STATISTICAL ANALYSIS PLAN

Randomized Controlled Trial of Home Therapy with Caffeine Citrate in Moderately Preterm Infants with Apnea of Prematurity

(NCT03340727)

SAP VERSION: 1.0

SAP DATE: December 12, 2022

SPONSOR: NICHD Neonatal Research Network

PREPARED BY: RTI International.
3040 Cornwallis Rd
Research Triangle Park, NC 27709-2104

AUTHOR (S): Benjamin Carper
Abhik Das

Contents

LIST OF ABBREVIATIONS

Abbreviation:	Deciphered:
AE	Adverse Event
ALTE	Apparent Life-Threatening Event
AoP	Apnea of Prematurity
BPM	Beats per Minute
CPAP	Continuous Positive Airway Pressure
DCC	Data Coordinating Center
DSMC	Data Safety and Monitoring Committee
GA	Gestational Age
IRB	Institutional Review Board
ITT	Intent-to-Treat
LOS	Length of Stay
NICHD	National Institute of Child Health and Human Development
NICU	Neonatal Intensive Care Unit
NIH	National Institutes of Health
NPO	Nothing by Mouth
NRN	Neonatal Research Network
PI	Principal Investigator
PMA	Post-menstrual Age
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System

1 BACKGROUND

Apnea of Prematurity (AoP) is a common diagnosis in moderately preterm infants, most exhibiting clinical signs of AoP as defined by the American Academy of Pediatrics [cessation of breathing for ≥ 20 seconds or < 20 seconds if accompanied by bradycardia and/or cyanosis in infants <37 weeks gestational age (GA)]. Moderately preterm infants comprise one fourth of NICU admissions. In a retrospective cohort study of moderately preterm infants, 92% (23/25) were diagnosed with AoP at 30 weeks post-menstrual age (PMA). Even at 33 weeks PMA, 48% (59/122) of infants had AoP, and apnea may persist beyond 37 weeks, even until ~44 weeks PMA. PMA at the time of resolution of apnea is inversely proportional to the GA at birth. Although respiratory maturation occurs with increased PMA[4], there is uncertainty regarding the range of PMA at which there is complete resolution of AoP. This uncertainty can delay discharge, as sufficiently mature respiratory control is one of the essential physiologic competencies before discharge of the infant as recommended by the American Academy of Pediatrics[9].

Three retrospective cohort studies of caffeine treatment at home for infants at risk for apnea and infants with apnea of infancy have been reported. The median PMA at last dose of caffeine was 34.4 weeks in a large clinical trial while in another large multicenter study, the average PMA at discharge was 35.7 weeks (encompassing infants born at 30 to 33 weeks GA). This illustrates many infants continue to receive treatment late in their hospital course. As apnea can still occur past these PMAs, continuing caffeine beyond the usual duration of treatment through hospital discharge may decrease late episodes of apnea that occur in the hospital. Often, these late apneas can prolong hospital stay due to an additional period of monitoring for further apnea. In addition, continuation of caffeine beyond hospital discharge may prevent complications from AoP that lead to hospital readmission or unplanned healthcare visits. A trial investigating the use of prolonged caffeine citrate treatment would thus provide high level of evidence and generalizability to determine the effects on length of hospitalization, readmissions, and sick visits in moderately preterm infants. This trial will evaluate infants for eight weeks following discharge, since infants would have reached at least a PMA of 40 to 44 weeks at the end of the evaluation period. Thus, the number of events of apnea and bradycardia should not differ when compared to full term infants of the same PMA. This is the reasoning for the eight-week evaluation period following hospital discharge.

2 PURPOSE OF THE ANALYSIS

This statistical analysis plan (SAP) contains detailed information about statistical analysis to be performed to evaluate the effect of continuing treatment with caffeine citrate in the hospital and at home in moderately preterm infants with resolved apnea of prematurity on days of hospitalization after randomization.

3 STUDY AIMS AND OUTCOMES

3.1 Study Aims

3.1.1 Primary Aim

To determine efficacy of continuing treatment with caffeine citrate compared to placebo in the hospital and at home in moderately preterm infants with resolved apnea of prematurity for the reduction in days of hospitalization after randomization.

3.1.2 Secondary Aims

1. To evaluate the effect of continuing treatment with caffeine citrate in the hospital on number of days to reach physiologic maturity from randomization to discharge up to 48 wks PMA.

2. To evaluate the effect of treatment with caffeine until 28 days after hospital discharge on all-cause readmissions to a hospital (including apneic and/or apparent life-threatening events) within the first 8 weeks after discharge by 48 wks PMA.
3. To evaluate the effect of treatment with caffeine until 28 days after hospital discharge on the number of all-cause sick visits to urgent care, emergency rooms, or health care provider's office (including apneic and/or apparent life-threatening events) within the first 8 weeks after discharge by 48 wks PMA.
4. To evaluate the side effects and adverse events of caffeine continuation in the hospital and after discharge up to 4 weeks post study drug administration.
5. To assess the hypothetical cost-savings by the reduction in number of days of hospitalization as a financial benefit of the trial's interventions.

3.2 Outcomes

3.2.1 Primary Outcome

The primary outcome of the study will be the number of days of hospitalization from randomization to discharge up to 48 weeks PMA, with censoring at time of transfer or death.

3.2.2 Secondary Outcomes

1. The number of days to physiologic maturity after randomization up to 48 wks PMA. Physiologic maturity will be defined as follows:
   a) Temperature: out of the incubator for at least 48 hours with normal body temperature
   b) Feeding: oral feeding at a volume of at least 140 ml/kg/day or growing on less than 140 ml/kg/day for at least 48 hours
   c) Respiratory: apnea-free for at least 5 consecutive days
2. Post menstrual age at discharge up to 48 wks PMA, censoring at time of transfer or death (status).
3. The number of all cause readmissions, including apneic events and apparent life-threatening events (ALTE), to the hospital within the first four weeks, second four weeks, and first eight weeks combined after discharge from the hospital by 48 wks PMA.
4. The number of all cause sick visits to urgent care, emergency room or health care provider office within the first four weeks, second four weeks, and first eight weeks combined after discharge from the hospital by 48 wks PMA.
5. Safety outcome measures pre-hospital discharge:
   a) Weight gain from randomization until status (discharge up to 48 wks PMA, with censoring at time of transfer or death). The weight gain may also be assessed using percentiles including birthweight percentiles.
   b) The number of days after randomization until status that infant had at least two consecutive heart rates >200 documented at least 3 hours apart (when infant not crying).
   c) Treatment for high blood pressure initiated after randomization until status.
   d) The number of episodes between randomization and status that infant was placed NPO for ≥ 24 hours.
   e) The use of anti-reflux medications started between randomization and status
   f) The number of days that significant apnea/bradycardia, as defined by documentation of infant receiving any of the following between randomization and status: open label caffeine, other methylxanthines, doxapram, continuous positive airway pressure (CPAP) or ventilatory support for apnea/bradycardia.
   g) The presence of documented and treated arrhythmias between randomization and status, not due to tachycardia or bradycardia.

h) The onset of documented seizures, as defined by treating with anti-convulsants, between randomization and status.

6. Death

4 STUDY METHODS

4.1 Overall Study Design and Plan

This study will enroll moderately preterm infants who are between $33^{0/7}$ and $35^{6/7}$ weeks GA (inclusive) who are being treated with caffeine for AoP and for whom plans are to discontinue caffeine. Once the infant has been screened for eligibility, consented, and the neonatologist has determined that caffeine will be stopped, infants will be eligible to be randomized if they can be started on the study medication within 72 hours after stopping caffeine. They will be randomized either to continuing caffeine citrate at a dose of 10 mg/kg/dose (5 mg/kg caffeine base) given daily (treatment group) or an equal volume of placebo given daily (placebo group), not less than 12 hours between doses, to continue during the hospital stay and for 28 days after hospital discharge. In-hospital dosages will be adjusted weekly based on weight. Post-discharge dosage will be based on weight obtained within 48 hours of discharge.

Trial Arms		Trial Intervention
Intervention Group	Caffeine	Caffeine citrate 10 mg/kg/day given once daily until 28 days after discharge
Control Group	Placebo	Placebo contains all of the excipients except for the active ingredient, caffeine citrate (a volume equivalent to 10 mg/kg of caffeine citrate) given once daily for 28 days after discharge.

The trial medication (caffeine citrate or placebo) will be weight-adjusted on a weekly basis until discharge. An apnea countdown will be done as per local practice before discharge. The duration of this countdown will be site dependent but pre-specified. Many infants may require hospitalization for observation after resolution of the apneas because of other discharge issues such as temperature control or feeding tolerance. If the attending neonatologist decides that re-initiation of open-label caffeine is clinically indicated, caffeine/placebo treatment per randomization will be put on hold until the physician decides to discontinue open label caffeine at which time, the infant will be treated as per randomization within 24-72 hours after open label caffeine is stopped. Infants on study medication who are still in the hospital at 44 weeks PMA will have the caffeine/placebo discontinued.

Prior to discharge, a weight obtained within the last 48 hours will be used to determine the post-discharge dosage for those infants continuing with caffeine citrate or placebo at home. Once deemed ready for discharge, the parent will be supplied with 28 numbered vials with oral caffeine citrate (intervention group) or placebo at an equivalent volume (placebo group). The parents will be educated by the research nurse, discharge nurse, physician, or pharmacist on storage and administration of study medication. Infants will be continued at home on the same dose of caffeine citrate or placebo for the first 28 days after hospital discharge. A member of the research team will contact the parents to obtain post-discharge information within 72 hours after discharge, once a week for the first 4 weeks, and biweekly during the weeks 5 to 8 after discharge.

Caffeine citrate therapy or placebo, depending on the randomized treatment allocation, will be prepared by the study pharmacist for administration in the NICU. Study pharmacist will prepare

daily doses of study drug (caffeine or placebo) for administration while in the hospital. At discharge, study pharmacist will dispense a 28-day supply of either caffeine or placebo vials for home administration. All other research staff, health care staff, and family members will be blinded to the treatment allocation. The study medication can be prepared by the research pharmacist within 48 hours of anticipated discharge date.

4.2 Study Population

4.3 Participant Characteristics

The study population is defined by the following eligibility criteria:

Inclusion Criteria:

1. $33^{0/7}$ - $35^{6/7}$ weeks PMA at the time of randomization
2. Receiving caffeine with plan to discontinue treatment or just discontinued caffeine treatment
3. Receiving full feeds (defined by a volume of $\geq$120 ml/kg/day) by oral and/or tube feeding
4. Ability to start study medication within 72 hours after stopping caffeine

Exclusion Criteria:

1. On respiratory therapy (oxygen more than room air equivalent for high altitude sites, nasal cannula, continuous positive pressure ventilation, and/or mechanical ventilation)
2. Infants who would otherwise be discharged home on apnea monitor due to underlying disease or family history, including history of a sibling with sudden infant death syndrome
3. Parental request for apnea monitor
4. Congenital heart disease other than atrial septal defect, ventricular septal defect, or patent ductus arteriosus
5. Neuromuscular conditions affecting respiration
6. Major congenital malformations and/or genetic disorders
7. Plans to transfer to a non NRN site before discharge
8. Unable to obtain parental or guardian consent

4.4 Study Arm Assignment and Randomization

When the attending neonatologist has discontinued clinical caffeine on an eligible and consented infant that is $33^{0/7}$ - $35^{6/7}$ weeks PMA, the infant can be randomized if it is possible to start study medication within 72 hours after stopping caffeine.

Infants will be randomized 1:1 using a stratified randomly permutated block design with a centralized procedure. Infants will be stratified by study center into two GA groups; infants born at 29 to 30 6/7 weeks GA and 31 to 33 6/7 weeks GA. Randomization will occur through the Data Coordinating Center (DCC) of RTI International (Research Triangle Park, North Carolina). Treatment allocation will be revealed only to the research pharmacy staff, who will prepare either caffeine citrate at 10 mg/kg (base of 5 mg/kg) or an equivalent volume of placebo to be administered once a day. The dosage will be based on the weight closest to randomization and then adjusted weekly for weight. Because both caffeine and placebo are colorless, they are indistinguishable and, therefore, blinding will be accomplished for all care providers and research staff. The oral delivery should begin as soon as clinically possible.

Randomization of twins and higher order multiples (when eligible) will be done independently.

4.5 Masking and Data Lock

4.6 General Masking Procedures

This is a double-blind, placebo-controlled clinical trial. As such, neither site staff nor the parent/guardians that administer the study drug to the patient will be aware of the treatment assignment. Only the research pharmacists at the clinical sites and designated personnel at the DCC (Study Manager, Data Manager and Study Statistician) will be unblinded. All clinical caregivers and investigators, including the DCC PI will remain masked to the study drug assignment. Any intentional or unintentional unmasking will be reported as a protocol deviation in the data management system.

4.7 Database Lock

The database will be locked after all enrolled infants have completed their post-discharge follow-up, died or transferred, or reached 48 wks PMA. Randomization assignment will not be unmasked until after database lock is finalized. Data cleaning and edit check may continue after the database lock and final versions of the analysis files will be archived.

4.8 In-hospital and Post-hospital Procedures

Study subjects are infants in the NICU at one of the principal or associated sites of the NRN. Once the attending physician has discontinued caffeine therapy, the infant will be randomized as described in section 5.2.2. The research pharmacist will be notified of enrollment and the infant will receive a once-a-day dose of caffeine citrate at 10 mg/kg/dose (5 mg/kg caffeine base) as part of the treatment group or receive placebo at the same volume and time every 24 hours and not given less than 12 hours apart. The doses will be weight adjusted on a weekly basis. If the infant is later put NPO, the oral caffeine citrate or placebo will be held until feeding resumes and at the discretion of the attending neonatologist.

In-hospital apnea and bradycardia events will be monitored using individual study site monitor devices. There are no universal definitions for apnea and bradycardia and therefore, diagnoses of these will be left at the discretion of each hospital group of neonatologists (pre-specified) and the neonatal team caring for the infant. The data on apnea and bradycardia will be used to determine physiologic maturity and for the local decision about discharging the babies home. The definitions of apnea and bradycardia will be decided locally. However, for the purposes of this study, significant apnea/bradycardia will be defined as apnea/bradycardia events receiving treatment. Treatment is defined as ordered by the attending physician (e.g., re-starting caffeine, not just stimulation for example). Using local definitions will increase generalizability and interfere less with local decision making.

Data will be collected from the infant's electronic or physical chart. The number of infants who receive CPAP, mechanical ventilation, and/or open label caffeine, other methylxanthine or doxapram for the treatment of apnea/bradycardia after randomization will be noted for the study.

The protocol will allow for individual study site centers to use their own discharge criteria, but for the purposes of the definition of secondary outcome of the trial, infants must have reached physiologic maturity, which includes a) out of an incubator for at least 48 hours, b) oral feeding at volume of at least 140 ml/kg/day or growing on less than 140 ml/kg/day for at least 48 hours and be apnea-free for at least 5 consecutive days. Infants who are still in the hospital at 44 weeks PMA will have the study medication (caffeine/placebo) discontinued. Analysis will be done by intention to treat. If an infant is

discharged before meeting the definition of physiologic maturity, it will be noted for analysis at the conclusion of the trial. The weight within 48 hours of discharge will be used for the caffeine/placebo home administration. The dose will not be adjusted for weight gain after discharge. The practice of discharging infants to home on open-label caffeine will not be a part of the trial. However, should it occur, it will be noted as a protocol deviation and will be noted for subsequent analysis.

Prior to discharge, a research nurse, physician, or pharmacist will educate the parents on storage and administration of the medication to the infant. Other instructions to the parents include preparation of daily dose and if a dose is missed, it may be given at any time on that day if the dose can be given at least 12 hours before the next day's dose is due. If it was not given at all, the parents will be instructed to discard the entire vial that was missed and record it on the medication form. If an infant spits or vomits after a dose is given, another dose should not be given on the same day. Follow-up forms (Medication and Follow up Forms) will be supplied to allow the parents to keep track of daily medication administration and missed doses. The parent will be given a stamped envelope to return the forms to the NRN research coordinator at the end of the 8 weeks. If the study site allows for the use of a gift card, it will be mailed to the parent or guardian when the documents are returned to the research coordinator. The gift card will encourage the family member to return the follow up documents at the end of the 8-week follow up period. Because it will be supplied at the end of the study period, it will have no consequences to the enrollment or treatment of the infant. In addition, the gift card does not constitute payment to the family member for any service provided as part of the trial.

Of note, infants who are deemed to require apnea monitoring at home after enrollment will be allowed its use and this will be noted on a follow-up form. Data from the home monitor recordings will not be used in the trial.

The NRN research coordinator and staff will obtain maternal demographics data including age, race, education, marital status, and parity. In addition, infant data will also be recorded, including weight at birth, randomization, and status; length/head circumference at birth and status; date of status; date of last apnea episode recorded after randomization; use of CPAP or mechanical ventilation; date of last NPO episode lasting >24 hours as well as the number of episodes of NPO >24 hours following randomization; date of last orogastric or nasogastric feed; date of last day in incubator; dates of initiation and termination of open label caffeine therapy; and information related to side effects and adverse events.

Following discharge, the research team will contact the family within 72 hours after discharge, once a week for the first 4 weeks, and biweekly during the weeks 5 to 8 after discharge to ensure the study medication is being administered adequately and to gather information on doctor or hospital visits.

5 ANALYSIS POPULATIONS

Intention-to-Treat Population (ITT)

The intention-to-treat population will comprise of all randomized participants who have at least one efficacy measurement. Participants who receive study drug but do not complete the study will be used in all analyses for which data are available. Participants who are withdrawn for further follow-up will be treated as lost to follow-up at the time that consent was withdrawn, and data will be used up to the time of study withdrawal (unless parents/guardians expressly forbid use of any collected data).

Safety Population

The safety population will comprise of all randomized participants regardless of treatment received or collection of efficacy information.

6 SAMPLE SIZE DETERMINATION

To obtain a two-day difference in the primary outcome of days from randomization to discharge, from 14 days to 12 days, with a power set at 90%, a significance of 0.05, and a 5% attrition, the trial would require 587 infants per group (intervention and placebo) for a total of 1174 infants, assuming a two-tailed, non-parametric test comparing median LOS in the two groups. With a power set at 80%, under the same assumptions and significance, the trial would require 439 infants per group for a total of 878 infants. The minimum required sample size for this trial will be 878 (439 in each treatment group), but enrollment will continue until either the available supply of study drug is exhausted or expires in approximately August 2023 (accounting for full in-hospital and post-discharge follow-up), whichever comes first.

There are approximately 3000 infants born between $29^{0/7}$ and $33^{6/7}$ weeks gestation per year in the NRN. A review of 100 infants in the NICU at one of the largest centers in the NRN showed that 37% of infants within this cohort were treated with a methylxanthine. This is similar to NRN data showing that approximately 1000 (one-third) of infants receive caffeine therapy. The following table shows scenarios for approximate time to completion of the trial as a function of consent rate and NICHD center involvement:

Consent Rate (%)	Time to Completion (mo), *all* sites	Time to Completion (mo), *half* sites
50	24	48
60	20	40
70	18	36
80	15	30

Effect Size (days)	Significance	Power (%)	Attrition (%)	Sample Size
14 to 10	0.05	80	5	226
14 to 10	0.05	90	5	302
14 to 12	0.05	80	5	878
14 to 12	0.05	90	5	1174

7 STATISTICAL/ANALYTICAL ISSUES

7.1 General Rules

All statistical computations performed and data summaries created will be accomplished using SAS 9.4 or higher. If additional statistical packages are required, these will be discussed in the study report. 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.

7.2 Adjustments for Covariates

Indicator variables for the study stratification of site and gestational age will be included as covariates in most efficacy analyses performed for this study. Additionally, demographic and baseline characteristics for subjects will be compared between study arms using analysis of covariance techniques for continuous measures, Mantel-Haenszel mean score test using standardized mid-rank scores for ordinal measures, and Cochran Mantel-Haenszel chi-square tests for general association for categorical measures. If sample sizes allow, these analyses will control for the study stratification factors. If these analyses suggest that substantial differences exist among arms, the use as covariates of these parameters on which the arms differ will be explored in exploratory analyses of the efficacy data.

7.3 Handling of Dropouts and Missing Data

Missing data will be monitored in real-time. In the event an assessment or data form is missed, DCC staff will contact site coordinators to enter the required information from the in-hospital portion of the study as quickly as possible. For the post-hospital data collection, missing data will be queried by DCC staff to determine if the site has collected the desired information from the parents/guardians. Given that recruitment into the study and the primary outcome are all performed/collected during the initial hospital stay, it is not anticipated that there will be any missing data for the length-of-stay primary outcome.

7.4 Interim Analyses and Data Monitoring

Statistical interim monitoring for this trial will be conducted by the independent NRN Data Safety and Monitoring Committee (DSMC) using data and analyses provided by the NRN DCC at RTI. We recommend that the DSMC meet formally to review interim safety, efficacy, and futility, as well as other aspects of accruing data (such as enrollment, data completeness and protocol compliance) four times during the study, after 100 infants complete the eight-week follow up and then roughly after 25%, 50% and 75% of subjects complete the eight-week follow up. Formal DSMC reports will be prepared by the DCC for these meetings. These reports will include the following information:

- Brief summary of the trial design, including primary and secondary hypotheses and outcomes, study population, inclusion and exclusion criteria, recruitment, screening, randomization and study intervention procedures, and statistical considerations for trial design and analysis
- Interim monitoring plan
- Enrollment, including screening, consent, randomization, and study exit
- Completeness of data and edit queries. Baseline study population characteristics, overall and by treatment group
- Drug adherence and compliance
- Primary efficacy outcomes by treatment group
- Secondary efficacy outcomes by treatment group
- Safety outcomes, including death, by treatment group
- Protocol deviations and violations, overall and by treatment group.

These DSMC reports will be blinded, with treatments labeled as group A and group B. As per the NRN DSMC Charter, the DSMC may, however, request to be un-blinded to perform their duties. If the DSMC recommends modification or cessation of the study protocol due to safety concerns, NICHD will make the final determination.

Interim Safety Monitoring

Statistical interim safety monitoring will be conducted by the DCC after 100 infants complete the eight-week follow up and then after 25% (300), 50% (600), and 75% (900) of the infants enrolled complete the eight-week follow up period. Formal safety assessment will be based on the number of sick visits, emergency room visits, apparent life-threatening event episodes or hospitalizations related to AoP. Relatively liberal Pocock safety bounds based on 5 interim safety looks at the data will be constructed to guard against any occurrence of false positives while at the same time allowing for stopping at reasonable levels of evidence. Thus, at each of these safety looks, an increased incidence of the above safety events in the caffeine group with $p < 0.0158$ will be considered as statistically significant evidence of harm that the DSMC can use to recommend suspension of the trial for safety reasons. In addition to the formal safety outcome, other safety outcomes collected in this trial (see section 6.3.1) will also be examined by treatment group. If a death occurs in either of the groups, the DSMC will review the cause of death and determine if the cause is linked to the study intervention. If it is deemed to be related to the intervention as an increased risk, the DSMC may recommend stopping the trial. After the DCC conducts each of these interim safety analyses, it will inform the Chair of the DSMC (or her designee) if there is any suggestion of a trend in the safety data. The DSMC Chair will then determine any future course of action, including convening a meeting of the entire panel to discuss the safety data.

Interim Efficacy Monitoring

Statistical interim efficacy monitoring will be conducted by the DCC after 25%, 50% and 75% of the babies enrolled into the trial complete eight-week follow up period. At each of these interim looks, the primary outcome will be compared by (blinded) treatment group using the methods outlined in Section 9. O'Brien Fleming stopping bounds for efficacy will be used to preserve an overall Type-I error rate of 0.05 for the primary outcome. Thus, stopping bounds for efficacy will be established if $p < 5 \times 10^{-5}$ at 25%, $p < 0.004$ at 50% and $p < 0.018$ at 75% of interim data accrual. In the event that these stopping bounds are reached, the DSMC may recommend termination of the trial for efficacy. In addition to the primary outcome, secondary efficacy outcomes collected in this trial will also be examined by treatment group.

Interim Futility Monitoring

Statistical interim futility monitoring will be conducted by the DCC after 50% and 75% of the babies enrolled into the trial complete the eight-week follow up period. At each of these two looks, we will estimate conditional power, i.e., the probability to detect a statistically significant treatment effect for caffeine, given the observed data and the hypothesized treatment effect for the unobserved data. The DSMC may recommend suspension of the trial for futility if this probability is less than 15%. In addition to the conditional power analysis, the DSMC may also consider other pertinent aspects included in the interim report such as the quality of the data, protocol violations and treatment adherence, rate of enrollment, rate of attrition, in determining futility.

Additional Data Monitoring

Additional monitoring of data for this study will occur in real-time and will include study drug and home kit reconciliation with randomization, pharmacy log reconciliation with patient administration log, missing and incomplete forms, tracking of primary and secondary outcomes, and reporting of protocol deviations and adverse events.

Expedited serious adverse events will be reported per local IRB guidelines, and within 24 hours to the NRN DCC at RTI.

The trial will record side effects and any study related serious adverse events will be reported to the DSMC periodically. Expedited serious adverse events will be reported per local IRB guidelines, and within 24 hours to the NRN DCC at RTI. All protocol deviations and violations will be reported.

Infants will be monitored prospectively while in the NICU. Interim efficacy and safety will be formally evaluated by the independent DSMC.

7.5 Masked Data Review

A masked data review of the primary outcome and secondary outcomes for this study will be performed by the protocol team. This review will occur prior to completion of the primary analyses prior to any official unblinding. This will include a presentation of descriptive statistics (e.g., means, standard deviations, percentiles for continuous variables and counts and percentages of categorical variables) of the selected outcomes and model predictor variables.

7.6 Multicenter Studies

For this multicenter study, randomization of study participants is stratified within center and by gestational age. Consequently, for all model-based primary and secondary analyses, center will be included as a fixed effect in the models. As an ancillary analysis associated with the primary outcome, we will examine descriptively whether the treatment effect varies across sites; however, no other analyses will assess site differences in treatment effect because sample sizes are inadequate to support evaluation of site-level effects.

7.7 Multiple Comparisons and Multiplicity

The primary hypothesis will be tested at a nominal two-sided type I error of 0.05. All p-values for any baseline and demographic characteristic comparisons, secondary outcomes, and safety parameters will be for descriptive purposes only.

7.8 Examination of Subgroups

Treatment heterogeneity by hospital center, gestational age, and sex will be examined as part of the primary analysis. If any interaction effects are determined to be significant, subgroup analysis will be performed to assess the treatment effects in the appropriate subpopulations. See Section 9.3 for more details.

8 STUDY PARTICIPANT CHARACTERIZATION

8.1 Participant Disposition

Participant eligibility status will be summarized, and overall disposition of study participants will be described using a standard CONSORT diagram. The number of participants randomized and those completing or discontinuing from study therapy will be summarized. Reasons for study withdrawal will be listed.

8.2 Protocol Deviations

Protocol deviations will be listed with information such as type of deviation, time of occurrence, and reason. Depending on the number of deviations reported, the number of protocol deviations will also be summarized overall and for each protocol deviation category.

8.3 Study Drug Exposure and Compliance

Adherence to medication regimens is affected in a multi-factorial manner. A review of 51 studies with 19 chronic diseases found a wide range of factors for lack of compliance, falling within categories of socioeconomic, patient-related, and therapy-related determinants[40]. The conclusion of this large systematic review points to the use of multiple methods to better attain compliance with medications at home.

For this trial, we will combine two recommended methods to assure adequate monitoring and improve compliance with medication administration:

a. While the infant remains inpatient, the medication will be administered by the nurse as directed and the research nurse will perform a chart review at least twice a week to ensure study medication is being received after randomization and prior to discharge.
b. After discharge:
    i. A Medication form will be supplied to the parents. This will include a daily item to mark if medication was administered. In addition, parents can use this form to record sick visits to the Health Care Provider, emergency room, or hospitalizations as well as weights obtained during well-child or sick visits and missed doses of study medication. Any new medications prescribed will also be noted on this form.
    ii. A Follow-up form will be supplied. This will be similar to the Medication form but will be used during the second four-week period discharge and will not include a medication administration section. Safety outcomes (such as Health Care Provider, Emergency visits, and re-admissions to a hospital) will be recorded on this form also.
    iii. A member of the research team will contact the parents within the first 72 hours after discharge to answer questions and assess medication compliance.
    iv. A member of the research team will contact the parents once a week during the first 4 weeks and biweekly during weeks 5-8 to address concerns and evaluate medication compliance*. During this contact, the parents will be asked about any well or sick visits made to the health care provider, urgent care, emergency room, or hospital since the last contact and the reason for these visits.

*To ensure both safety and compliance, vials will be marked with a number corresponding to the beginning day of medication to be administered at home. For example, the next day after discharge, the parent is to administer the study drug from the vial marked #1 by oral syringe. The following day, the vial marked #2 should be given. This procedure, using an oral syringe, should be followed until day 28 after discharge when the last medication vial marked #28 should be given. If a dose is missed, it may be given at any time on that day as long as it is given at least 12 hours before the next dose is due the next day. If it was not given at all, the parents will be instructed to discard the dose that was missed and record it on the medication form. The schedule of medication administration will then continue with the next numbered vial. This information will be noted for further analysis at the conclusion of the trial. Because of the long half-life of caffeine, single missed doses will have little clinical effect.

8.4 Demographic and Baseline Characteristics

Demographic and baseline clinical characteristics for the study participants will be summarized by study arm using the general analysis rules describe above. Variables of interest include mother demographics, gender, gestational age, race and ethnicity, and in-hospital clinical outcomes at the time of birth.

9 EFFICACY ANALYSES

9.1 Overview of Efficacy Analysis Methods

- All efficacy analyses will be performed on the ITT population.
- All efficacy variables will be summarized by treatment group. Count (N), mean, standard deviation, minimum, and maximum will summarize continuous efficacy variables (e.g., post-menstrual age); median, 1st quartile, 3rd quartile, minimum, and maximum will be calculated for count efficacy variables (e.g., number of days of hospitalization); whereas number and percent will summarize categorical efficacy variables (e.g., death).
- Analyses of the primary outcome are separately discussed in Section 9.3. Analyses of other dichotomous outcomes will be performed using robust Poisson regression models, with both log and linear link functions, to estimate adjusted relative risks and risk differences. Efficacy models will be adjusted for stratification by clinical site and gestational age. Unless otherwise noted, all analyses of continuous efficacy variables will be performed using median or robust Poisson regression.

  Models will be adjusted for clinical site and gestational age. If there are not enough patients per clinical site to include the variables in the models as fixed effects, clinical site will be included as a random intercept to account for correlation between outcomes of patients treated by the same clinical site or similar sites will be combined.

9.2 Efficacy Variables

Primary and secondary efficacy variables are described in the table below.

Variable	Type	Definition
Primary Outcome		
Number of days of hospitalization	Count	The number of days between randomization and hospital discharge. This outcome is censored at 48 weeks PMA and at time of transfer or death
Secondary Outcomes		
Number of days to physiologic maturity	Count	Physiologic maturity is defined: i. Temperature: out of the incubator for at least 48 hours with normal body temperature ii. Feeding: oral feeding at a volume of at least 140 ml/kg for 48 hours or growing on less than 140 ml/kg/day for at least 48 hours iii. Respiratory: apnea-free for at least 5 consecutive days This outcome is censored at 48 weeks PMA.

Variable	Type	Definition
Number of days to physiologic maturity - Temperature	Count	Physiologic maturity is defined: i. Temperature: out of the incubator for at least 48 hours with normal body temperature This outcome is censored at 48 weeks PMA.
Number of days to physiologic maturity – Feeding	Count	Physiologic maturity is defined: i. Feeding: oral feeding at a volume of at least 140 ml/kg for 48 hours or growing on less than 140 ml/kg/day for at least 48 hours This outcome is censored at 48 weeks PMA.
Number of days to physiologic maturity - Respiratory	Count	Physiologic maturity is defined: i. Respiratory: apnea-free for at least 5 consecutive days This outcome is censored at 48 weeks PMA.
Post-menstrual age at discharge	Continuous	The post-menstrual age of the infant at discharge censored at 48 weeks PMA and at time of transfer or death
Number of all-cause readmissions within first 4 weeks post-discharge	Count	Number of hospital readmissions including apneic events and apparent life-threatening events during first 4 weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
Number of all-cause readmissions within second 4 weeks post-discharge	Count	Number of hospital readmissions including apneic events and apparent life-threatening events during second 4 weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
Number of all-cause readmissions within first 8 weeks post-discharge	Count	Number of hospital readmissions including apneic events and apparent life-threatening events during first 8 weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
Number of all-cause sick visits within first 4 weeks post-discharge	Count	Number of all cause sick visits, urgent care, emergency room, or other non-scheduled/routine health care provider office visits during first 4 weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
Number of all-cause sick visits within second 4 weeks post-discharge	Count	Number of all cause sick visits, urgent care, emergency room, or other non-scheduled/routine health care provider office visits during last 4

Variable	Type	Definition
		weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
Number of all-cause sick visits within first 8 weeks post-discharge	Count	Number of all cause sick visits, urgent care, emergency room, or other non-scheduled/routine health care provider office visits during first 8 weeks following hospital discharge. The outcome is censored at 48 weeks PMA.
All-cause Mortality	Categorical	Demise of Infant

9.3 Primary Analysis Methods

For the primary analysis, the number of days of hospitalization between randomization and discharge will be compared between the caffeine citrate and placebo groups using either median or robust Poisson regression models. The model will include a fixed effect for treatment group. The model will also be adjusted for the design effect of stratification by site and gestational age. Estimates, p-values and 95% confidence intervals will be presented for treatment group comparisons. In an intent-to-treat fashion, all data will be included in the primary model based on their randomized treatment, regardless of the treatment administered to the patient. If there are important baseline characteristics that differ between randomized treatment groups by chance, additional sensitivity analyses may be conducted to evaluate the impact of those imbalances.

We will examine treatment heterogeneity for the primary outcome by adding interaction terms between treatment and study center, as well as treatment and gestational age group, and following NIH guidelines, treatment and sex. If any of these interaction terms are close to statistical significance ($p<0.1$), for treatment heterogeneity by study center, we will attempt to identify centers where the treatment effect may be different (although the statistical power available to conduct such analyses will be fairly limited). If there is indication of treatment heterogeneity by GA group and/or sex, we will conduct stratified analyses by GA group and/or sex to quantify the treatment effect separately by GA group/sex.

9.4 Secondary Analysis Methods

Secondary outcomes will be compared between treatment groups using models similar to the primary outcome for continuous measures and analogous generalized linear models for categorical outcomes.

Additionally, secondary analyses will evaluate whether other models better fit the distributions for both the primary and secondary outcomes (for example, survival analysis models for the primary outcome and negative binomial or log linear models for suitable secondary outcomes).

9.5 Exploratory Analysis Methods

Not Applicable.

10 SAFETY ANALYSES

10.1 Overview of Safety Analysis Methods

All safety analyses will be performed using all participants who were randomized and evaluated based on the treatment they received. The frequency of toxicities potentially attributable to study treatment will be determined. Treatment and resolution of all safety endpoints will also be documented. Treatment groups will be compared for incidence of events with a chi-squared or Fisher's exact test, and incidence density (events per person-months of exposure) will be calculated.

10.2 Safety Variables

Safety outcome variables are described in the table below.

Variable	**Type**	**Definition**
Safety Outcomes		
Weight gain from randomization until status	Continuous/ Categorical	Change in weight between randomization and in-hospital status. Outcome is censored at 48 weeks PMA and time of transfer or death. May also be assessed using percentiles.
Number of days after randomization until status that infant had at least two consecutive heart rates >200 documented at least 3 hours apart	Count	Number of days between randomization and status with at least 2 consecutive heart rates > 200 bpm at least 3 hours apart. Outcome excludes when infant was crying.
Treatment for high blood pressure initiated after randomization until status	Categorical	Presence of treatment for high blood pressure that was started between randomization and status.
Number of episodes between randomization and status that infant was placed NPO for ≥ 24 hours	Count	Number of days between randomization and status in which infant was placed NPO for at least 24 hours
Use of anti-reflux medications between randomization and status	Categorical	Presence of use of anti-reflux medication that were started between randomization and status.
Number of days of significant apnea/bradycardia between randomization and status	Count	Number of days between randomization and status with infant showing significant apnea/bradycardia. Significant apnea/bradycardia documented by infant receiving open label caffeine, doxapram, other methylxanthines, CPAP or ventilatory support for apnea/bradycardia

Variable	Type	Definition
Presence of documented and treated arrhythmias between randomization and status	Categorical	Presence of treated arrhythmias between randomization and status. Outcome excludes arrhythmias due to tachycardia or bradycardia
Onset of documented seizures between randomization and status	Categorical	Presence of documented seizures between randomization and status. Documented seizures are defined by treating with anti-convulsants.

10.3 Adverse Events

Per the protocol, specific adverse events will be collected on an adverse event log. Adverse events will be listed and summarized by event type. Summaries will be of the number of events and number of individuals experiencing events by treatment group and will be created for all adverse events, by severity, and by relationship to treatment.

10.4 Deaths and Serious Adverse Events

A serious adverse event (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 and treatment-related SAEs will be summarized in the manner mentioned in Section 10.3 if there are enough events to summarize. Deaths will be listed.

11 ANALYSIS OF OTHER OUTCOMES

At this time, there are no other outcomes under consideration. Additional analyses motivated by observations in the final outcome data may be conducted, but these will be clearly identified as post-hoc analyses in any publication and will not be specifically addressed via updates to the SAP.

12 REPORTING CONVENTIONS

Unless required otherwise by FDA standards, 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.

13 CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

None at this time. Additional analyses motivated by observations in the final outcome data may be conducted, but these will be clearly identified as post-hoc analyses in any publication and will not be specifically addressed via updates to the SAP.

14 REFERENCES

None.

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
Demographic and Baseline Characteristics (Demographic and baseline summaries mothers and infants)
Participant Disposition (Number of individuals consented, randomized, and completing in-hospital and post-discharge, as well as final disposition based on final study status.)
Dosing Summary (Number of doses taken in hospital, proportion of expected doses taken post-discharge missed)
Primary Efficacy Model Results
Secondary Efficacy Model Results
Exploratory Efficacy Model Results
Safety Summary
Figures
CONSORT Diagram
Boxplots of numbers of days
Survival plot for primary outcome by treatment arm
Effect estimates and 95% confidence intervals summarizing treatment heterogeneity for primary outcome by GA group and sex
Additional figures may be produced for manuscripts.
Data Listings will be created for all critical study variables