include:

1. In vivo persistence of the ex vivo expanded cord blood product
2. Duration of initial hospitalization
3. Number of ICU-free days within the first 100 days post HCT
4. Relapse/recurrence of the underlying malignancy, defined as follows:
    a. For subjects with a pre-transplant diagnosis of AML or ALL, relapse is defined as the presence of ≥ 5% leukemic blasts in the marrow or peripheral blood by morphology and/or flow cytometry. Extramedullary disease that requires initiation of induction chemotherapy, such as presence of leukemic blasts in the CNS or testicles, will also be considered relapsed disease. Similarly, subjects will also be considered to have relapsed disease if at any time point post-

transplant they require initiation of chemotherapy to treat the underlying malignancy, even with < 5% blasts in the bone marrow or peripheral blood.

b. For subjects with a pre-transplant diagnosis of MDS, recurrent/relapsed disease is defined as the presence of abnormal findings in the peripheral blood or bone marrow sufficient to diagnose MDS per French American System (FAB) or World Health Organization (WHO) classifications and requiring systemic therapy, or the presence of >10% leukemic blasts in the peripheral blood or bone marrow. The use of single agent chemotherapy (such as Azacitidine) as prophylaxis against disease recurrence will not be an indication of relapsed disease.

c. For subjects with pre-transplant diagnosis of CML, relapsed disease is defined as leukemic blast crisis and/or accelerated phase with ≥ 5% leukemic blasts in the bone marrow or peripheral blood. Subjects with evidence of molecular or cytogenetic disease alone is not considered disease relapse unless the disease is refractory to TKIs.

5. Disease-free survival
6. GVHD-free, relapse-free survival (GRFS), with GVHD defined as either severe acute or chronic GVHD as defined by the 2014 National Institutes of Health (NIH) criteria
7. CTCAE Grade ≥ 3 infusion toxicity, defined as grade 3+ events within the first day after completion of infusion of expanded cells (experimental arm only)
8. Graft failure: Primary and secondary
9. Immune reconstitution: kinetics of immune system recovery as measured by T and B cell subsets, TREC, and TCR sequencing
10. Time to platelet engraftment of 50,000/μL, defined as the first of at least 2 separate days with platelets ≥50,000/μL that span a period of at least 7 days without intervening platelet transfusion

9.9.1. Exploratory Efficacy Endpoint 1: In Vivo Persistence of the Ex-Vivo Expanded Cord Blood Product

Persistence of the ex vivo expanded product is assessed at days 28, 80, and one year for Bone marrow (BM) sample for Unsorted Chimerism; weekly through day 28, then days 42, 56, 80, 100, 180, 1 and 2 years for Peripheral Blood sample for Sorted Chimerisms.

In vivo persistence of the expanded cord blood product will be summarized using descriptive statistics. Results will be presented by individual cell immunophenotype (CD3+, CD14+ or CD15+, CD33+, CD19+ or CD20+, CD56+). This endpoint only pertains to the experimental arm.

9.9.2. Exploratory Efficacy Endpoint 2: Duration of Hospitalization

Duration of initial hospitalization is assessed daily.

Duration of initial hospitalization will be presented, with particular attention paid to subjects who die without being discharged. While this is expected to be uncommon, to avoid a potential artificial reduction in hospitalization length due to "early" deaths while in the hospital, the effect of treatment on the number of days alive and out of the hospital during the 100 days post HCT will also be evaluated.

The duration of initial hospitalization will be compared between treatment groups using an analysis of variance (ANOVA) with fixed effects for treatment groups and randomization stratification. Number of days alive and out of the hospital will be analyzed in the same manner.

9.9.3. Exploratory Efficacy Endpoint 3: Intensive Care Unit-Free Days

The effect of treatment on the number of days alive and out of the ICU during the first 100 days post HCT will be evaluated in a manner consistent with the duration of initial hospitalization endpoint (Section 9.9.2).

9.9.4. Exploratory Efficacy Endpoint 4: Relapse/Recurrence of the Underlying Malignancy

The cause-specific hazards of relapse will be analyzed using the same analysis methods specified for the primary endpoint (Section 9.7). Relapse will be captured for each subject until end of follow-up, defined as 2 years after HCT. Deaths without relapse will be censored at time of death. In addition, the cumulative incidence functions[6] of relapse will be compared using Gray's test[4,5], with death without relapse regarded as a competing-risk event. Subjects without relapse will be censored at the last known date alive up to 2 years follow-up.

9.9.5. Exploratory Efficacy Endpoint 5: Disease-Free Survival

Disease-free survival is defined as survival without the occurrence of relapse. This endpoint will be analyzed using the same analysis methods specified for the primary endpoint (Section 9.7). Subjects without relapse will be censored at the last known date alive up to 2 years follow-up.

9.9.6. Exploratory Efficacy Endpoint 6: GVHD-free, Relapse-free Survival (GRFS)

GVHD-free, relapse-free survival (GRFS), defined as survival without relapse, severe acute GVHD, or chronic GVHD, will be analyzed using the same analysis methods specified for the primary endpoint (Section 9.7). Subjects without failure at the time of database lock will be censored at the last known date alive.

9.9.7. Exploratory Efficacy Endpoint 7: Grade ≥ 3 Infusion Toxicity

Grade 3 infusion toxicity is defined as the occurrence of grade 3 events within the first day following completion of the expanded cell infusion and be presented as a simple proportion as well as table with MedDRA coded AE terms by system organ class and preferred term. This endpoint will only be described for the experimental arm.

9.9.8. Exploratory Efficacy Endpoint 8: Graft Failure

Graft failure is assessed daily within the first 55 days following HCT. Graft failure will be estimated using simple proportions, with graft failure (both primary and secondary) as defined below:

Primary Graft Failure: Subjects may be considered primary graft failure/rejection provided they meet any criteria listed below:

- Absence of 3 consecutive days with neutrophils ≥ 500/µL combined with host CD3 peripheral blood chimerism ≥ 50% by day + 42
- Absence of 3 consecutive days with neutrophils ≥ 500/µL under any circumstances at day 55
- Death after day 28 with neutrophil count < 100 µL without any evidence of engraftment (< 5% donor CD3)
- Primary autologous count recovery with < 5% donor CD3 peripheral blood chimerism at count recovery and without relapse

Secondary Graft Failure/Rejection: Decline of neutrophil count < 500/µL with loss of donor chimerism after day 55

Primary and secondary graft failure will be analyzed separately using the same analysis methods specified for the severe acute GVHD endpoint (Section 9.8.1.3).

9.9.9. Exploratory Efficacy Endpoint 9: Immune Reconstitution

Immune reconstitution will be examined pre-conditioning, days 28, 56, 80 or 100 (per institutional practice), 6 months, 1 and 2 years.

Immune reconstitution will be assessed graphically by plotting chimerism values longitudinally at the time points listed above, with the "average" values obtained by fitting non-linear continuous curves through the longitudinal data (e.g., using restricted cubic splines).

9.9.10. Exploratory Efficacy Endpoint 10: Time to Platelet Engraftment ≥50,000/µL

As another measure of evaluation of kinetics of immune reconstitution, time to platelet engraftment to ≥50,000/µL will be evaluated. This endpoint is defined as the first of at

least 2 separate days with platelets ≥50,000/µL that span a period of at least 7 days without intervening platelet transfusion. This endpoint will be analyzed using the same analysis methods as the secondary endpoint of time to platelet engraftment to ≥20,000/µL (Section 9.8.1.1).

10. SAFETY

Safety assessments will consist of summarizing transplant and infusion details; all AEs collected as per the protocol, including SAEs,; and laboratory values including (complete blood count, blood chemistry, immunoglobulin, and lymphocyte panel).

All safety analyses will be conducted on the Safety Set, unless specified otherwise. Summaries of all safety assessments will be reported at visits where data are collected. All safety data will be presented in individual subject data listings.

10.1. EXTENT OF EXPOSURE AND FOLLOW-UP

10.1.1. Exposure

The following exposure variables will be summarized continuously for each treatment using the SS:

- Duration of infusion (minutes) during transplant for Unit 1 and Unit 2,
- Volume (mL) infused for Unit 1 and Unit 2,
- Volume (mL) of ex vivo expanded product infused
- Duration (minutes) of infusion of ex vivo expanded product

10.1.2. Follow-up

The duration of follow-up for each subject will be calculated as the length of time from date of HCT to date of completion/discontinuation from study in days. Similarly, the duration from randomization to HCT will be calculated in days. Both will be summarized by treatment group as continuous variables with descriptive statistics.

Duration of follow-up will also be categorized as follows:

- ≥ 6 months
- ≥ 1 year
- ≥ 2 years

These categories are not mutually exclusive. Counts and percentages of subjects in each of these categories will be summarized for each treatment group using the SS.

10.2. TREATMENT COMPLIANCE

N/A.

10.3. ADVERSE EVENTS

See the protocol Section 12 for the definition of adverse events (AEs), serious AEs, related/unrelated AEs, and other definitions relating to AEs. See protocol appendix D for a list of potential adverse events associated or expected with hematopoietic cell transplantation.

Adverse events (AEs) defined by study protocol will be collected on the CRF and coded using the MedDRA coding dictionary version 21.0 or later. Severity will be graded using the CTCAE Version 4.0.

A treatment-emergent adverse event (TEAE) is any untoward medical occurrence that either occurs or worsens at any time on or after HCT and that does not necessarily have to have a causal relationship with the treatment or surgical procedure. Please see Section 6.3 for handling missing or incomplete dates in determining whether an AE is a TEAE. For analyses purposes, only TEAEs occurring during the first 100 days after HCT will be presented in summary tables.

All collected AEs and SAEs will be presented in individual listings. The listing will contain the following information: treatment group, verbatim term, system organ class, preferred term, related to the expanded cell product, relationship to transplant, date and day of onset, date and day of resolution, treatment given to treat the adverse event, the outcome, whether the event was an SAE, CTCAE grade, immediate toxicity to Unit 1/Unit 2, immediate toxicity to ex vivo expanded product, and whether it is a TEAE. Listings will be sorted by subject identification number, onset date, system organ class, and preferred term. If the onset date is completely missing, then these events will be presented first. If the onset date is missing a month or a day, then these events will be presented before any complete dates.

Attribution - The following are definitions for determining whether an adverse event is related to a treatment or procedure:

- An AE is considered “related” to treatment or “related” to research procedures if in the opinion of the principal investigator, it was considered possibly, probably, or definitely related on the CRF.
- An AE is considered “unrelated” if in the opinion of the principal investigator it was considered unlikely related or unrelated on the CRF.

Missing relatedness will be summarized as treatment or procedure related events. Missing CTCAE grade will be summarized using a grade 3.

At each level of summarization, the summaries will include the number of subjects by system organ class and preferred term, a subject is counted once at the system organ class

and once at each preferred term within the system organ class level. For summaries by system organ class, preferred term, and maximum relationship, a subject is counted once at the strongest relationship level for which the event occurred at the system organ class level and the highest relationship level for each unique preferred term within that system organ class level. Therefore, subjects may only contribute once to each preferred term and once to each system organ class level. Summaries by CTCAE grade would be handled similar to the summaries by relationship to AE.

Summaries presenting frequency of AEs by system organ class and preferred term will be ordered by overall descending frequency of system organ class and then, within a system organ class, by overall descending frequency of preferred term.

An overall summary of all TEAEs will be presented by treatment group and will include the number and percentage of subjects experiencing the following:

- Any TEAE
- TEAE related to the expanded cell product
- TEAE by strongest relationship to transplant
- TEAE by greatest CTCAE grade
- TEAE leading to study withdrawal
- Treatment emergent SAE (TESAE)
- TEAE resulting in death

In addition to the overall AE table above, the following AE tables will also be provided.

- TEAEs overall and by system organ class and preferred term
- Expanded Cell Product-related TEAEs overall and by system organ class and preferred term
- TEAEs by maximum CTCAE grade, overall and by system organ class and preferred term
- TEAEs by maximum relationship to expanded cell product , overall and by system organ class and preferred term
- TEAEs by maximum relationship to transplant, overall and by system organ class and preferred term
- Serious TEAEs, overall and by system organ class and preferred term
- TEAEs resulting in death, overall and by system organ class and preferred term

- TEAES involving bleeding as determined by the study team prior to database lock and unblinding of the sponsor for the primary analysis

Listings will be provided for the following:

- All AEs, with the verbatim term of non-TEAEs flagged
- CTCAE Grade 3 or higher AEs (this is a subset of AEs where severity is marked as CTCAE grade 3, 4, or 5)
- Expanded Cell Product-Related AEs
- SAEs (this is a subset of the AEs where serious is marked as "Yes")
- AEs leading to death (This is a subset of the AEs where outcome is indicated as "Death" or the CTCAE grade is 5)

10.4. LABORATORY EVALUATIONS

Blood samples will be collected for laboratory safety tests as specified in the schedule of assessments (Section 3.7). Due to the design of the study, local laboratories are used by the sites involved in the study; therefore different units of measure may be used to report results for the same laboratory parameter. Standard laboratory values required for the statistical analyses of laboratory measures will therefore be obtained from originally entered data values using proper conversion factors. Laboratory testing is described below.

Complete Blood Count (CBC)

White blood cell count, hemoglobin, hematocrit, platelet count, neutrophils (absolute), lymphocytes (absolute), monocytes (absolute), eosinophils (absolute), basophils (absolute), neutrophils (%), lymphocytes (%), monocytes (%), eosinophils (%), and basophils (%).

Blood Chemistry

Sodium, potassium, chloride, carbon dioxide, glucose, blood urea nitrogen (BUN), creatinine, phosphate, albumin, magnesium, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, total bilirubin, direct bilirubin, total protein, and lactic dehydrogenase (LDH).

Immunoglobulin

Immunoglobulin A (IgA), Immunoglobulin G (IgG), and Immunoglobulin M (IgM).

Lymphocyte Panel

CD3, %CD3, CD4, % CD4, CD8, %CD8, CD4/CD8, CD19, %CD19, CD56, and %CD56.

Panel Reactive Antibody (PRA)

PRA will be collected prior to HCT, at engraftment, and at unscheduled visits.

Descriptive statistics by treatment group and time of assessment will be presented for each laboratory parameter. Changes from baseline will be presented. All laboratory values will be classified as low, normal, or high based on normal ranges supplied by each local laboratory. For purposes of analyses, laboratory results based upon Standard International system (SI) of units will be used.

For each summary, the number of non-missing observations, mean, median, standard deviation, inter-quartile range, minimum, and maximum values will be presented by treatment group.

10.5. PERFORMANCE STATUS

ECOG, Lansky, and Karnofsky performance status scores will be collected and presented descriptively by type of performance status.

11. CHANGES FROM PROTOCOL PLANNED ANALYSES

This SAP is based on the most recent version of the protocol. The only substantive changes to planned protocol analyses include the addition of several exploratory endpoints to further evaluate the clinical and pharmacoeconomic impact of NLA101 treatment, consistent with the protocol-defined secondary and exploratory objectives of the study. These additional exploratory endpoints include comparisons of disease-free survival, GVHD-free, relapse-free survival, ICU-free days within the first 100 days post HCT, and related sensitivity analyses.

The protocol specifies a secondary endpoint of severe bacterial and fungal infections. However, as the study halted prior to completion of blinded review and cleaning of this endpoint, any meaningful analysis is not possible as consistency standards and queries have not been addressed. Therefore, this SAP has removed this endpoint from the analysis plan as the data would not be interpretable.

In addition, subgroup analyses, sensitivity analyses, and study populations have been added in order to evaluate the robustness and generalizability of the results from the primary and key secondary endpoints.

12. REFERENCES

1. Guidance for Industry Statistical Principles for Clinical Trials (ICH E9), September 1998.

2. Guidance for Industry Structure and Content of Clinical Study Reports (ICH E3), July 1996.

3. Guidance for Industry Good Clinical Practice (ICH E6), April 1996.

4. Fine JP, Gray RJ. (1999). A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc; 94:496 –509.

5. Gray, R. (1988). A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. The Annals of Statistics, 16, 1141–1154

6. SAS macro %CIF. http://support.sas.com/kb/45/997.html

13. PROGRAMMING CONSIDERATIONS

The following are Syneos Health's standard programming considerations with additional Nohla Therapeutics specifications.

All TLFs, and statistical analyses will be generated using SAS® for Windows, Release 9.4 (SAS® Institute Inc., Cary, NC, USA), unless specified otherwise. Computer-generated TLF output will adhere to the following specifications.

13.1. GENERAL CONSIDERATIONS

- One SAS program can create several outputs.
- Each output will be stored in a separate file.
- Output files will be delivered in Word format.
- Numbering of TFLs will follow ICH E3 guidance

13.2. TABLE, LISTING, AND FIGURE FORMAT

13.2.1. General

- All TLFs will be produced in landscape format, unless otherwise specified.
- All TLFs will be produced using the Courier New font, size 9
- The data displays for all TLFs will have a minimum 3/4-inch margin at the top and 3/8 of an inch on all other sides.
- Headers and footers for figures will be in Courier New font, size 9.
- Legends will be used for all figures with more than 1 variable, group, or item displayed.
- TLFs will be in black and white (no color), unless otherwise specified
- Specialized text styles, such as bolding, italics, borders, shading, and superscripted and subscripted text, will not be used in the TLFs, unless otherwise specified. On some occasions, superscripts 1, 2, or 3 may be used (see below).
- Only standard keyboard characters will be used in the TLFs. Special characters, such as non-printable control characters, printer-specific, or font-specific characters, will not be used. Hexadecimal-derived characters will be used, where possible, if they are appropriate to help display math symbols (e.g., μ). Certain subscripts and superscripts (e.g., cm^2, C_{max}) will be employed on a case-by-case basis.
- Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats, as appropriate.

13.2.2. Headers

- All output should have the following header at the top left of each page:

  <Sponsor Name> Protocol XXX (Syneos Health study number xxx)
  Draft/Final Run <date>

- All output should have Page n of N at the top or bottom right corner of each page. TLFs should be internally paginated in relation to the total length (i.e., the page number should appear sequentially as page n of N, where N is the total number of pages in the table).
- The date output was generated should appear along with the program name as a footer on each page.

13.2.3. Display Titles

- Each TLF should be identified by the designation and a numeral. (i.e., Table 14.1.1). ICH E3 numbering will be used. A decimal system (x.y and x.y.z) should be used to identify TLFs with related contents. The title is centered. The analysis set should be identified on the line immediately following the title. The title and table designation are single spaced. A solid line spanning the margins will separate the display titles from the Column headers. There will be 1 blank line between the last title and the solid line.

Table x.y.z
First Line of Title
Second Line of Title if Needed
ITT Analysis Set

13.2.4. Column Headers

- Column headings should be displayed immediately below the solid line described above in initial upper-case characters.
- In the case of efficacy tables, the variable (or characteristic) column will be on the far left followed by the treatment group columns and total column (if applicable). P-values may be presented under the total column or in separate p-value column (if applicable). Within-treatment comparisons may have p-values presented in a row beneath the summary statistics for that treatment.
- For numeric variables, include “unit” in column or row heading when appropriate.
- Analysis set sizes will be presented for each treatment group in the column heading as (N=xx) (or in the row headings if applicable). This is distinct from the ‘n’ used for the descriptive statistics representing the number of subjects in the analysis set.
- The order of treatments in the tables and listings will be by randomized treatment group with EXP first, followed by SOC, and then screen failures where applicable., followed by a total column (if applicable).

13.2.5. Body of the Data Display

13.2.5.1. General Conventions

Data in columns of a table or listing should be formatted as follows:

- alphanumeric values are left-justified;
- whole numbers (e.g., counts) are right-justified; and

- numbers containing fractional portions are decimal aligned.

13.2.5.2. Table Conventions

- Units will be included where available.
- If the categories of a parameter are ordered, then all categories between the maximum and minimum category should be presented in the table, even if n=0 for all treatment groups in a given category that is between the minimum and maximum level for that parameter. For example, the frequency distribution for symptom severity would appear as:

Severity Rating	N
severe	0
moderate	8
mild	3

Where percentages are presented in these tables, zero percentages will not be presented and so any counts of 0 will be presented as 0 and not as 0 (0%).

- If the categories are not ordered (e.g., Medical History, Reasons for Discontinuation from the Study, etc.), then only those categories for which there is at least 1 subject represented in 1 or more groups should be included.
- An Unknown or Missing category should be added to any parameter for which information is not available for 1 or more subjects.
- Unless otherwise specified, the estimated mean and median for a set of values should be printed out to 1 more significant digit than the original values, and standard deviations should be printed out to 2 more significant digits than the original values. The minimum and maximum should report the same significant digits as the original values. For example, for systolic blood pressure:

N	XX
Mean	XXX.X
Std Dev	X.XX
Median	XXX.X
Minimum	XXX
Maximum	XXX

- P-values should be output in the format: “0.xxx”, where xxx is the value rounded to 3 decimal places. Any p-value less than 0.001 will be presented as <0.001. If the p-value should be less than 0.0001 then present as <0.0001. If the p-value is returned as >0.999 then present as >0.999
- Percentage values should be printed to one decimal place, in parentheses with no spaces, one space after the count (e.g., 7 (12.8%), 13 (5.4%)). Values that round down to 0.0 would be displayed as '<0.1', or as appropriate with additional decimal places. Unless otherwise noted, for all percentages, the number of subjects in the analysis set for the treatment group who have an observation will be the denominator. Percentages after zero counts should not be displayed and percentages equating to 100% should be presented as 100%, without any decimal places.
- Tabular display of data for medical history, prior / concomitant medications, and all tabular displays of adverse event data should be presented by the body system, treatment class, or system organ class with the highest occurrence in the active treatment group in decreasing

order, assuming all terms are coded. Within the body system, drug class and system organ class, medical history (by preferred term), drugs (by ATC1 code), and adverse events (by preferred term) should be displayed in decreasing order. If incidence for more than 1 term is identical, they should then be sorted alphabetically.

- Missing descriptive statistics or p-values which cannot be estimated should be reported as "-".
- The percentage of subjects is normally calculated as a proportion of the number of subjects assessed in the relevant treatment group (or overall) for the analysis set presented. However, careful consideration is required in many instances due to the complicated nature of selecting the denominator, usually the appropriate number of subjects exposed. Describe details of this in footnotes or programming notes.

- For categorical summaries (number and percentage of subjects) where a subject can be included in more than one category, describe in a footnote or programming note if the subject should be included in the summary statistics for all relevant categories or just 1 category and the criteria for selecting the criteria.

- Where a category with a subheading (such as system organ class) has to be split over more than one page, output the subheading followed by "(cont)" at the top of each subsequent page. The overall summary statistics for the subheading should only be output on the first relevant page.

13.2.5.3. Listing Conventions

- Listings will be sorted for presentation in order of treatment groups as above, treatment group, subject number, visit/collection day, and visit/collection time.
- Missing data should be represented on subject listings as either a hyphen ("-") with a corresponding footnote ("- = unknown or not evaluated"), or as "N/A", with the footnote "N/A = not applicable", whichever is appropriate.
- Dates should be printed in SAS® DATE9.format ("ddMMMyyyy": 01JUL2000). Missing portions of dates should be represented on subject listings as dashes (--JUL2000). Dates that are missing because they are not applicable for the subject are output as "N/A", unless otherwise specified.
- Units will be included where available

13.2.5.4. Figure Conventions

- Unless otherwise specified, for all figures, study visits will be displayed on the X-axis and endpoint (e.g., treatment mean change from Baseline) values will be displayed on the Y-axis.

13.2.6. Footnotes

- A solid line spanning the margins will separate the body of the data display from the footnotes.
- All footnotes will be left justified with single-line spacing immediately below the solid line underneath the data display.
- Footnotes should always begin with "Note:" if an informational footnote, or 1, 2, 3, etc. if a reference footnote. Each new footnote should start on a new line, where possible.

- Subject specific footnotes are to be avoided, where possible.
- Footnotes will be used sparingly and must add value to the table, figure, or data listing. If more than six lines of footnotes are planned, then a cover page may be used to display footnotes, and only those essential to comprehension of the data will be repeated on each page.
- The last line of the footnote section will be a standard source line that indicates the name of the program used to produce the data display, date the program was run, and the listing source (i.e., 'Program : myprogram.sas Listing source: 16.x.y.z').

14. QUALITY CONTROL

SAS programs are developed to produce output such as analysis data sets, summary tables, data listings, figures or statistical analyses. Syneos Health SOP 03.010 and 03.013 provide an overview of the development of such SAS programs.

Syneos Health SOP 03.009 describes the quality control procedures performed for all SAS programs and output. Quality control is defined here as the operational techniques and activities undertaken to verify the SAS programs produce the output by checking for their logic, efficiency and commenting and by review of the produced output.