STATISTICAL ANALYSIS PLAN

INS-3: A Phase 3, Randomized, Double-Masked, Placebo-Controlled Study of the Efficacy and Safety of myo-Inositol 5% Injection to Increase Survival without Severe Retinopathy of Prematurity (Reduce-ROP) in Extremely Premature Infants

IND:	70510
SAP VERSION:	0.48
SAP DATE:	July 1, 2013
SPONSOR:	NICHD
PREPARED BY:	RTI International 3040 Cornwallis Rd. Research Triangle Park, NC 27709-2104
AUTHOR (S):	Abhik Das, PhD Tracy Nolen, DrPh

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	Adverse event
AN	Abbott Nutrition and/or their designated successor company
BPD	Bronchopulmonary dysplasia
CA	Chronological age
CPAP	Continuous positive airway pressure
DCC	Data Coordinating Center
DSMC	Data and safety monitoring committee
GA	Gestational age
GDB	Generic database
GI	Gastrointestinal
ITT	Intent–to-treat
IV	Intravenous
IVH	Intraventricular hemorrhage
LPM	Liters per minute
MRI	Magnetic resonance imaging
MI	Multiple imputation
NDA	New drug application
NDI	Neurodevelopmental impairment
NEC	Necrotizing enterocolitis
NICHD	*Eunice Kennedy-Shriver* National Institute of Child Health and Human Development
NRN	Neonatal Research Network
NRN Status	Soonest of discharge/transfer, death or chronologic age of 120 days
PDA	Patent ductus arteriosus
PI	Principal investigator
PO	Per mouth, orally (in the neonatal population "oral" drugs are often delivered by nasogastric or gastric tubes, as are feedings because the infants are too neurologically immature to feed from a nipple)
PMA	Postmenstrual age
PVL	Periventricular leukomalacia
RDS	Respiratory distress syndrome
ROP	Retinopathy of prematurity
ROP-EAC	Retinopathy of Prematurity Endpoint Adjudication Committee
sROP	Severe ROP
RTI	Research Triangle Institute, Research Triangle Park, NC
SAE	Serious adverse event
SAP	Statistical analysis plan

1 BACKGROUND AND PROTOCOL HISTORY

Previous trials have shown that supplementation of inositol soon after birth results in a reduction of death, chronic lung disease, and retinopathy of prematurity (ROP) among premature infants with respiratory distress syndrome (RDS). These are common and serious morbidities among this population and warrant further study of inositol. However, previously published data are in infants who, while preterm, were on average heavier and more mature than the population at risk today.

Additionally, there is no FDA-approved intravenous (IV) or oral (PO) inositol formulation to prevent or reduce the severity of ROP in premature infants. Given the favorable results observed in published trials of this compound and its strong safety profile, the *Eunice Kennedy-Shriver* National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN) initiated a research program to determine if inositol supplementation administered soon after birth in extremely premature infants is safe and reduces the incidence of severe ROP, as well as other common morbidities seen in this population. Two Phase 2 pharmacokinetic studies (INS-1 and INS-2) evaluating the single- and multiple-dose pharmacokinetics, safety, and dose-ranging of myo-Inositol 5% Injection administration in premature infants have been completed by the NICHD NRN. Results of these studies indicate that the treatment is safe and support the inositol dose of 80 mg/kg/day selected for evaluation in the Phase 3 program. As such, this current trial is intended to formally assess the efficacy and safety of intravenous and enteral use of myo-Inositol 5% Injection at the dose of 80 mg/kg/day.

2 PURPOSE OF THE ANALYSES

This statistical analysis plan (SAP) contains detailed information about statistical analyses to be performed to evaluate the efficacy and safety of multidose myo-Inositol 5% Injection using data from a phase III trial. The results of these analyses will be included in the clinical study report and primary manuscript.

Different statistical analytic approaches for assessing efficacy are planned for the scientific publication of this study conducted by the NICHD NRN and the subsequent regulatory submission undertaken by Abbott Nutrition (AN). In this SAP, when the regulatory approach is different from the publication approach, the regulatory approach will be specified as well. Otherwise, the content applies to both purposes.

3 STUDY OBJECTIVES AND OUTCOMES

3.1 Study Objectives

3.1.1 Primary Study Objective

The primary objectives are:

- To determine the efficacy of myo-Inositol 5% Injection compared to placebo for increasing the probability of survival without severe ROP in premature infants <28 weeks' gestation.

- To determine the safety of myo-Inositol 5% Injection compared to placebo in premature infants <28 weeks' gestation.

3.1.2 Secondary Study Objectives

The secondary objective is to determine if myo-Inositol 5% Injection compared to placebo has an effect on other common morbidities of extreme premature birth.

3.2 Outcomes

3.2.1 Primary Outcome

The primary efficacy outcome for this study is survival without severe ROP (sROP) with eye examinations up to 55 weeks post menstrual age (PMA), if needed for final/acute ROP determination.

- The primary sROP outcome is defined as at or prior to 55 weeks PMA reaching either
    - favorable ROP status defined as in both eyes meeting one or more of the following criteria
        - fully vascularized to the ora serrata on 1 examination, and PMA is greater than 35 weeks
        - or the retinal vessels have grown into Zone III, observed on two sequential examinations showing only immature vessels, regressing ROP or Stage 1 or Stage 2 ROP
        - or unequivocally Zone III vessels on one examination in an infant of 35 weeks PMA or older, if there has been no previous ROP in Zone I or Zone II, and this examination is not the only examination the infant has had
        - or PMA has reached 50 weeks and there is no prethreshold or worse ROP present where prethreshold ROP is defined as (ETROP 2003, CRYO-ROP 1988):

- Zone II: Stage 3 ROP or
- Zone I: any stage ROP

- unfavorable ROP status defined as one or both eyes showing any of the following findings/circumstances, confirmed by a second examiner (or documented by fundus photograph)
  - meeting criteria "Type 1 ROP" established by the ETROP study, 2003
  - or ROP that has progressed to stages worse than ETROP Type 1 ROP, including classic CRYO-ROP threshold, or any degree of retinal detachment (CRYO-ROP 1988)
  - or death before ROP favorable or unfavorable status is reached.

- Death occurring after favorable ROP status has been met will not count towards an unfavorable outcome.

Additional details of the primary outcome definition including criteria used to classify "Type 1 ROP" or worse as well as how statistical endpoints will be developed based on this outcome and how missing outcomes will be handled for the publication and regulatory analyses are provided in Sections 9.2 and 12 respectively.

3.2.2 Secondary Outcomes

- Efficacy
  - Bronchopulmonary dysplasia (BPD): NICHD Physiologic Definition: Requiring oxygen to maintain an oxygen saturation of ≥90% while breathing room air at 36 weeks PMA (Walsh 2004).
  - BPD (Physiologic Definition) or Death from BPD prior to 37 weeks PMA, with cause of death certified by the Center PI as BPD for the primary cause, or a significant co-contributing cause of death.
  - All cause death (through NRN status [earliest of discharge/transfer, death or chronologic age of 120 days], prior to final ROP status, as well as through 55 weeks PMA): defined as death from any cause following randomization. The primary, underlying cause of death will be certified by the Principal Investigator (PI) at each center who may also list co-contributing causes.
  - Any ROP: defined as ROP of any severity that is observed on at least 2 examinations in either eye through the time that Acute/Final ROP status is reached. ROP that is observed on only 1 of multiple eye examinations will not be considered "any ROP". An infant who has only one eye examination prior to death, and that examination shows any stage ROP, will be also be considered as having "any ROP".

- Type 2 ROP through the time that Acute/Final ROP status is reached: defined as 1 or both eyes reaching Type 2 ROP (ETROP 2003), but not Type 1. Type 2 ROP is defined as: (ETROP 2003).
  - In Zone II: Stage 3 ROP without Plus Disease, or
  - In Zone I: Stage 1 or 2 ROP without Plus Disease
- Severe intraventricular hemorrhage (IVH): IVH Grades 3 or 4 on either side of the brain, or extensive periventricular leukomalacia (PVL). IVH will be classified as described by Papile (Papile 1978). PVL will be graded by the characteristics of the periventricular white matter on ultrasound or magnetic resonance imaging (MRI) done between 4 to approximately 6 weeks after birth according to usual practice (Barkivich 2000) and/or porencephalic cyst.

- Safety
  - Occurrence of adverse events (AEs) and serious adverse events (SAEs), in addition to the clinical diagnoses through 7 days post last dose of study therapy.
  - Common complications of prematurity through the time of NRN status. These include:
    - Occurrence of necrotizing enterocolitis (NEC): Stage II or worse, whether treated (medically or surgically) and if the infant survived (modified Bell's classification [Walsh 1986]).
    - Isolated gastrointestinal perforation judged not to be due to NEC.
    - Late onset sepsis: culture positive septicemia/bacteremia (≥72 hours of age) treated with antibiotics for ≥ 5 days or died before treatment was completed.
    - Occurrence of clinically significant patent ductus arteriosus (PDA), as well as PDA receiving intervention with prostaglandin inhibitors, and/or surgery. This does not include prophylactic Indomethacin or Ibuprofen administered without the diagnosis of clinically significant PDA.
    - Seizures treated with an anticonvulsant for >72 hours.
    - Total days on parenteral nutrition (including amino acids and/or lipids).
    - Days on oxygen, days on ventilator.
    - If hearing screening was performed, did the infant fail in one or both ears (defined as never passing a hearing screening in one or both ears).

- Long term Outcomes Collected Using the NICHD Follow-up Protocol

- Data collection for other common complications of prematurity performed during long-term follow-up through 22-26 months will be conducted using the established NICHD-NRN Protocol. These data include:

  - Neurodevelopment at 22-26 months corrected age (i.e., 22-26 months past due date) using the Bayley Scales of Infant Development III.

  - Vision loss as diagnosed by an ophthalmologist as legally blind, and subdivided into "ophthalmic origin", or "not ophthalmic origin" (i.e., cortical blindness is non-ophthalmic in origin and indicates that there is no retinal detachment or other abnormal fundus or ocular finding, except optic atrophy. Such cases will be considered central (neurologic) in origin.)

  - Subsequent ophthalmic intervention/treatment related to ROP, or other ophthalmic treatments, including glasses, contacts, scleral buckle, vitrectomy, cataract surgery, strabismus surgery, glaucoma treatment, etc. This will form a final confirmation of the clinical findings and sequelae of ROP in each child following discharge. To facilitate collecting this information from outpatient ophthalmology visits (anticipated in all infants, and to involve serious medical issues in about 20%), explicit consent to obtain office records will be included in the consent form.

  - Cerebral palsy by severity category (absent/mild/moderate/severe).

  - Hearing loss requiring that hearing aids be prescribed.

  - Neurodevelopmental Impairment (NDI).

  - Overall health status per recall from the parent/guardian (including survival, re-hospitalizations, surgeries, ongoing medications, and chronic illnesses).

Additional details of the secondary outcomes definitions including how statistical endpoints will be developed based on these outcomes and how missing outcomes will be handled are provided in Sections 9.2, 10, and 11.2 respectively.

4 STUDY METHODS

4.1 Overall Study Design and Plan

This will be a randomized, double-masked, placebo-controlled study. A total of 1760 infants will be randomized 1:1 to a dose of 80 mg/kg/day of myo-Inositol 5% Injection or placebo which will be given in two daily doses, 40 mg/kg/dose, starting within 72 hours of birth, and continuing until 34 weeks PMA (gestational age [GA] at birth plus chronological age [CA] in weeks), 10 weeks CA, or the time of discharge, whichever occurs first. Randomization will be stratified by center and GA. Myo-Inositol 5% Injection or placebo will be administered intravenously until enteral feedings are established, at which time the same dose and formulation will be administered enterally.

For publication purposes, the analysis of the primary efficacy outcome will consider the entire study population. In support of a new drug application (NDA) for use of myo-Inositol 5% Injection to improve the rate of survival without sROP in premature infants <28 weeks' gestation, the analysis of the primary efficacy outcome will be conducted within pre-specified regulatory sub-studies created by administratively splitting infants enrolled at each study center into two groups referred to as regulatory Study A and Study B as discussed in Section 12.1.

4.2 Study Population

4.2.1 Inclusion Criteria

An infant will be eligible for study participation if he/she meets the following criteria:

1. Inborn or out born infants of either gender or any race with best obstetrical estimate of gestation <28 weeks ($27^{6/7}$ weeks and younger). Gestational age will be determined by best obstetrical estimate per NRN Generic Database (GDB) protocol (using the hierarchy of best obstetrical estimate using early ultrasound dating, maternal menstrual dating confirmed by examination, or Neonatal gestational age assessment by physical examination). The GA assigned will be carried forward for all study purposes.

2. Alive at 12 hours.

3. Age in hours up to 72 hours, although we will seek enrollment as early as feasible after consent and 12 hours.

4. Informed consent signed and dated by parent and/or guardian, which includes likelihood of completing follow-up ophthalmic examinations as an outpatient, and long-term follow-up.

4.2.2 Exclusion Criteria

An infant will be excluded from the study if he/she meets any of the following criteria:

1. Major congenital malformations

2. Congenital malformations of the eye identified prior to randomization.

3. Overt evidence of intrauterine congenital infections ("TORCH") or life-threatening impairment of renal, hepatic, or cardiac function (considered moribund).

4.3 Study Arm Assignment and Randomization

Randomization will be performed at the overall study level and will be stratified by center and GA groups (<26 weeks vs. ≥26 weeks GA). Stratified randomization (1:1 to myo-Inositol 5% Injection or placebo) will be performed using randomly permuted blocks, with block sizes known only to the data coordinating center.

For regulatory analyses, the randomized study population will be administratively split within study center in a pre-determined manner as further detailed in Section 12.1 to define the regulatory sub-studies A and B.

4.4 Masking and Data Lock

4.4.1 General Masking Procedures

This study will be double-masked. The only persons unmasked at enrolling sites will be the pharmacists who are not otherwise involved in any other aspects of this study. Furthermore, non-site study team members including the study PI, the INS-3 subcommittee for the NICHD NRN, sponsor (NICHD NRN), and AN will also be masked while the study is ongoing. The independent NRN Data Safety and Monitoring Committee (DSMC) will review the data in a partially masked fashion (e.g. group IDs of X and Y rather than by treatment assignment) as specified in the protocol. The DSMC may request to be unmasked if the Committee felt that such information was necessary for the discharge of its duties. The Data Coordinating Center (DCC) study statistician will be responsible for reporting to the DSMC and as such will be unmasked. However, the DCC senior investigators will remain masked. In the case of partially missing data, the ROP endpoint adjudication committee (ROP-EAC) will provide a masked assessment and adjudication of the final ROP status for inclusion in the efficacy analyses.

4.4.2 Database Lock

At the end of the study, data lock and unmasking will occur in 2 stages to ensure complete and accurate data in a final locked data set, yet facilitate timely analysis of the primary outcome data.

1. Primary data

    i. Purpose: Primary efficacy outcomes and early safety during drug administration used for primary scientific publication as well as for regulatory submission

    ii. Data Included:

        a. All baseline data

        b. Drug administration from study start through the soonest of discharge, transfer, death, or maximum 70 days

        c. AEs and concurrent drug exposure through 7 days after last dose of study drug

        d. Hospital course data through NRN status (earliest of discharge/transfer, death, hospitalization up to 120 days chronologic age) including clinical outcomes

        e. ROP final/acute outcome: May occur up to 55 weeks PMA (up to 3 months after full term due date).

f. Death form

2. Long Term Follow up data

i. Purpose: Long term outcome evaluation – safety

ii. Data Included:

a. All data from I. above without change

b. Neuro-developmental assessment and Health survey at 22-26 months corrected age (after the date when the infant would have been born at full term).

This sequential data lock does not result in an interim analysis, but rather an analysis of each of the two target result sections [1) primary clinical outcomes, and 2) long term safety] being started when the data are complete for that section. In general, no summaries or analyses by treatment group will be provided to any study team member for any data prior to the data being locked. Furthermore, no individuals other than the statisticians at Research Triangle Institute (RTI), the DCC for the NRN, will have access to individual treatment assignment until the end of the first stage of data lock.

Specifically, at the time of the first stage of the data lock, the planned analyses for the associated data will be conducted by DCC statisticians and distributed to the study PI, study subcommittee, NICHD, and AN. At this point, these study team members may become unmasked with treatment assignment for individual study subjects. However, the Follow up investigators who conduct neuro-developmental assessments at 22-26 months and the clinical staff and families will remain masked as to study drug assignment of individual subjects throughout this period. Accordingly, while summary results may be distributed to larger audiences at this time (e.g. as part of new drug application [NDA] submission, conference presentations), any summary results that would be accessible to Follow up investigators, clinical staff or families would not include any information that could unmask individual treatment assignment.

After the completion of the final stage of database lock, follow-up investigators and clinical staff may be unmasked to individual treatment assignment, if requested. Additionally, parents may request and receive information about which treatment their infant received at this time.

Any deviations from this plan will be discussed in the clinical study report. For example, the clinical study report will include details of any emergency unmasking of individual study subjects due to safety concerns (e.g. a suspected adverse drug reaction). Likewise, if the study is halted early for safety, futility or efficacy, some aspects of treatment assignment unmasking may also occur in an expedited fashion.

4.5 Study Flow Chart of Assessments and Evaluations

Procedures	12-72[a] hours after birth	Study Day / Study Week	1	2-35 / 1-5	36-70 / 6-10	71-105 / 11-15	106-140 / 16-20	141-175 / 21-25	176-210 / 26-30	22-26 months corrected
Informed Consent	X									
Randomization[a]	X									
Cranial Ultrasound[b]				X	X					
Study Drug Administration[c]			X	X	X					
Ophthalmologic Examination[d]				(X)	X	X	X	(X)	(X)	
Infant Growth[e]				X	X	X	X	(X)	(X)	X
Concurrent Medications[f]	X		X	X	X	X				
Adverse Event Assessments[g]	X		X	X	X	X				
Clinical Diagnoses[h]	X		X	X	X	X	X	X	X	X
Follow-up Evaluation[i]										X

[a] With allowance for up to 72 hours; however, the goal is to enroll as soon as practicable after 12 hours
[b] Cranial ultrasounds will be performed at times consistent with clinical practice to evaluate for IVH and PVL. Otherwise they will not be mandated by the study
[c] Daily study drug administration will begin following consent and end at 34 completed weeks PMA, 10 weeks chronologic age, or at the time of discharge, whichever occurs first.
[d] Ophthalmologic examinations will begin at Week 31 PMA and continue every 1 to 3 weeks based on findings until Acute/Final ROP status is achieved.
[e] Weight is recorded weekly, head circumference is recorded every 2 weeks, and length is recorded every 4 weeks.
[f] Recorded from 24 hours prior to randomization until discharge, transfer, or 7 days after the last dose of study drug administration, whichever occurs first.
[g] Recorded from treatment initiation until 7 days after last dose of study drug
[h] Clinical diagnoses will be collected from the medical record throughout hospitalization and again at the follow-up visit.
[i] Follow-up evaluations and interval medical data will be collected at 22-26 months corrected age per the NRN Follow-up Protocol (described separately)

5 ANALYSIS POPULATIONS

Depending on type, each analysis will be conducted within one or more of the following analysis populations. For the overall study analyses for publication, each population will include all infants that meet the population definition. For the regulatory submission analyses conducted within regulatory sub-studies A and B, each sub-study will have a set of analysis populations per the definitions below that only comprise the infants assigned to that regulatory sub-study.

Safety Population

The safety population will include all infants who were randomized and received at least 1 dose of study drug. The safety population will be used for all safety analyses and infants will be grouped according to actual treatment received.

Intent-to-Treat Population

The ITT population is the primary population for formal efficacy analyses. This population includes all subjects randomized. For these analyses, subjects will be analyzed as part of the study arm to which they are assigned by randomization, regardless of actual therapy they received.

Per-Protocol Population

The per-protocol population will be used for secondary sensitivity analyses of efficacy. This population includes all subjects who received treatment according to randomized assignment and per-protocol (through first of 34 completed weeks PMA (GA at birth plus CA in weeks), 10 weeks CA, or the time of discharge with study drug discontinuation or hold occurring only as specified in the protocol.

6 SAMPLE SIZE DETERMINATION

Sample size calculations are based on the primary analysis of the single trial intended for scientific publication; however considerations for available statistical power are also presented for the submission analyses with the regulatory sub-studies A and B.

The underlying prevalence of unfavorable sROP status (meeting criteria for Type I ROP or worse or expiring before ROP status is reached) is assumed to be 0.30 (SUPPORT 2010). Based on a 2-sided Mantel-Haenszel test of the primary hypothesis with α=0.05, a total sample size of 1672 infants (836 per group) provides a 90% power for detecting a treatment effect of $\geq$0.07 absolute reduction in unfavorable ROP status prevalence in the inositol arm compared to placebo.

It is also estimated that approximately 3-5% of the enrolled population may be lost to follow-up without sufficient information to obtain an adjudicated ROP endpoint. These subjects will be non-evaluable.

Accordingly, the final minimum required sample size to retain 90% power at an overall α level of 0.05 with one formal interim analysis for efficacy for detecting a $\geq$0.07 reduction in unfavorable ROP status prevalence is 1760 infants.

As detailed further in Section 12, if the findings of the primary publication analysis demonstrate a significant benefit, additional analyses for regulatory purposes (e.g. NDA submission) will be conducted. This prioritization of analyses precludes the necessity of adjusting alpha for the two study purposes. For the regulatory analyses, the study population will be administratively split in a pre-determined manner (Section 12.1). Within each

'regulatory sub-study', the primary endpoint will be tested using a Mantel-Haenszel chi-square test for ordinal data with outcomes of favorable, indeterminate (i.e., partial ROP information available in that ROP assessments conducted but final status not met) and unfavorable. Missing outcomes (i.e., no available ROP data) will be imputed via multiple imputation (MI). To assess the power for this analysis, simulations were conducted assuming that there will be 880 subjects within each 'regulatory sub-study', an underlying prevalence for sROP of 0.30, 5% subjects with completely missing outcome information and varying prevalence of indeterminate outcome and treatment effect (while the overall study is powered to test for minimally clinically significant treatment reduction, prior studies suggest a larger treatment effect may exist). The resulting power estimates are provided in Table 1. In general, the power estimates are fairly robust to the level of indeterminate outcomes assuming that level is not markedly more than currently anticipated. If the true treatment effect is a 7% reduction in the myo-Inositol 5% Injection group, each sub-study analysis only has between a 56 to 62 percent chance of demonstrating a significant benefit for that sub-study. However, if the true treatment effect is an 8-9% reduction, the power for each sub-study analysis increases to between 69-83%.

Table 1: Probabilities of Detecting Potential Treatment Effect for Regulatory Analyses within Each 'Regulatory Sub-Study'

True Underlying Treatment Effect	Indeterminate Outcome Prevalence		Probability of Detecting Effect
	Would have been favorable if observed	Would have been unfavorable if observed	
0.07	0.03	0.03	0.62
	0.03	0.09	0.61
	0.09	0.03	0.56
	0.09	0.09	0.56
0.08	0.03	0.03	0.73
	0.03	0.09	0.73
	0.09	0.03	0.70
	0.09	0.09	0.69
0.09	0.03	0.03	0.82
	0.03	0.09	0.83
	0.09	0.03	0.80
	0.09	0.09	0.79

NOTE: For simplicity, missing outcomes were excluded from the simulations,

Although the study is only powered to formally test 1 primary hypothesis of efficacy (treatment effect on survival without sROP), a variety of secondary outcomes will be assessed in an exploratory manner. Due to the exploratory nature of these secondary analyses, no additional adjustments for multiplicity outside of accounting for the interim analysis are planned. Given a sample size of 1760 (effective sample size of 1672 infants), the probabilities of detecting a potential treatment effect (i.e., 30% reduction or increase) for the overall study population in common clinical outcomes of secondary interest are summarized in Table 2.

Table 2: Probabilities of Detecting Potential Treatment Effect on Secondary Outcomes

	Underlying Prevalence	Absolute Effect For 30% Change	Probability of Detecting	
			Reduction	Increase
Death all cause	0.18	0.054	0.87	0.78
BPD physiologic definition	0.38	0.114	0.99	0.99
BPD or death from BPD before 37 weeks PMA	0.49	0.147	1.00	1.00
Any ROP	0.65	0.195	1.00	1.00
Type 2 ROP	0.25	0.075	0.96	0.92
Severe IVH (Grade III/IV or shunt or PVL)	0.20	0.060	0.90	0.83

7 STATISTICAL / ANALYTICAL ISSUES

7.1 General Rules

Data will be summarized by treatment group. For summaries of study data, categorical measures will be summarized in tables listing the frequency and the percentage of subjects in each study arm; continuous data will be summarized by presenting mean, standard deviation, median and range; and ordinal data will be summarized by only presenting median and range.

Most statistical computations will be performed and data summaries will be created using SAS 9.3 or higher. If additional statistical software is required, this will be discussed in the study report.

7.2 Adjustments for Covariates

In general, summaries and analyses will be stratified by or adjusted for gestational age strata. Specifically, table summaries will be presented for all subjects and for each gestational age strata used for randomization. All model-based analyses and test-statistics examining the treatment effect will be adjusted for study center and gestational age strata where possible. For example, the primary outcome will be tested using a robust Poisson regression model controlling for strata defined by study center and gestational age. Additionally the primary outcome as well as other binary outcomes will be analyzed using robust Poisson regression to obtain adjusted relative risk estimates for the treatment effect. All other demographic and baseline characteristics for subjects will be compared between treatment groups. If analyses of these characteristics suggest that substantial differences exist for some of these characteristics between treatment arms at baseline, their use as covariates will be explored in the adjusted exploratory analyses of efficacy and safety data.

7.3 Handling of Dropouts and Missing Data

The primary analysis for publication for the overall study population as detailed in Section 9.3 will be conducted using adjudicated survival without sROP (favorable vs. unfavorable ROP status) without any statistical imputation for missing data due to loss to follow-up or indeterminate final status. Accordingly, it is assumed for the primary analyses that the remaining missing primary outcomes after adjudication are missing completely at random. Additional sensitivity analyses will be conducted to assess the impact of adjudication and of the missing/indeterminate data on the estimates of the treatment effect. Additional analyses will include at a minimum excluding adjudicated endpoints as well as treating missing values as 'treatment failures' or as 'treatment successes.'

The analyses for regulatory submission as detailed in Section 12.2 will use a three level ordinal outcome that treats indeterminate outcomes as a mid-level response in-between favorable and unfavorable and will also use multiple imputation to include subjects with completely missing outcomes.

Otherwise, analysis of secondary efficacy and safety data will generally include available data such that no data obtained within the study assessment windows will be discarded and no imputation for missing data will be done.

7.4 Interim Analyses and Data Monitoring

While the study is ongoing, an independent NRN DSMC will examine accumulating data, to ensure protection of subjects' safety; while assuring that the study's scientific goals are being met.

The NRN DSMC will monitor safety at a preplanned frequency when approximately 25%, 50% and 75% enrolled subjects have completed study therapy as detailed in the protocol. During these reviews, the DSMC will review reports including tabular summaries of study data to determine whether there are any safety concerns that may impact continuation of the trial, or evidence that study procedures should be changed or the trial should be halted, only for reasons relating to the safety of the study subjects or inadequate trial performance (e.g., poor recruitment of subjects). The DSMC can decide to recommend suspending or stopping enrollment if a safety concern is identified. No formal analyses of efficacy data will be conducted as part of these safety monitoring activities and the DSMC will not be able to recommend stopping the study for efficacy at these reviews. Therefore, no α will be spent for these safety reviews.

Additionally, one formal interim analysis of efficacy and futility will be conducted after approximately 1,000 infants have been enrolled and reached primary endpoint. Interim analyses will be conducted using the overall study population and the primary analytic approach detailed in Section 9.3. Cut-off p-values for testing for efficacy at this interim analysis will be determined based on a Bonferroni-type correction for multiple comparisons with a nominal α of 0.0001 at the interim analysis and a nominal α of 0.0499 for the final analysis (Casella and Berger 2002). Specifically, if the p-values are < 0.0001 for the

treatment versus placebo comparison for the primary outcome and <0.001 for the treatment comparison for mortality alone at the interim analysis, then the DSMC can recommend stopping enrollment. The DSMC can recommend stopping enrollment for futility if the conditional power for the primary test of treatment effect on survival without sROP is less than 0.3 at this time.

Recommendations from the DSMC are addressed to the Director of NICHD who has the ultimate responsibility to make decisions to alter or halt this NRN study.

7.5 Masked Data Review for Adjudication

Since an estimated 5-8% of primary outcomes are expected to be missing primarily due to final acute ROP status not being determined by 55 weeks (i.e. favorable or unfavorable status as defined in Section 3.2.1), where possible, for the primary analysis for publication an adjudication process will be used to determine final acute status for those with incomplete or missing final acute ROP status. This assessment will be used in defining the survival without sROP (favorable vs. unfavorable ROP status), as follows:

- Algorithm
  - Death
    - If final acute status is reached prior to death, the primary outcome is determined by final acute status (either favorable or unfavorable ROP status)
    - If death occurs prior to any ROP examinations, the primary outcome is determined by death and subject will be considered to have unfavorable ROP status
    - If death occurs after ROP examinations but before final acute status, the primary outcome is determined by death and subject will be considered to have unfavorable ROP status
  - Lost to follow up prior to final acute status, final acute status not met until after 55 weeks PMA or received treatment for ROP before meeting anatomic definition of unfavorable status
    - If no ROP examinations are available, the primary outcome is missing
    - If ROP examinations have occurred, the primary outcome is indeterminate and will go to adjudication
- Adjudication Process
  - Four ophthalmologists that are skilled and experienced in ROP examinations and familiar with the definitions specified by the Study Protocol and the Manual of Procedures will comprise the ROP-EAC.
    - Three of the four ophthalmologists will review each case requiring adjudication.
    - No ophthalmologist will adjudicate cases from their own center.
  - The ROP assessments and essential clinical information such as GA at birth and birth weight will be provided to the reviewers in a masked fashion (no site, subject or treatment identifying information will be provided). If an infant receives intervention for sROP without meeting the anatomic definition of unfavorable status then all ROP treatment information and ROP assessments

occurring after treatment will be excluded from the information provided to the reviewers,

- Each adjudicator will independently judge each subject (each eye separately) as:
  - Eye most likely did NOT reach criteria for intervention for ROP
  - Eye most likely did reach criteria for intervention for ROP
  - Insufficient data to determine whether this eye may have reached criteria for sROP.
- For discrepancies between adjudicators
  - If two of three adjudicators select 'most likely did reach criteria' and
    - The 3rd selected 'insufficient data', the eye will be assigned to 'most likely did reach criteria'.
    - The 3rd adjudicator selected 'most likely did NOT reach criteria' then the case will go to the second stage of adjudication
  - If two of three adjudicators select 'most likely did NOT reach criteria' and
    - The 3rd selected 'insufficient data', the eye will be assigned to 'most likely did NOT reach criteria'
    - The 3rd adjudicator selected 'most likely did reach criteria' then the case will go to the second stage of adjudication
  - If two of three adjudicators select 'insufficient data' then the case will go to the second stage of adjudication
- The second stage of adjudication will be a conference call among the three adjudicating ophthalmologists. If consensus is reached, that will be the adjudicated outcome. If it is not, the decision of the majority will be used. A three-way split is not expected to be an option that will be observed.
- The adjudication process will be completed in relative real time such that adjudicated endpoints can be used for the specified interim analysis.

7.6 Multicenter Studies

Based on historic NRN data, heterogeneity among centers in the incidence of clinical outcomes including the Phase III primary efficacy outcome is expected. Consistency of study results across centers will be assessed as described in section 9.3 for the primary outcome of survival without sROP and its components (i.e. survival and sROP).

7.7 Multiple Comparisons and Multiplicity

There is only one formal hypothesis test for this Phase III study, related to detecting a treatment effect in the primary outcome. The primary analysis for publication planned for the overall study population is described in Section 9.3. Inference based on the primary analyses for regulatory submission purposes described in Section 12.2 are conditional upon detecting a significant effect in the primary publication analyses. This prioritization of analyses precludes the necessity of adjusting alpha for the two study purposes. Additionally, the regulatory sub-studies are treated as two individual studies (A and B).

All other analyses of outcomes are exploratory in nature; therefore, resulting p-values and confidence intervals will generally be provided for descriptive purposes only. As such, the only adjustment for multiplicity will be for the planned interim analysis of efficacy performed after approximately 1,000 (57%) study subjects reach endpoint.

For model building activities, p-values will be used to identify the best fitting model as well as covariates to be included in the final models. For these model building activities, p-values <0.05 will generally determine significance. However, due to the exploratory nature of these models, less rigid standards may be considered and would be described fully in the final clinical study report and study manuscript.

7.8 Assessment Windows

For the primary efficacy outcome, final acute status of ROP is considered indeterminate if not obtained by 55 weeks PMA and thus, for the analyses for publication, requires endpoint adjudication. Any ROP assessments after 55 weeks PMA (including events of sROP) will be excluded from summaries/analyses of this outcome; however these assessments will be provided to the adjudicating ophthalmologist when determining survival without sROP outcomes (favorable vs. unfavorable ROP status) unless intervention was implemented prior to meeting the anatomic definition of unfavorable status prior to these assessments. Otherwise, all other data will be summarized and analyzed as collected. Additionally, the number of assessments obtained outside of window (e.g. final acute status for ROP determined after 55 weeks PMA) will be compared among study arms. If there are differences among study arms then sensitivity analyses that include/exclude assessments outside of study window will be conducted to evaluate if any results are sensitive to timing of assessments.

8 STUDY SUBJECT CHARACTERIZATION

8.1 Subject Disposition

Subject eligibility status will be summarized and listed by study arm. The number of subjects randomized; completing or discontinuing from study drug; reaching NRN status (defined in section 3.2.2) via being discharged, remaining in hospital, dying or transferring to another hospital; and completing 22-26 month follow-up (overall and also whether or not long term ROP status assessment was completed) will be summarized by study arm. Reasons for study drug discontinuation will be listed. Additional variables to be derived, listed and/or summarized include:

- Time until NRN status by type of status event (discharge, transfer, death): Date of NRN status – Date of Birth

8.2 Protocol Deviations

Protocol deviations are identified by site staff, monitors at monitoring visits, and also via automated checks of the clinical database. Protocol deviations will be listed by center with

information such as type of deviation, time of occurrence, and reason. Incidence rate of protocol deviations will also be summarized overall and for each protocol deviation category by center.

- Incidence rate of protocol deviations: number of deviations divided by the number of subject weeks at the center

8.3 Study Drug Exposure

Characteristics of study drug exposure will be summarized by study arm. Summaries will be created for all doses together as well as by dose type (i.e. IV vs. enteral). Characteristics include:

- Duration (days): Date of last study drug dose – Day 1 Date + 1 (Day 1 date is the date of study drug initiation)
- Days any therapy received: the number of days during which either of the 2 scheduled doses were given
- Total doses received
- Missed/held doses prior to permanent discontinuation
- Reasons for missed doses
- Average daily dose (mg): the sum of all reported doses divided by the days any therapy is received
- Average daily dose (mg/kg): similar to average daily dose (mg) except that daily dose (mg/kg) is calculated as reported dose/estimated daily weight where estimated daily weight is based on a linear interpolation between the previous and next available weight measurement (for dosing occurring after the last measured weight, the last measured weight will be carried forward for dose calculations

8.4 Demographic and Baseline Characteristics

Demographic and baseline characteristics for the study subjects will be summarized by study arm. Variables of interest include: Gestational age (weeks), gestational age stratum, birth weight, head circumference and length, age at start of study drug, sex, race, ethnicity, use of prenatal steroids, chorioamnionitis, delivery by cesarean section, 1 and 5 minute Apgar scores, use of chest compressions or resuscitation drugs in the delivery room, and early onset sepsis (<72 hours).

- Age at start of study drug will be calculated as: Day 1 Date– Date of birth

9 EFFICACY ANALYSES

9.1 Overview of Efficacy Analyses Methods

All efficacy analyses will be performed using the ITT population unless otherwise specified.

9.2 Efficacy Variables

Variable	Type	Definition
Primary Outcomes		
9.2.1.1 Death through NRN status	Binary indicator	=1 if death reported before NRN status (e.g. before discharge or transfer or 120 days CA) on CRF INS-3-13 =0 otherwise
9.2.1.2 Death prior to acute final ROP outcome (up to 55 weeks PMA)	Binary indicator	=1 if death occurred prior to acute final ROP status (up to 55 weeks PMA) reported on CRF INS-3-13 and CRF INS-3-08 =0 otherwise
9.2.1.3 sROP (or worse in at least 1 eye)	Binary indicator	=1 if Type I ROP or worse in either or both eyes (Type I ROP reported or retinal detachment reported as ROP of stage 4a, 4b or 5) during an assessment occurring prior to 55 weeks PMA on CRF INS-3-08 =0 if infant reached favorable status in both eyes at or prior to 55 weeks PMA where favorable status defined as fully vascularized to the ora serrata at one exam and PMA is greater than 35 weeks (vessel growth status reported as mature on CRF INS-3-08 on or after 35 weeks PMA and the highest stage of ROP for that assessment in any zone is 0-No ROP), or the retinal vessels have grown into Zone III, observed on 2 sequential exams, showing only immature vessels, regressing ROP or Stage 1 or Stage 2 ROP (lowest zone not fully vascularized is zone III for two sequential exams with highest stage of ROP in any zone on these exams ≤2 on CRF INS-3-08) or unequivocally Zone III vessels on one examination in an infant of 35 weeks PMA or older, if there has been no previous ROP in Zone I or Zone II, and this examination is not the only examination the infant has had (lowest zone not fully vascularized is zone III for one exam on or after 35 weeks PMA, and there is at least one prior exam reported with all prior exams where the highest stage of ROP in any zone is 'No ROP' on CRF INS-3-08) or PMA has reached 50 weeks without meeting sROP criteria and there is no prethreshold or worse ROP present where prethreshold ROP is defined as 1) Zone II: Stage 3 ROP, or Zone I: Any Stage ROP. (infant has assessment through 50 weeks PMA on CRF INS-3-08 and for those assessments, there is no plus disease. Additionally, there is no assessment where the lowest zone with any ROP is 'Zone I' and no assessment where the highest stage of ROP in any zone is 'Stage 3'). Otherwise, the endpoint will be considered missing (no available ROP assessment data) or incomplete (available ROP assessment data but final status definition not met). Refer to Sections 9.3, 9.4 and 12 to see how the missing and incomplete outcomes will be handled
9.2.1.4 ROP or Death before acute final ROP status (through	Binary indicator	=1 if either 9.2.1.2 Death prior to acute final ROP status or 9.2.1.3 sROP occur =0 if neither 9.2.1.2 Death prior to acute final ROP nor 9.2.1.3 sROP occur (both have values of 0)

Variable	Type	Definition
55 weeks PMA)		
Secondary Outcomes		
9.2.2.1 Any ROP	Binary indicator	=1 if either ROP of any stage in any zone of either eye determined on at least two assessments prior to 55 weeks PMA reported on CRF INS-3-08 (NOTE: assessments do not have to be consecutive but do have to occur in same eye) Or if there is only one examination prior to death, and it has any ROP reported on CRF INS-3-08 =0 otherwise if infant does not have ROP reported in either eye on any examination (or only on 1 examination per eye) and reached acute/final ROP status at or prior to 55 weeks PMA *Infants not meeting either criterion will be excluded from the efficacy analyses for this endpoint
9.2.2.2Type 2 ROP (or worse)	Binary indicator	=1 if one or both eyes reach Type 2 ROP or worse (Type 1 ROP is worse) prior to 55 weeks PMA reported on CRF INS-3-08. Type 1 ROP or worse is defined in **9.2.1.3 sROP (or worse in at least one eye)** and Type 2 ROP is defined as Stage 3 ROP without Plus Disease in Zone II, or Stage 1 or 2 ROP without Plus Disease in Zone I (there is an assessment on CRF-INS-3-08 prior to 55 weeks PMA where lowest zone with any ROP is 'Zone I' or where the highest stage of ROP in any zone is 'Stage 3'). =0 otherwise if infant does not meet Type 2 or Type 1 ROP definitions in either eye and reached acute/final ROP status at or prior to 55 weeks PMA *Infants not meeting either criterion will be excluded from the efficacy analyses for this endpoint
9.2.2.3 Traditional BPD (including deaths due to BPD)	Binary indicator	Defined similar to GDB traditional BPD definition except that deaths reported with contributing cause of BPD prior to 37 weeks PMA will also be classified as having BPD *Information will be obtained from GDB NG-07, INS-03-7 BPD supplemental and INS-03-13
9.2.2.4 BPD (physiologic definition)	Binary indicator	Traditional BDP definition is oxygen required at 36 weeks (including deaths reported with contributing cause of BPD prior to 37 weeks PMA). Physiologic BPD is defined similar to traditional BPD with two exceptions. 1. Infants receiving support via ventilator or CPAP at 36 weeks PMA are considered to have BPD regardless of whether they are receiving supplemental oxygen or room air. 2. Infants receiving low levels of supplemental oxygen (≤30%) at 36 weeks PMA may be eligible for a physiologic challenge in which there is an attempt to wean the infant to room air. Specifically, infants are eligible for the challenge if at 36 weeks PMA they are receiving effective oxygen <27% and have majority saturation ≥90%, or they are receiving effective oxygen 27-30% and have majority saturation ≥96%, or they are receiving room air by nasal cannula. The challenge takes place between 36 and 37 weeks PMA. Infants who are successfully weaned to room air during the challenge do not have BPD. Those who are weaned to room air before the challenge can take place also do not have BPD. Those who are not

Variable	Type	Definition
		challenged because their level of support increases (support with CPAP or vent or increased oxygen) are considered to have BPD, as are those who fail the challenge. Infants who are eligible for challenge but who are not challenged because of instability (including surgery or sepsis) or other reasons (such as personnel not available) are classified based on their level of support at 36 weeks. Infants receiving supplemental oxygen by any means at 36 weeks are considered by default to have BPD. Those receiving room air by nasal cannula at 36 weeks with flow >0.5 LPM will be classified as having BPD and those with flow ≤0.5 LPM as not having BPD. Infants who are receiving supplemental oxygen at 36 weeks PMA and who are not eligible for the physiologic challenge are considered to have BPD, and those who are breathing room air on their own (with no support) at 36 weeks do not have BPD. Infants who are transferred or discharged before 36 weeks are classified based on the support they are receiving at that time, and the follow up at 36 weeks PMA. For the physiologic definition: Those receiving supplemental oxygen or positive pressure support via CPAP or ventilator have BPD, and those breathing room air on their own do not have BPD. Those receiving room air via nasal cannula at discharge at a flow rate of less than 0.5 LPM will not have BPD, and those receiving it at 0.5 LPM or greater will have BPD. If BPD would otherwise be missing, an infant who is transferred or discharged on supplemental oxygen, ventilator or CPAP at ≤37 weeks PMA will be considered to have BPD (this is used to classify very few cases, if any). For INS-3, additional information is obtained about oxygen/respiratory support status of infants at 36 weeks if they are transferred or discharged to home prior to 36 weeks PMA and therefore don't have challenge results. This information will be used to augment the network definition of physiologic BPD in the following manner. Specifically, if such an infant is off all support at 36 weeks or on nasal cannula respiratory support at 36 week s PMA but flow ≤0 .5 LPM, they will be considered to not have BPD. Otherwise, if the infant is on nasal cannula respiratory support and flow >0 .5 LPM at 36 weeks they will be considered to have BPD. *Information will be obtained from GDB NG-07, INS-02-07 BPD supplemental and INS-02-13

<table>
<tr><th>Variable</th><th>Type</th><th>Definition</th></tr>
<tr><td colspan="3">Primary Outcomes</td></tr>
<tr><td>9.2.2.5 BPD (National Institutes of Health (NIH) Consensus definition)</td><td>Categorical variable with values of ‘none’, ‘mild’, ‘moderate’ and ‘severe’</td><td>The first 2 steps define the original consensus definition:
Step 1: If < 28 days of support and in room air (FiO_2=0.21) on Day 28, then assignment is No BPD
Step 2: If infant received oxygen or support for all 28 of the first 28 days, and
Is in room air at 36 weeks (no nasal cannula either), then assignment is mild BPD
Is receiving effective oxygen of < 30% (0.21 < FiO_2 < 0.30) and is not receiving ventilation or CPAP, then assignment is moderate BPD
Is receiving effective oxygen of ≥ 30% (FiO2 ≥ 0.30) and/or is receiving positive pressure ventilation (ventilation or CPAP) at 36 weeks, then assignment is severe BPD
The following refinements are for infants where BPD is still undefined/inconsistent:
<table>
<tr><th>Status At Day 28</th><th>Status At 36 Weeks</th><th>BPD Status According to Original Definition</th><th>Classification for NRN and this protocol [1]</th></tr>
<tr><td><28 days & Not in RA</td><td>≥28 days & Not in RA</td><td>Not defined</td><td>Moderate or Severe</td></tr>
<tr><td><28 days & Not in RA</td><td><28 days & Not in RA</td><td>Not defined</td><td>Moderate or Severe</td></tr>
<tr><td><28 days & Not in RA</td><td>≥28 days & in RA</td><td>Not defined</td><td>Mild</td></tr>
<tr><td><28 days & in RA</td><td>≥28 days & in RA</td><td>Defined/questionable</td><td>None</td></tr>
<tr><td><28 days & in RA</td><td>≥28 days & Not in RA</td><td>Defined/inconsistent</td><td>Moderate or Severe</td></tr>
<tr><td><28 days & in RA</td><td><28 days & Not in RA</td><td>Defined/inconsistent</td><td>Moderate or Severe</td></tr>
<tr><td><28 days & Not in RA</td><td><28 days & in RA</td><td>BPD=not clear</td><td>None</td></tr>
</table>
RA = room air
[1] For those with classification of Moderate or Severe, severity is determined from the oxygen concentration level (FiO_2) at 36 weeks in accordance with the published definition.

*Information will be obtained from GDB NG-07 and INS-02-7 BPD supplemental NG-07 form</td></tr>
<tr><td>9.2.2.6 Severe IVH</td><td>Binary indicator</td><td>=1 if grade 3 or 4 IVH reported as defined in Section 10.4 or cystic PVL reported regardless of whether sonogram was reported or shunted hydrocephalus
=0 if subject was reported as not having grade 3 or 4 IVH or cystic PVL or shunted hydrocephalus
*Information will be obtained from GDB NG-07</td></tr>
<tr><td>9.2.2.7 Any ROP or Death before acute final ROP status (through 55 weeks PMA)</td><td>Binary indicator</td><td>=1 if either 9.2.1.2 Death prior to acute final ROP or 9.2.2.1 Any ROP occur
= 0 if neither 9.2.1.2 Death prior to acute final ROP nor 9.2.2.1 Any ROP occur (both have values of 0)</td></tr>
<tr><td>9.2.2.8 Type 2 ROP or death before acute final ROP status (through 55 weeks PMA)</td><td>Binary indicator</td><td>=1 if either 9.2.1.2 Death prior to acute final ROP or 9.2.2.2 Type 2 ROP occur
= 0 if neither 9.2.1.2 Death prior to acute final ROP nor 9.2.2.2 Type 2 ROP occur (both have values of 0)</td></tr>
</table>

Variable	Type	Definition
Primary Outcomes		
9.2.2.9 Traditional BPD (including deaths due to BPD) or death (prior to 37 weeks PMA)	Binary indicator	=1 if either **9.2.2.3 Traditional BPD (including deaths due to BPD)** or death prior to 37 weeks PMA as reported on CRF INS-2-15 occur = 0 if neither **9.2.2.3 Traditional BPD (including deaths due to BPD)** or death prior to 37 weeks PMA reported on CRF INS-2-15 occur
9.2.2.10 BPD (physiologic definition) or death (prior to 37 weeks PMA)	Binary indicator	=1 if either **9.2.2.4 BPD (physiologic definition)** or death prior to 37 weeks PMA as reported on CRF INS-3-13 occur = 0 if neither **9.2.2.4 BPD (physiologic definition)** or death prior to 37 weeks PMA reported on CRF INS-3-13 occur
9.2.2.11 Severe IVH or death through network status	Binary indicator	=1 if either 9.2.2.6 Severe IVH or 9.2.1.1 Death through network status occur = 0 if neither 9.2.2.6 Severe IVH nor 9.2.1.1 Death through network status occur
9.2.2.12 CA at Time of Death through network status	Time to event variable for survival analysis	Date of 9.2.1.1 Death through network status– Date of Birth Infants surviving through network status (discharged, transferred or 120 days CA) will be censored at the time of network status.
9.2.2.13 CA at Time of sROP Diagnosis or Death before acute final ROP status (through 55 weeks PMA)	Time to event variable for survival analysis	[Date of whichever of the following outcomes occur 9.2.1.3 sROP (or worse in at least 1 eye) or 9.2.1.2 Death prior to acute final ROP outcome (up to 55 weeks PMA)] – Date of Birth Infants not meeting criteria for either event will be censored at last recorded ROP assessment prior to 55 weeks (or last assessment prior to intervention if intervention given prior to 55 weeks and meeting anatomic sROP definition).
9.2.2.14 CA at Time of Any ROP or Death before acute final ROP status (through 55 weeks PMA)	Time to event variable for survival analysis	[Date of whichever of the following outcomes occur 9.2.2.1 Any ROP or 9.2.1.2 Death prior to acute final ROP outcome (up to 55 weeks PMA)] – Date of Birth Infants not meeting criteria for either event will be censored at last recorded ROP assessment prior to 55 weeks.

9.3 Primary Analyses Methods for Scientific Publication

Incidence of survival without sROP (outcome 9.2.1.4) will be listed and summarized. The primary analysis for overall study of this outcome will include adjudicated endpoints such that adjudicated outcomes of 'most likely did NOT reach criteria' for surgery threshold are treated as 'favorable' outcomes and 'most likely did reach criteria' for surgery threshold are treated as 'unfavorable' outcomes. It is anticipated that <5% of infants will have adjudicated outcomes of 'insufficient data' and also that these outcomes will be approximately evenly divided between treatment groups. Accordingly, adjudicated outcomes of 'insufficient data' will be treated as missing completely at random and will be excluded from the primary analyses.

Robust Poisson regression adjusting for center and GA strata will be used to estimate adjusted relative risk (and associated 95% CI) for myo-Inositol 5% Injection vs. placebo (McNutt, Wu, Xue , and Hafner 2003; Greenland 2004; Zou 2004). The primary hypothesis will be tested based on a score test using the robust variance estimator of the treatment effect in the Poisson regression model where the null hypothesis of no treatment effect on survival without severe ROP is rejected at either the interim or final analysis with a p-value consistent with an overall Type 1 error rate of 0.05 and Bonferroni-type correction for multiple comparisons (Casella and Berger 2002).

These analyses will be implemented with the following SAS code where treatment assignment, RNDTRT is 0 for placebo and 1 for active treatment and a binary sROP value of 0 denotes a positive outcome (i.e. favorable status) and 1 a negative outcome (i.e. unfavorable outcome or death prior to final ROP status). The mean estimate and confidence limits provided by the ESTIMATE statement are the adjusted relative risk and associated confidence interval.

```
PROC GENMOD;
CLASS RNDTRT CENTER GASTRATA SUBJID;
MODEL SROP = RNDTRT CENTER GASTRATA/ DIST=POISSON LINK=LOG;
REPEATED SUBJECT= SUBJID;
ESTIMATE 'TREATMENT EFFECT' RNDTRT -1 1 / EXP;
RUN;
```

The analysis methods for the primary efficacy outcome for regulatory submission are detailed in Section 12. Formal inference based on the regulatory analyses is dependent on the findings of the primary analysis for scientific publication demonstrating a benefit (i.e. null hypothesis rejected at either the interim or final analysis with a p-value consistent with an overall Type 1 error rate of 0.05 and Bonferroni-type correction for multiple comparisons). This prioritization of analyses precludes the necessity of adjusting alpha for the two study purposes. Sensitivity analyses will also be performed to assess the impact of missing/indeterminate outcomes including a repeat of the primary analysis excluding infants with adjudicated outcomes, using multiple imputation for all missing and indeterminate outcomes, and lastly treating all missing outcomes after adjudication a) as treatment failures and also b) as treatment successes.

If the results of the primary analysis of this multi-centered trial are positive then sensitivity analyses of the primary outcome will also be performed to examine internal consistency of the study results. Specifically, consistency of treatment effect across subsets defined by GA strata and center separately will be assessed using the robust Poisson regression model. For GA strata, an interaction between GA strata and treatment will be added to the model. A p-value for the interaction >0.2 will be indicative of no interaction effect. If heterogeneity of treatment effect is present then the likelihood ratio test detailed in Gail and Simon (1985) will be used to determine whether the interactions are qualitative in nature (i.e. a test of the null hypothesis that myo-Inositol 5% Injection is at least as good as placebo in every subset of patients vs. the alternative hypothesis that a cross-over effect exists such that placebo outperforms myo-Inositol 5% Injection in at least one of the subset of patients). This same approach will also be used to assess consistency of treatment effect across centers. However, to ensure adequate power for identifying qualitative interactions among centers, the consistency analyses will be conducted with centers clustered into four subgroups defined by baseline prevalence of unfavorable ROP status (high vs. low) and expected enrollment size (large vs. small).

9.4 Secondary Analyses Methods

Incidence of individual components of the primary outcome (outcomes 9.2.1.2 and 9.2.1.3) and of secondary efficacy outcomes for other binary indicators of mortality, ROP, BPD and IVH (outcomes 9.2.1.1 and 9.2.2.1-9.2.2.11 excluding 9.2.2.5) will be summarized and analyzed in a similar manner to the primary analyses methods of 10.3. BPD using the NIH consensus definition (outcome 9.2.2.5) will also be summarized and analyzed. P-values for this outcome will be based on a Mantel-Haenszel mean score test using modified ridit scores. Additional descriptive summaries of ROP will be constructed including the number and percent of subjects falling into the following categories experiencing: 1) Type 1 ROP, 2) Type 2 ROP but no worse (overall as well as for those that reached final acute status and those that did not), 3) Any ROP less than Type 2 ROP (overall as well as for those that reached final acute status and those that did not), and 4) never reported any ROP (overall as well as for those that reached final acute status and those that did not).

Kaplan-Meier survival curves by study arm will be created for the PMA outcomes (outcomes 9.2.2.12-9.2.2.14). These outcomes will also be listed and summarized by median and range. No p-values will be calculated for these outcomes.

10 SAFETY ANALYSES

10.1 Overview of Safety Analyses Methods

All safety analyses will be performed using the safety population (i.e., as treated) unless otherwise specified. Descriptive p-values comparing the study arms will be provided on most safety table summaries and will be obtained using chi-square tests for binary outcomes. If the number of events allow, a 2-sided Cochran Mantel-Haenszel test controlling for strata defined by study center and gestational age will be used to obtain the p-values.

10.2 Adverse Events

Reportable AEs include events starting or worsening in severity after start of study drug through 7 days after last study drug dose. AE will be reported and graded using Toxicity Table for Premature Neonates: NICHD NRN (Appendix E of the Manual of Procedures) that was modified from the ©Baylor College of Medicine Neonatal Toxicity Table (2001). Using this table, events will be listed and summarized by category and preferred event term. Summaries will be of the number of individuals experiencing events and will be created for all AEs, AEs by severity, and AEs by relationship to treatment. Summaries will be done for the number and percent of subjects per study arm experiencing an AE. Any events starting outside of the reportable timeframe will be included in separate listings and will be excluded from summary tables.

For the displays above, only on-study AEs will be included. On-study AEs include events starting on or after Day 1 and prior to 7 days after last dose. If a complete onset date is unknown and it cannot be confirmed that the event occurred during this time period, then the event will be considered an on-study AE.

10.3 Deaths and Serious Adverse Events

An SAE is any event that is life threatening, results in death, causes or prolongs hospitalization, leads to a disability or birth defect, or requires an intervention to prevent a disability.

SAEs will be listed and SAEs, treatment-related SAEs and SAEs with an outcome of death will be summarized in the manner mentioned in Section 10.2 pending there are enough events to summarize. Separate displays listing and summarizing deaths occurring after start of study drug through 7 days after last study drug dose including age at death and cause of death (including primary and contributing causes) will also be created.

10.4 Clinical Outcomes

In addition to being captured and summarized as part of the adverse events if occurring during the AE reporting period, clinical outcomes are also reported as part of the NRN Generic Database (GDB) and therefore will also be summarized separately from the other adverse event summaries. These clinical outcomes are collected from birth to time of reaching NRN status. These include:

- Death (from birth to time of NRN status, i.e. maximum of 120 days CA)
- BPD (oxygen required at 36 weeks PMA)
- Respiratory distress (w/i first 24 hours of life): any, demonstrating clinical features, requiring >6 hours oxygen or positive pressure
- Pulmonary hemorrhage

- PDA: any, treated with indomethacin and/or ibuprofen, treated with surgery
- IVH: any, grade III/IV defined blood/echodensity in the parenchyma (grade IV) or ventricular size enlarged and blood/echodensity in the ventricle (grade III)Seizures (treated for >72 hours)
- Cystic area(s) in the brain parenchyma
- Sepsis: early onset, late onset
- NEC: any defined as Bell's stage II or higher, or treated with surgery
- Spontaneous gastrointestinal (GI) perforation: without NEC
- Hearing test failure: never passed in one or both ears

Most of these clinical outcomes are reported via yes/no indicators in the GDB; for those that are not, standard NRN GDB definitions will be used for their derivation. Specifically, see Section 9.2 for details of the definitions that will be used for BPD (BPD is considered both a clinical outcome as well as a secondary efficacy outcome).

Clinical outcomes will be listed and the number of individuals experiencing each type of event will be summarized. Tables will also include 2-year historical rates of these morbidities among infants of matching GA born at all of the centers of the NICHD NRN. Death through NRN status summaries for age at death and cause of death (including primary and contributing causes) will also be created.

10.5 Concomitant Medications

Concomitant medications starting from 24 hours prior to randomization until earliest of discharge, transfer, or 7 days after the last dose of study drug administration will be recorded. These medications will be listed and the number of subjects receiving each medication will be summarized by each reported term as well as drug class. Any medications starting outside of the reportable period will be listed separately and excluded from summary tables. Summaries of concomitant medications by treatment group will be created for all study subjects together and by GA strata.

11 ANALYSES OF OTHER OUTCOMES

11.1 Short Term Growth

Weight (g), length (cm), head circumference (cm) and z-scores for growth from birth through end of study drug will also be collected as part of this study and will be listed and plotted by study arm over time using the ITT study population.

- Z-scores for growth are based on the growth charts provided by Riddle and DonLevy.

11.2 Long Term Outcomes Collected Using the NICHD Follow-up Protocol

Mortality through 22-26 months of age, completion of 22-26 month follow-up visit and specific outcomes of interest collected at the 22-26 month follow-up visit will be listed and summarized. Interval medical history includes incidence of the following events: rehospitalization; operations; repeated receipt of medication in past 3 months; and current use of apnea monitor, oxygen, ventilator/CPAP, gastrostomy tube and/or tube feeding, tracheostomy, or pulse oximeter at home or in chronic use facility. Other outcomes include growth measures of weight (kg), recumbent length (cm), and occipital-frontal circumference (cm); Bayley III cognitive, receptive and expressive communication, and fine and gross motor function classification system (GMFCS) scores; incidence and severity of cerebral palsy; as well as neurodevelopmental impairment (NDI). Additionally, ROP diagnoses and their treatment as reported by parents at the 22-26 month follow-up will be listed and summarized. The majority of listings and summaries for outcomes other than mortality will only include infants in the ITT population completing the 22-26 month follow-up.

However, additional summaries of the following outcomes will also be created that include all subjects in the ITT population: ROP or Death (death being prior to ROP acute status and prior to 22-26 months of age) and NDI and/or Death (prior to 22-26 months of age). Robust Poisson regression adjusting for center and GA strata will be used to estimate adjusted relative risk and associated 95% CI for these outcomes.

- Occurrence of ROP will include any ROP reported during the 22-26 month follow-up visit as well as any ROP reported prior to 55 weeks PMA (described in section 9)

- NDI will be defined as occurrence of any of the following:

    - Moderate or severe cerebral palsy

    - Bayley III cognitive composite score < 85

    - Bayley III motor composite score < 85

    - Limited vision with correction, unilateral blind or bilateral blind

    - Permanent hearing loss that does not permit child to understand directions of the examiner and communicate despite amplification with cochlear implant or hearing aid

Ophthalmic treatments (including all eye surgeries) as reported by the family will be listed and summarized.

12 ANALYSES FOR REGULATORY SUBMISSION PURPOSES

In support of a new drug application (NDA) for myo-Inositol 5% Injection for prevention of sROP, the primary outcome of survival without sROP will be analyzed within sub-studies created by administratively splitting the infants enrolled at each of the study centers into two groups (i.e. ‘regulatory sub-studies’) based on the pre-defined criteria described below.

12.1 Regulatory Sub-Study Administrative Split

For regulatory submission purposes the study population will be administratively split in a pre-defined manner into two regulatory sub-studies designated "Study A" and "Study B". These studies will be analyzed separately. The administrative split will be performed within each center to minimize center effects between the split studies which become more likely when the whole study is enrolled from only 18 study centers that are geographically dispersed unevenly across the country. This approach is preferred because this method of splitting the data ensures that variations in the racial, ethnic makeup of the regional population, and baseline risks for ROP would be distributed reasonably evenly between Study A and Study B. In addition, approximately equal numbers would be expected in each sub-study, permitting more accurate estimate of the power of each sub-study.

As a part of the randomization system each study subject will be assigned a "Study Allocation" number based on a Uniform (0, 1) distribution using a pseudo random number generator. That assignment will be maintained in the randomization file, with no study staff other than the randomization statistician and programmer having access to the Study Allocation number until the study database has been locked. After database lock, subjects will be allocated to either "Study A" or "Study B" using the following algorithm.

- Within each cell defined by center, gestational age stratum, and treatment arm, subjects will be rank ordered by their Study Allocation number.

- For cells with an even number of subjects those subjects that are in the 50% with the smaller Study Allocation numbers will be assigned to "Study A" and those subjects that are in the 50% with the larger Study Allocation numbers will be assigned to "Study B."

- For those cells with an odd number of subjects, those subjects with Study Allocation numbers less than the median Study Allocation number will be assigned to "Study A" and those subjects with Study Allocation numbers greater than the median Study Allocation number will be assigned to "Study B."

- The subject with the median Study Allocation number will be assigned to "Study A" if the Study Allocation number is less than or equal to 0.5 and will be assigned to "Study B" if the allocation number is greater than 0.5.

While this process does not allocate subjects to "Study A" or "Study B" at the time of enrollment, the Study Allocation number is assigned at the time of treatment randomization and therefore regulatory sub-study assignment will be determined *a priori*.

12.2 Analysis Method

As suggested by the FDA, for the primary regulatory analyses, the primary outcome will be treated as a three level ordinal outcome where the unfavorable outcomes are assigned a value

of -1, indeterminate outcomes are assigned a value of 0 and favorable outcomes are assigned a value of +1. See Table 3.

Table 3: Subject Disposition by Outcome Category (All randomized patients)

Outcome Status	Placebo (N1)	Inositol (N2)	Difference
Favorable (1)	n_{11} (n11/N1)	N_{21} (n21/N2)	XX
Unfavorable (-1)	n_{12} (n12/N1)	N_{22} (n22/N2)	XX
Indeterminate (0)	n_{13} (n13/N1)	N_{23} (n23/N2)	XX
Missing (by MI)	n_{14} (n14/N1)	N_{24} (n24/N2)	XX

Within each sub-study, the primary endpoint will be tested using a 2-sided Mantel-Haenszel chi-square test using modified ridit scores for ordinal data controlling for strata defined by study center and gestational age.

Individuals with completely missing outcomes will be imputed using the multivariate technique for multiple imputation (MI) described by Raghunathan et. al. (2001). Imputations will be based on the demographic and baseline characteristics listed in Section 8.4: gestational age (weeks), birth weight, head circumference and length, age at study drug start, sex, race, ethnicity, use of prenatal steroids, chorioamnionitis, delivery by cesarean section, 1 and 5 minute Apgar scores, use of chest compressions or resuscitation drugs in the delivery room, and early onset sepsis. If there are randomized infants that are not treated, then age at time of randomization will be used in place of age at study drug start for all randomized infants in the MI models. If any of these characteristics are missing for any infants, then values for these characteristics will be imputed as part of the conditional MI modeling so that the characteristics can be included as predictors for imputation of the missing outcomes. The conditional MI regression model used for each characteristic requiring imputation will be based on the type of data requiring imputation (e.g. linear regression model for continuous variables, logistic model for binary variables, and generalized logit model for categorical data). Specifically, an ordinal logistic regression model will be used to impute values of -1, 0 or 1 for the primary outcome. The MI process will be repeated 10 times resulting in 10 imputed datasets. The mean score for the myo-Inositol 5% Injection treatment group across strata and its variance, the basis for the Mantel-Haenszel chi-square test statistic for ordinal data, will be obtained for each imputed data set. The overall mean score and variance estimate will then be calculated based on formulae presented in Rubin (1978) and used to calculate the final Mantel-Haenszel chi-square test statistic and corresponding p-value.

The sensitivity analyses specified for the overall study analysis in Section 9.3 for assessing the impact of indeterminate and missing outcomes will also be repeated for each regulatory sub-study. The Robust Poisson regression analysis described in Section 9.3 will also be provided as part of the regulatory submission as an overall assessment of the strength of the efficacy results.

13 REPORTING CONVENTIONS

Unless required otherwise by a journal, the following rules are standard:

- Moment statistics including mean and standard deviation will be reported at 1 more significant digit than the precision of the data.
- Order statistics including median, min and max will be reported to the same level of precision as the original observations. If any values are calculated out to have more significant digits, then the value should be rounded so that it is the same level of precision as the original data.
- Following SAS rules, the median will be reported as the average of the two middle numbers if the dataset contains even numbers.
- Test statistics including t and z test statistics will be reported to two decimal places.
- P-value will be reported to 3 decimal places if > 0.001. If it is less than 0.001 then report '<0.001'. Report p-values as 0.05 rather than .05.
- No preliminary rounding should be performed, rounding should only occur after analysis. To round, consider digit to right of last significant digit: if <5 round down, if >=5 round up.

14 CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

There are no changes to the analyses planned in the protocol to date.

15 LIST OF POTENTIAL DISPLAYS

Data displays may be added, deleted, rearranged or the structure may be modified after finalization of the SAP. Such changes require no amendment to the SAP as long as the change does not contradict the text of the SAP.

Tables
Subject Eligibility
Subject Disposition
Protocol Deviations
Study Drug Exposure
Demographic and Baseline Characteristics
Number of Subjects Experiencing Each AE (overall, by severity, by relationship to study drug)
Number of Subjects Experiencing Each SAE (overall, fatal, related to study drug)
Mortality
Clinical Outcomes
Concomitant Medications
Primary Efficacy Results
Primary Efficacy by GA Strata
Primary Efficacy by Center Cluster

Sensitivity Analyses for Primary Efficacy Results Secondary Efficacy Outcomes Regulatory Efficacy Analysis Results Sensitivity Analyses for Regulatory Efficacy Analysis Results Long Term Outcomes
Figures
Survival Curves of Time to Primary and Secondary Efficacy Outcomes (plotted for each outcome) Short Term Growth Over Time (plotted for each growth outcome)
Data Listings
Subject Eligibility Subject Disposition Protocol Deviations Study Drug Exposure Demographic and Baseline Characteristics Reportable, On-study Adverse Events Adverse Events Not Included In Summaries Serious Adverse Events Mortality Clinical Outcomes Concomitant Medications Primary and Secondary Efficacy Outcomes Short Term Growth Long Term Outcomes

16 REFERENCES

Barkovich AJ. Brain and spine injuries in infancy and childhood. In Barkovich AJ. *Pediatric Neuroimaging.* 3rd ed. Philadelphia, PA: Lippencott, Williams, &Wilkins; 2000:179-84.

Casella G. and Berger R.L. *Statistical Inference.* California: Duxbury. 2002.

Cryotherapy for Retinopathy of Prematurity Cooperative Group. Multicenter Trial of Cryotherapy for Retinopathy of Prematurity: Preliminary Results. *Arch Ophthal.* 1988;106:471-479.

(ETROP) Early Treatment for Retinopathy of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity. Arch Ophthlamol. 2003;121:1684-96.

Gail M. and Simon R. Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics.* 1985 Jun: 41(2): 361:72.

McNutt LA, Wu C, Xue X, Hafner JP. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. Am J Epidemiol 2003; 157(10):940-3.

Papile L-A, Burstein J, Burstein R, et al: Incidence and evolution of subependymal and intraventricular hemorrhage: A study of infants with birth weights less than 1500 gm. *J Peds.* 1978;92:529-34.

Riddle, W. R., and S. C. DonLevy. Generating expected growth curves and Z-scores for premature infants. *Journal of Perinatology.* 2010:30.11:741-750.

Sander Greenland, Model-based Estimation of Relative Risks and Other Epidemiologic Measures in Studies of Common Outcomes and in Case-Control Studies, American Journal of Epidemiology 2004;160:301-305.

SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network, Carlo WA, Finer NN, Walsh MC, Rich W, Gantz MG, Laptook AR, Yoder BA, Faix RG, Das A, Poole WK, Schibler K, Newman NS, Ambalavanan N, Frantz III ID, Piazza AJ, Sanchez PJ, Morris BH, Laroia N, Phelps DL, Poindexter BB, Cotten CM, Van Meurs KP, Duara S, Narendran V, Sood VBG, O'Shea TM, Bell EF, Ehrenkranz RA, Watterberg KL, Higgins RD. Target ranges of oxygen saturation in extremely preterm infants. New England Journal of Medicine. 2010;362:1959-69.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology*, 27(1), 85-96.

Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, Inc. 1978.

Walsh, MC, Kliegman R. Necrotizing enterocolitis: Treatment based staging criteria. *Pediatric Clinics of North America.*1986; 33(1):179-201.

Walsh MC, Yao Q, Gettner P, et al. Impact of a physiologic definition on bronchopulmonary dysplasia rates. *Pediatrics.* 2004;114(11):1305-11. PMID: 15520112.

Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004; 159(7):702-6.