STATISTICAL ANALYSIS PLAN

PRAGMATIC, RANDOMIZED, BLINDED TRIAL TO SHORTEN PHARMACOLOGIC TREATMENT OF NEWBORNS WITH NEONATAL OPIOID WITHDRAWAL SYNDROME (NOWS)

Short Title: NOWS Weaning

SAP VERSION: Version 2.0

SAP DATE: 02/26/2024

PI/SPONSORS: Jessica Snowden, University of Arkansas for Medical Sciences
Abhik Das, RTI International

PREPARED BY: RTI International

AUTHOR (S): Rouba Chahine

Contents

1 ADMINISTRATIVE INFORMATION

1.1 Trial Name

1.1.1 Trial Registration

ClinicalTrials.gov identifier: NCT04214834

NRN ID number: NICHD-ACT-NOW-0003

1.2 Protocol Version

Version V-10 (2023-05-26)

1.3 SAP Revision History

Version	Justification for Change	Date
1.0	Initial Version	2024-01-31
2.0	Version 1.0 shared with the larger group and based on feedback the following changes were made after 02/22/2024 meeting: 1. Added section 6.6.2 regarding handling missing follow-up data. 2. Removed NDI terminology from Bayley's descriptive tables.	2024-02-26

1.4 List of Abbreviations

Abbreviations	Definition:
AE	Adverse event
DCC	Data Coordinating Center
DSMB	Data and Safety monitoring board
FAS	Full analysis set
GA	Gestational age
hr	Hours
SC	Steering Committee
IRB	Institutional Review Board
ITT	Intent-to-treat
mITT	Modified intent-to-treat
N/A	Not Applicable
NAS	Neonatal abstinence syndrome
NDI	Neurodevelopmental Impairment
NICHD	National Institute of Child Health and Human Development
NICU	Newborn intensive care unit
NIH	National Institute of Health
NNNS	NICU Network Neurobehavioral Scale
NOWS	Neonatal opioid withdrawal syndrome
NRN DCC	Neonatal Research Network Data Coordinating Center
PP	Per protocol
RS	Research site
SAE	Severe adverse event
SAP	Statistical analysis plan
wks	Weeks

2 Introduction

The NOWS Weaning trial is a pragmatic, randomized, multi-center, masked trial to compare effects of a rapid-wean oral opioid intervention to a slow-wean oral opioid intervention as the primary treatment for neonatal opioid withdrawal syndrome (NOWS). The primary outcome is the number of days of opioid treatment (used as primary treatment), including escalation, resumption, and spot treatment, from first weaning dose to cessation of opioid.

2.1 Background and Rationale

The incidence of maternal opioid use in the United States has increased substantially since 2000. This includes an increase of opioid use during pregnancy including prescription opioids and illicit drugs, as well as a rise in opioid substitution programs for addiction treatment. As a consequence of opioid use during pregnancy, the incidence of neonatal opioid withdrawal syndrome (NOWS) has increased five-fold between 2002 and 2012. NOWS is a clinical syndrome that reflects signs of withdrawal from opioids in a newborn following *in-utero* exposure. Signs typically occur in the first 5-7 days following birth and reflect dysfunction of the brain, gastrointestinal tract, and autonomic regulation. Simultaneously during this rise in opioid use, the pattern of use has shifted from an inner city, indigent population to a more socioeconomically diverse population.

A systematic literature review indicated rural pregnant women have higher rates of polysubstance abuse, as compared to urban pregnant women. Pregnancy complicated by opioid use disorder is associated with high rates of polydrug use, mental health disorders, infectious diseases, poor nutrition, chronic illnesses, and limited social support. Associated risks for newborns beyond NOWS include preterm birth and fetal growth restriction. Pregnancy represents an opportunity for entry into the healthcare system and initiation of interventions for the mother-infant dyad. However, there are many knowledge gaps in the care of infants with NOWS.

At the time of protocol development, a *Journal of Pediatrics* editorial emphasized the rapid rise of NOWS in the United States and provided a framework to target research initiatives and care delivery innovations for infants with NOWS. Specifically, research and quality improvement initiatives should be safe, effective, patient centered, timely, efficient, and equitable. High-quality research is needed to ensure that NOWS care is evidence-based, eliminates non-beneficial practices, and achieves the overarching goals of limiting ongoing opioid exposure for infants, minimizing separation of the mother-infant dyad, and reducing healthcare expenditures. To date, the research community has not rigorously evaluated, through randomized clinical trials, many aspects of NOWS treatment regimens.

2.2 Study Objectives

Among infants receiving an opioid (defined as morphine or methadone) as the primary treatment for neonatal opioid withdrawal syndrome (NOWS), a rapid-wean intervention will reduce the number of days on opioid treatment from the first weaning dose to cessation of opioid compared to a slow-wean intervention.

2.2.1 Study Hypotheses

Primary objective

1) **Research Objective 1**: To compare the numbers of days of opioid treatment from the first weaning dose to cessation between rapid- and slow-wean regardless of medication use (morphine or methadone). Specifically, we expect that the Number of days of opioid treatment from the first weaning dose to cessation will be lower in rapid-wean than in slow-wean treatment regardless of medication used.

   H_0: Number of days of opioid treatment from the first weaning dose to cessation will not differ between rapid- and slow-wean treatment regardless of medication used.

   H_a: Number of days of opioid treatment from the first weaning dose to cessation will differ between rapid- and slow-wean treatment regardless of medication used.

Main secondary objectives

2) **Research Objective 2**: To compare the numbers of days of opioid treatment from the first weaning dose to cessation of opioid between rapid- and slow-wean interventions among infants treated with morphine. Specifically, we expect that the number of days of opioid treatment from the first weaning dose to cessation will be lower in rapid-wean than in slow-wean treatment among infants treated with morphine.

   H_0: Number of days of opioid treatment from the first weaning dose to cessation will not differ between rapid- and slow-wean treatment among infants treated with morphine.

   H_a: Number of days of opioid treatment from the first weaning dose to cessation will differ between rapid- and slow-wean treatment among infants treated with morphine.

3) **Research Objective 3**: To compare the numbers of days of opioid treatment from the first weaning dose to cessation of opioid between rapid- and slow-wean intervention among infants treated with methadone. Specifically, we expect that the number of days of opioid treatment from the first weaning dose to cessation will be lower in rapid-wean than in slow-wean treatment among infants treated with methadone.

   H_0: Number of days of opioid treatment from the first weaning dose to cessation will not differ between rapid- and slow-wean treatment among infants treated with methadone.

   H_a: Number of days of opioid treatment from the first weaning dose to cessation will differ between rapid- and slow-wean treatment among infants treated with methadone.

4) **Research Objective 4**: To compare the proportion of infants in the rapid- and slow-wean intervention arms who have an escalation or resumption of opioid during weaning.

   H_0: The rate of escalation or resumption of opioid during weaning will not differ between rapid- and slow-wean treatment regardless of medication used.

H_a: The rate of escalation or resumption of opioid during weaning will differ between rapid- and slow-wean treatment regardless of medication used.

5) **Research Objective 5**: To compare the total amounts of opioid from the first weaning dose to cessation of opioid among infants in rapid- and slow-wean intervention arms.
Statistical Hypothesis 4: The total amounts of opioid used from the first weaning dose to cessation of opioid will be lower in rapid-wean group that in slow-wean group regardless of medication used.

   H_0: Total amounts of opioid used from the first weaning dose to cessation will not differ between rapid- and slow-wean treatment regardless of medication used.

   H_a: Total amounts of opioid used from the first weaning dose to cessation will differ between rapid- and slow-wean treatment regardless of medication used.

6) **Research Objective 6**: To compare the proportion of infants who experience initiation and/or escalation of second-or-third line drugs to treat NOWS signs from the first weaning dose to cessation of opioid in the rapid- and slow-wean intervention arms.

   H_0: Proportion of infants who experience initiation and/or escalation of second-or-third line drugs to treat NOWS signs from the first weaning dose to cessation will not differ between rapid- and slow-wean treatment regardless of medication used.

   H_a: Proportion of infants who experience initiation and/or escalation of second-or-third line drugs to treat NOWS signs from the first weaning dose to cessation will differ between rapid- and slow-wean treatment regardless of medication used.

   Additionally, for this objective a descriptive table with second-or-third line drugs to treat NOWS signs initiated before start of study drug will be provided.

Safety secondary outcome, Table 2

7) **Research Objective 7:** To compare the proportion of infants in each intervention arm with any safety outcome (seizure, excessive stool, respiratory disturbance, and/or feeding intolerance).
   A) **Objective 7.a:** To compare the proportion of infants in each intervention arm with seizures (clinical or EEG)
   B) **Objective 7.b**: To compare the proportion of infants in each intervention arm with excessive stool output
   C) **Objective 7.c**: To compare the proportion of infants in each intervention arm with respiratory disturbances
   D) **Objective 7.d**: To compare the proportion of infants in each intervention arm with feeding intolerance

Additional secondary objectives

8) **Research Objective 8**: To compare infants in each treatment arm with an atypical neurobehavioral profile prior to discharge on the NICU Network Neurobehavioral Scale (NNNS).

   H_0: Behavioral profiles will not differ between rapid- and slow-wean treatment regardless of medication used.
   H_a: Behavioral profiles will differ between rapid- and slow-wean treatment regardless of medication used.

9) **Research Objective 9**: To compare the lengths of hospital-stay between the treatment intervention arms. Specifically, we expect that the length of hospital stay will be lower in the rapid-wean group then in the slow-wean group regardless of medication use.

   H_0: The length of hospital stay will not differ between rapid- and slow-wean treatment regardless of medication used.

   H_a: The length of hospital stay will differ between rapid- and slow-wean treatment regardless of medication used.

10) **Research Objective 10**: Compare the following assessments between rapid- and slow-wean groups:
    - Statistical Hypothesis 10A: To compare maternal well-being using PROMIS score at 24-month between rapid- and slow-wean groups regardless of medication used adjusting for PROMIS score at one month.

      H_0: PROMIS score at 24-month will not differ between rapid- and slow-wean treatment regardless of medication used.

      H_a: PROMIS score at 24-month will differ between rapid- and slow-wean treatment regardless of medication used.

    - Statistical Hypothesis 10B: To compare maternal-infant attachment measured using MPAQ total scores at one month post discharge

      H_0: MPAQ total scores will not differ between rapid- and slow-wean treatment regardless of medication used.

      H_a: MPAQ total scores will differ between rapid- and slow-wean treatment regardless of medication used.

11) **Research Objective 11**: To compare the growth over the first 24 months of age in each intervention arm.
    - Statistical Hypothesis 11A: To compare the weight z-score between rapid- and slow-wean groups regardless of medication used.

      H_0: Infants z-scores for weights will not differ between rapid- and slow-wean treatment regardless of medication used.

H_a: Infants z-scores for weights will differ between rapid- and slow-wean treatment regardless of medication used.

- Statistical Hypothesis 11B: To compare the height z-score between rapid- and slow-wean groups regardless of medication used.

  H_0: Infants z-scores for heights will not differ between rapid- and slow-wean treatment regardless of medication used.

  H_a: Infants z-scores for heights will differ between rapid- and slow-wean treatment regardless of medication used.

- Statistical Hypothesis 11C: To compare the head circumference z-score between rapid- and slow-wean groups regardless of medication used.

  H_0: Infants z-scores for head circumference will not differ between rapid- and slow-wean treatment regardless of medication used.

  H_a: Infants z-scores for head circumference will differ between rapid- and slow-wean treatment regardless of medication used.

12) **Research Objective 12**: To assess infant wellness after discharge and initial 24 months of age in each intervention arm, by looking at ER visits, acute care facility visits over time, and at the BITSEA scores.

- Statistical Hypothesis 12a: To compare the number of ER and acute care facility visits between rapid- and slow- wean groups regardless of medication used.

  H_0: The rate of ER and acute care facility visits will not differ over time between rapid- and slow-wean groups regardless of medication used

  H_a: The rate of ER and acute care facility visits will differ over time between rapid- and slow-wean regardless of medication used

- Statistical Hypothesis 12b: To compare the number of ER and acute care facility visits due to NOWS over time between rapid- and slow- wean groups regardless of medication used.

  H_0: The rate of ER and acute care facility visits due to NOWS will not differ over time between rapid- and slow-wean groups regardless of medication used.

  H_a: The rate of ER and acute care facility visits due to NOWS will differ over time between rapid- and slow-wean groups regardless of medication used.

- Statistical Hypothesis 12c: To compare the number of hospitalizations after an ER and/or acute care facility visits over time between rapid- and slow- wean groups regardless of medication used.

  H_0: The rate of hospitalization after an ER and/or acute care facility visits will not differ over time between rapid- and slow-wean groups regardless of medication used.

H_a: The rate of hospitalization after an ER and/or acute care facility visits will differ over time between rapid- and slow-wean groups regardless of medication used.

- Statistical Hypothesis 12d: To compare BISTEA percentile scores administered during the 24-month visits between rapid- and slow-wean groups regardless of medication used.

  H_0: The percentile ranking for BITSEA will not differ between rapid- and slow-wean groups regardless of medication used.

  H_a: The percentile ranking of BITSEA will differ between rapid- and slow-wean groups regardless of medication used.

For these objectives descriptive table for reason of ER visits will be provided.

13) **Research Objective 13**: To assess infant development at 24 months of age in each intervention arm.

- Statistical Hypothesis 13 a. To compare Bayley cognitive scores between rapid- and slow-wean groups regardless of medication used.

  H_0: Bayley cognitive score will not differ between rapid- and slow-wean groups regardless of medication used.

  H_a: Bayley cognitive score differ between rapid- and slow-wean groups regardless of medication used.

- Statistical Hypothesis 13 b. To compare Bayley language scores between rapid- and slow-wean groups regardless of medication used.

  H_0: Bayley language scores will not differ between rapid- and slow-wean regardless of medication used.

  H_a: Bayley language scores differ between rapid- and slow-wean groups regardless of medication used.

- Statistical Hypothesis 13 c. To compare Bayley motor score between rapid- and slow-wean groups regardless of medication used.

  H_0: Bayley motor score will not differ between rapid- and slow-wean groups regardless of medication used.

  H_a: Bayley motor score differ between rapid- and slow-wean regardless of medication used.

- Additionally, for this objective a descriptive table will be included using the following cutoffs for impairment.

Domain	Normal	Moderate Impairment	Severe/Profound Impairment
Bayley III Cognitive or Bayley 4 Cognitive	≥ 85	70-84	<70 / ≤ 54

Bayley III Motor or Bayley 4 Motor	≥ 85	70-84	<70 / ≤ 46

And the following cutoffs for language delay

Domain	Normal	Moderate Delay	Severe/Profound Delay
Bayley III Language or Bayley 4 Language	≥ 85	70-84	<70 / ≤ 54

3 Study Methods

This statistical analysis plan (SAP) contains detailed information about statistical analysis to be performed to assess the above-mentioned primary and secondary objectives once the final NOWS Weaning analysis datasets are constructed.

3.1 Trial design

A pragmatic, randomized, blinded, multicenter trial to compare a rapid-wean intervention (15% decrements from the stabilization dose) to a slow-wean intervention (10% decrements from the stabilization dose) for infants with NOWS. Randomization is stratified by hospital, with the study protocol commencing after NOWS signs have been controlled with an opioid. Weaning pharmacology treatments are either morphine or methadone administered orally.

3.2 Study Intervention and Process

Participating hospitals must provide opioid replacement therapy with either morphine or methadone as the primary drug for treating NOWS. Hospitals may change use of these two opioids during the trial period. The study protocol will commence after NOWS signs have been controlled with an opioid (stabilization) and weaning of pharmacologic treatment is to be started. At or before each 24-hour interval, clinical team members will evaluate and score infants, per hospital practice, for signs of NOWS to determine if the infant will tolerate weaning of the study drug.

- If the infant can tolerate weaning and is in the rapid-wean intervention arm, the clinical team will reduce the study drug by 15% of the stabilization dose. The clinical team will terminate the study drug when the infant can tolerate 25% of the stabilization dose without NOWS signs.

- If the infant can tolerate weaning and is in the slow-wean intervention arm, the clinical team will reduce the study drug by 10% of the stabilization dose. The clinical team will terminate the study drug when the infant can tolerate 20% of the stabilization dose without NOWS signs.

- If infants cannot tolerate weaning in either intervention arm, infants will enter a 12-hour period of study protocol guideline that will mandate either weaning or escalating the study drug by the end of the 12-hour interval. If the clinical team escalates the study drug, infants will receive opioid using the prior step of the assigned intervention arm.

To maintain blinding of study drug dose during the interventions, the volume of the syringe will be constant and equal the volume of the opioid at stabilization. As the clinical team decreases the study drug during the interventions, the pharmacist will add normal saline to keep a constant syringe volume. Only the pharmacy will be aware of the opioid dose. The use of placebo (normal saline without opioid) in the rapid-wean intervention arm will ensure comparable duration of both weaning interventions.

As part of a pragmatic trial, clinical teams will follow hospital practice for other care practices related to NOWS treatment (type of scoring system, threshold to initiate treatment, duration of stabilization, use of second line and third line-drugs, rooming in, breast milk, etc.). After study drug cessation, the clinical team will observe infants in the hospital for at least 48 hours prior to discharge, which is similar to clinical practice. A trained examiner will administer the Neonatal Intensive Care Unit (NICU) Network Neurobehavioral Scale (NNNS) to assess neurobehavioral profiles after infants cease study drug and prior to discharge.

At one month post discharge, primary caregivers will complete the Parent-Reported Outcome Measure Information System (PROMIS) Measures, the Maternal Postnatal Attachment Questionnaire (MPAQ) and a caregiver questionnaire. The site research team will contact the primary caregiver(s) to update contact information and/or complete questionnaires when the infant is 6 months, 12 months, 18 months, and 24 months of age. The questionnaires will assess infant wellness, visits to emergency room facilities, hospitalizations, and caregiver well-being. At 24 months, the infants will be seen during which a, certified developmental specialists, blinded to the intervention, will administer the Bayley Scales of Infant and Toddler Development, Fourth Edition (Bayley-4) to assess infant neurodevelopment. The PROMIS Measures and the Brief Infant Toddler Social Emotional Assessment (BITSEA) will also be administered during the 24-month visit along with measures of growth.

3.3 Randomization

Stratification: Randomization is stratified by hospital. This is critical to minimize the chance of differences between intervention arms in hospital practices, provider practices, and maternal characteristics.

Randomization: After written informed consent or oral consent (written informed consent to be obtained later) was obtained, qualifying infants were randomized to either rapid- or slow-wean intervention. The Neonatal Research Network Data Coordinating Center (NRN DCC) centrally randomize participants. They developed an allocation sequence with randomly varying block sizes, and they implement the sequence through a central process that is available 24 hours each day. The NRN DCC independently randomizes multiple births. Pharmacy personnel of each participating hospital are the only staff with access to group assignment.

3.4 Sample Size

The projected effect size is a 2.0-day difference in the duration of opioid treatment between infants randomized to a rapid-wean intervention and infants randomized to a slow-wean intervention. The sample size was derived by using results from a recently completed trial comparing morphine-treated infants and methadone-treated infants in which the standard

deviation was 6.9 days for the morphine-treated infants and 8.0 days for methadone-treated infants (see protocol for more details). A standard deviation of 6.9 days was used to derive the sample size, anticipating that more infants will be treated with morphine. However, the primary analysis compares treatment regardless of drug used, and sample size was calculated irrespective of the proportion of infants treated with morphine or methadone.

The study will have two intervention arms (rapid-wean and slow-wean) with 251 morphine/methadone treated infants per intervention arm, for a total of 502 morphine/methadone treated infants. This will achieve 90% power to reject the null hypothesis with a significance level of 0.05 using a two-sided two-sample t-test. The null hypothesis will be morphine or methadone treated infants in both arms will have equal means when the population difference in the duration of opioid treatment is 2.0 days with a standard deviation of 6.9 days. Sample size recalculation was performed in July 2022, using a standard deviation of 6.2 based on data from enrolled infants up to June 2022. It was determined that a sample size of 404 is needed to reach 90% power to reject the null hypothesis with a significant level of 0.05 using a two-sided two-sample t-test, and 302 to reach 80% power. We chose this latter number as our target enrollment.

Subsequently, given continued enrollment challenges faced by this trial, the DSMC and NICHD decided to terminate the trial after December 31, 2023.

3.5 Analysis Structure

The main goal of this analysis is to evaluate whether rapid-wean intervention (15% decrements from the stabilization dose) for infants with NOWS will shorten the days on treatment as compared with a slow wean intervention (10% decrements from the stabilization dose). The secondary goals of analysis will be completed to better understand the effect of rapid-wean by specific pharmacologic intervention (morphine vs methadone), proportion of infants with escalation/resumption of opioid during the intervention, and total amount of opioid used among infants in the rapid and slow wean intervention.

3.6 Statistical Interim Analysis and Stopping Guidance

All interim analyses will utilize Bayesian modeling and predictive posterior inference based on neutral, enthusiastic, and skeptical priors.

Safety Interim Analysis

Per protocol, safety will be assessed after 25%, 50%, and 75% of the enrolled infants are medically ready for discharge. At the 25% mark, we only had one infant with non-related seizure. Interim safety analysis was not performed.

Futility and Efficacy Analysis

Per protocol, efficacy and futility will be assessed after 50% of the enrolled infants are medically ready for discharge.

Important implications and recommendation

Due to slow study enrollment, the data safety and monitoring committee (DSMC) met on August 5, 2022. At the end of this meeting, DSMC recommended enrollment into ACT NOW

Weaning to continue until a target date of December 31, 2023. The DSMC will continue to review recruitment and safety data every 6 months. The DSMC will not look at the primary outcome data again until the end of the study.

Based on this decision, no interim analysis was performed after August 5, 2022.

3.7 Timing of Interim and Final Analysis

Interim analysis

It will not be performed, see section 3.6.

Final analysis

Final analyses are planned to occur after March 2023. We are expecting data lock for the primary and some of the secondary hypotheses by end of March 2024. Follow up data is expected by end of June 2026. Consequently, data analysis for the first phase is expected to commence in April 2024, and for the second phase (follow-up secondary outcomes) in July 2026. . The analysis database is constructed using all enrolled and consented infants.

4 Statistical Principles

4.1 Modeling Process

For the primary outcome (H1), linear mixed model with site as random effect and intervention as a fixed effect will be used.

For safety outcomes, binomial proportions and their corresponding 95% CI by intervention arm for seizure (clinical and EEG), excessive stool output, respiratory disturbances, and feeding intolerance will be reported. Chi-square/Fisher exact tests will be used to compare between the intervention arms. If we end up with a very small number of adverse events and severe adverse events, analysis of these measures will be descriptive only.

For the secondary linear outcomes (H2, H3, H5 & H12d), linear mixed model with site as random effect will be used. For secondary binary outcomes (H4, H6), generalized linear mixed model with site as random effect will be used. For count outcomes (H12a-12c), generalized linear mixed models using either Poisson or negative binomial distribution and site as random effect will be used. Intervention will be included in the models as a fixed effect. Secondary outcome models will be adjusted for breastfeeding status, maternal treatment, and stabilization dose.

Hypotheses H9 – H11 will be analyzed in a similar manner to H2. For the PROMIS model (H10), we will adjust for 1-month after discharge and if mother is the primary care giver as fixed effect covariates.

Hypothesis H8 will be analyzed using latent profile analysis as specified in Liu et al. (2). Model fit indices and likelihood ratio tests will inform the determination for the most-appropriate number of classes.

4.2 Protocol Violations/Deviations

Protocol violations (PV) and protocol deviation (PD) for pharmacy and non-pharmacy will be descriptively summarized in terms of infants per arm with any PV or PD and in terms of number of PV-PD per arm.

4.3 Analysis Datasets

The analyses for each outcome will be performed using an intention-to-treat (ITT) framework, defined as the inclusion of all infants randomized and consented into the study and analyzed according to the trial arm to which the infant was randomized to.

5 Trial Population

5.1 Screening population

Research personnel will screen medical records of pregnant mothers to identify mothers who use opioids, and screen charts of infants with known opioid exposure in utero and infants treated for NOWS in all areas of the hospital in which infants may receive care. After an infant is born from this population, they will be evaluated against the inclusion and exclusion criteria. Infants will be tracked to ensure eligibility by verifying the following: NOWS signs occurred, clinical team-initiated morphine or methadone treatment, and weaning after stabilization on an opioid dose has not started.

5.2 Withdrawal / Follow-up

Caregiver of infants withdrawn from the study will be given three options: 1. Discontinuation of study intervention with permission to continue study specific assessment and data collection, 2. Withdrawn with permission to continue data collection per protocol, or 3. Withdrawn with discontinuation of further data collection. If withdrawal status allows for continued data collection, then information about the number of days infants were treated for NOWS with opioids after withdrawal from the intervention will be collected.

5.3 Analysis Populations

The Full Analysis Set (FAS)[1] population will serve as the population for the analysis of efficacy data in this trial. The FAS population consists of all subjects randomly assigned to treatment who completed the treatment, this will include non-withdrawn infants as well as withdrawn infants with permission for continued data collection.

Survival analysis and safety analysis will utilize Intention-to-treat population, defined as the inclusion of all randomized participants. For the survival analysis, efficacy outcome for withdrawn infants with discontinuation of further data collected will be right censored.

5.4 Baseline Patient Characteristics

Infants and maternal baseline characteristics will be compared across treatment groups. Table 1 and 2, respectively. The table shells list out the characteristics that will be summarized. Other baseline characteristics may be investigated.

6 Statistical Analysis

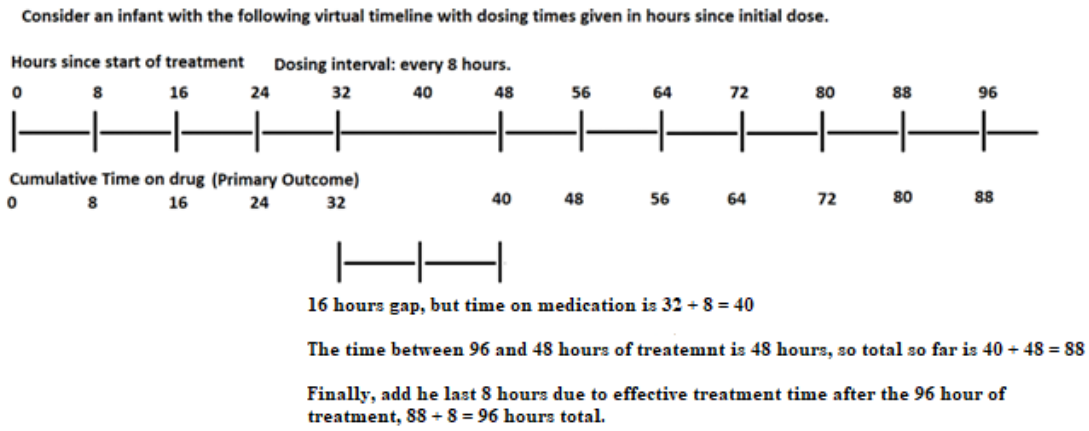
6.1 Primary Outcome

6.1.1 Description of Primary Outcome Calculation

In the protocol, the primary outcome was defined as “The number of days of opioid treatment (used as primary treatment), including escalation, resumption, and spot treatment, from the first weaning dose to cessation of opioid”. Additionally, the protocol specifies, “we will define days on a 24-hour basis, e.g., 18 hours will represent 0.75 days”. For data collection on Form Weaning07, data was not collected at this detailed level of time and date for every dose. Instead, data was collected by steps with given time and dose of first dose in each step and total number of doses received by patient at this protocol step.

Considering the factors mentioned above, several aspects must be considered when determining the primary outcome. Initially, we identified time gaps that were at least two times the dosing intervals and excluded them from the calculation of the time spent on medication. It is crucial to note that during the first interval of these gaps, the infant was on treatment, which we will refer to as 'effective time on treatment.' These gaps are technically part of the duration of effect of the medication. Consequently, the time on the drug needs to be adjusted by one dose interval for each identified time gap.

Applying a similar rationale, an infant is considered "effectively on treatment" for one dose interval following the administration of the last study dose. Therefore, an additional dose interval is added to the cumulative total obtained thus far. To enhance clarity, the following example illustrates the methodology used in calculating the primary outcome for a non-withdrawn infant: The following figure represented a partial virtual timeline for an infant who was not withdrawn and was on Methadone treatment at an 8-hour interval. For this infant, the value of the primary outcome after looking at the dosing times through the first 96 hours is 96 hours or 4 days. In this instance, notice that we have a gap of 16 hours between the 5th and 6th dose, but the 5th dose was effective through hour 40 (32 + 8 hours). Next, the infant was on medication from 48-hour timepoint to 96-hour timepoint, this means 48 hours of treatment. So total time on treatment so far is 40 + 48, resulting in 88 hours. Finally, we add another 8 hours due to the effective treatment time after the 96-hour timepoint. This results in the final calculation of 96 hours on treatment.



Withdrawn Infants: One additional step is needed for infants who withdrew but have consented to continue data collection. On the Weaning12 form, section B, Q5b, we have a field where we can get "number of days infant was treated for NOWS with opioids after infant was withdrawn from the intervention". This number represents the days on treatment after withdrawal. As such, for infants who are withdrawn and give consent for continued data collection, we add the value recorded on Weaning 12, section B, Q5b to the value calculated above. Primary outcome for infants who withdrew with discontinuation of data collection will be censored and only included in the survival model.

Finally, per the protocol, "Infants will exit the study intervention without unblinding (remain in the trial) if they have not weaned off study drug by 35 days (inclusive of the 35th day) form the first weaning dose." To accommodate this, the primary outcome will be truncated at 35 days regardless of the study arm. The non-truncated data will be utilized in the planned survival sensitivity analyses using cutoff at 35 days as censoring indicator.

For full details on primary outcome calculation, please refer to PrimaryOutcomeCalculationWeaning.pdf

6.2 Safety Outcomes (H7)

Safety outcomes to be considered in the analysis will include the following:

- Seizures, clinical and/or EEG
- Excessive stool output
- Respiratory disturbance
- Feeding intolerance

6.3 Secondary Outcomes

Secondary outcomes that will be considered during initial analysis will include but not limited to:

- Number of days on morphine from first weaning dose to cessation of opioid for slow- and rapid-wean groups (H2)

- Number of days on methadone from first weaning dose to cessation of opioid for slow- and rapid-wean groups (H3)
- Escalation or resumption of opioid during treatment (Y/N scenario)) (H4)
- Total amount of opioid from first weaning dose to cessation of opioid for non-withdrawn infants (H5)
- Initiation and/or escalation of second- or third- line drugs to treat NOWS signs from the first weaning dose to cessation (H6)
- Atypical neurobehavioral profile prior to discharge using the NICU Network Neurobehavioral Scale (NNNS)(H8)
- Length of hospital stay (H9)
- Maternal well-being measured using PROMIS, one month post discharge and at 24-month follow-up visit (H10)
- Maternal-infant attachment measured using MPAQ total scores, one month post discharge (H10)
    - Maternal well-being
    - maternal-infant attachment
- Infant growth (H11)
    - Weight, height, head circumference at 24-month follow-up visit
- Infant wellness after discharge and initial 24 months of age looking at ER and acute care facility visits over time (Y/N for time (1-4)) (H12)
- BISTEA consists of two multi-item scales, a problem scale (31 items) and competence scale (11 items), with high score on the problem scale or a low score on the competence scale being less favorable. (H12). These scores will be converted to percentiles based on BITSEA Examiner's Manual (BITSEA ph2 Manual MapiTrust.pdf)
- Infant development to 24 months of age (H13)
    - Bayley cognitive, language, and motor performed by a developmental specialist.

6.4 Frequentist Statistical Methods

6.4.1 Statistical Methods

The proposed analyses will be conducted using the intent to treat (ITT) population.

Consort Diagram

Patient accounting will be presented in a consort diagram. This diagram will include:

- Number of infants screened
- Number of infants eligible
- Number of infants consented
- Number of infants randomized
- Number of infants randomized by group
- Number of infants in each medication category
- Number of infants with primary outcome per group per medication

Descriptive Tables

Baseline characteristics and outcomes will be descriptively summarized in the following manner:

Binary (yes/no, 1/0 type) measures will be summarized by level with frequency out of the total non-missing values and percentage of total non-missing values. Percentages will be reported to one decimal place, nn(XX.X%) or nn/NN(XX.X%)

Multiple level discrete measures will be tabulated with one row per level and with a row indicating how many observations had missing data. Percentages will be calculated as percent of all observations associated with a given treatment.

Continuous measures will be summarized by number of non-missing values, means and standard deviation [n: XX.X (XX.X)], or by median with minimum and maximum, or by median with first and third quartile [n: XX.X (XX.X , XX.X)]. Mean, standard deviations, median, minimum, maximum, and quartiles will be reported to one decimal place.

Count data will be summarized using median (min, max).

Statistical Modeling and Inferences

Statistical modeling will be performed using linear mixed model for continuous measures and generalized linear mixed model for binary and count outcomes with logit and log link functions, respectively. All models will include site as random effect and intervention arm as a fixed effect. Secondary outcome models will also be adjusted for maternal treatment, stabilization dose, and breast feeding.

For the ith infant, all models (except the primary) will have the following form $g(\mu_i)$ with $g(.)$ being a monotonic link function and the following form:

$$g(\mu_i) = \beta_o + \beta_1 I(\text{intervention=fast}) + \beta_2 I(\text{Maternal Treatment=methadone}) + \beta_3 I(\text{Materna Treatment= Buprenorphine} \pm \text{naloxone}) + \beta_4 I(\text{Maternal Treatment=other opioids}) + \beta_5(\text{stabilization dose}) + \beta_6(\text{Maternal feed=Yes}) + \beta_j X_j$$

where:

μ_i = expected value of outcome y_i

β_0 = intercept, the expected mean number of treatment days when all covariates = 0

β_1 = Intervention effect of fast vs slow wean interventions

β_2, β_3, β_4 = effect of maternal treatment, compare to no treatment

β_5 = effect of stabilization dose

β_5 = effect of breastfeeding

β_j = effect of the jth covariate added to the model

For normal outcomes, $g(.)$ is the identity link function.

For binary outcomes, $g(.)$ is the logit link function and we will assume that y_i follows Bernoulli distribution. Alternatively, we can follow Zou et al. (2003) and use the natural log function, and we will assume that y_i follows Poisson or negative binomial distribution. Decision will be based on manuscript writing team, if they prefer to report risks vs odd ratios.

For count outcomes, $g(.)$ is the natural log function, and we will assume that y_i either follows Poisson or negative binomial distribution.

All hypothesis tests will be performed using a two-sided significance level of 0.05. Unless otherwise specified, results for each type regression performed will include the following:

- Normal regression
    - Unadjusted
        - Mean (and 95% confidence intervals (CI)) of the outcome in each treatment group.
        - Mean difference (and 95% CI) of the outcome between treatment groups.
    - Adjusted using pre-specified covariates.
        - Mean (and 95% CI) of the outcome in each treatment group.
        - Mean difference (and 95% CI and p-value) of the outcome between treatment groups.
- Logit Binomial regression
    - Unadjusted
        - Rate (and 95% CI) of the outcome in each treatment group
        - Odds Ratios (and 95% CI) of the outcome between treatment groups
    - Adjusted using pre-specified covariates.
        - Rate (and 95% CI) of the outcome in each treatment group
        - Odds Ratios (and 95% CI and p-value) of the outcome between treatment groups.

- Log-Binomial or Poisson regression
    - Unadjusted
        - Risk (and 95% CI) of the outcome in each treatment group
        - Relative risk (and 95% CI and p-value) of the outcome between treatment groups
    - Adjusted using pre-specified covariates.
        - Risk (and 95% CI) of the outcome in each treatment group
        - Relative risk (and 95% CI and p-value) of the outcome between treatment groups.
- Poisson or negative binomial regression
    - Unadjusted
        - Rate (and 95% CI) of the outcome in each treatment group
        - Relative rate (and 95% CI and p-value) of the outcome between treatment groups
    - Adjusted using pre-specified covariates.
        - Rate (and 95% CI) of the outcome in each treatment group
        - Relative rate (and 95% CI and p-value) of the outcome between treatment groups

6.4.2 Assumption Checks

For continuous outcomes, linearity will be assessed by plotting the conditional residuals from the models against the value of the linear predictor. No obvious trend should be evident (i.e., uniform horizontal band present). In the case we do not see a good fit, continuous outcome models will be fitted using alternative distributions.

For binary outcomes, if we do not have the needed sample (may run into convergence issues), then outcome will be analyzed using Fisher Exact.

For count outcomes, models will be fitted using either negative binomial or Poisson models. Choice of distribution will be based on dispersion and/or convergence issues.

Given the limited number of infants in some centers, if we run into convergence issues that cannot be resolved using standard statistical approaches, we will consider combining sites based on geographical locations. We will also consider reducing the number of covariates.

6.4.3 Covariates

All secondary regression models will include the stratification variable (site) as a random effect and will be adjusted for:

- Breastfeeding status
- Stabilization dose, continuous
- Maternal treatment which will be categorized as follow:
    - Methadone use
    - Buprenorphine ± naloxone
    - Prescribed opioids other than methadone and buprenorphine ± naloxone
    - None/Unknown

  In case of mixing among the 4 groups the following rules will be followed:

    - If methadone is used with other opioid, classify as methadone
    - If buprenorphine is used with other opioid, classify as Buprenorphine ± naloxone
    - If methadone is used with buprenorphine, classify based on which drug has been used most proximal to delivery.

Additionally, if the sample size allows, secondary analyses can be adjusted for critical covariates that are substantially imbalanced at baseline (both clinically and statistically significantly different)

6.4.4 Subgroup Analysis

For the primary outcome the following subgroup analyses will be performed:

1. By race/ethnicity
2. By sex

Each model will be similar to the analysis described in section 6.4.1 but will also include fixed effect for the stratification variable as well as an interaction between intervention and the stratification variable. If

p-value for interaction is less than 0.10 then we will report the treatment effect within the specified subgroup. We will not correct for multiple testing as all analyses other than the primary are considered exploratory.

6.4.5 Sensitivity Analysis

Because of censoring and/or partial data collected on withdrawn infants we will perform the following additional analyses:

1. Replicate the primary analysis (H1) to include maternal treatment, stabilization dose, and maternal feeding.
2. Given the learning curves needed to be successful at this trial, we will run a sensitivity analysis limited to centers that enrolled greater than 10 infants.
3. Replicate the primary outcome analysis to evaluate the effect of being part of Eat Sleep Console
4. Perform survival analysis using cutoff at 35 days as censoring indicator.
5. If sample size allows, competing risk analysis will be performed using Cox proportional hazard model adjusting for the following: parental withdrawal, physician withdrawal, and treatment failure defined as unable to wean by 35 days of methadone/morphine treatment. This analysis will allow for additional inquiry into the intervention effect on the primary outcome while accounting for competing safety and withdrawal risks.
6. For H13, perform a sensitivity analysis for all hypotheses by removing follow ups outside the 22-26 weeks window.

6.5 Bayesian Statistical Methods

6.5.1 Statistical Methods

Due to sample size limitation, the primary outcome (H1) and sensitivity analyses (1) and (2) will be analyzed using Bayesian methods in addition to the frequentist analyses to estimate the posterior distribution for all model parameters. Bayesian models will be fit using Markov chain Monte Carlo methods (MCMC) by using software such as PROC MCMC (SAS) or stan (RSTAN/RSTANARM package in R). Model parameter estimates will be generated by fitting three MCMC chains with randomly drawn starting chain. Each chain will involve 1,000 burn-in iterations and 1,000 additional iterations for each chain. Thinning will be used as necessary to reduce autocorrelation among the samples to produce good sampling of the posterior. Trace plots and Gelman-Rubin convergence diagnostics (Rhat <1.1) for all parameters will be used to monitor convergence. For all models, we will report posterior medians and 95% credible intervals (CrI) for the group difference and the probability that a rapid-wean will reduce the days of opioid treatment, compared to a slow-wean intervention.

Primary outcome (H1)

Linear regression model will be used for the primary outcome with treatment group as fixed effect and as hospital random effect. We will use the following priors:

trt effect $\sim Normal\ (\mu = 0, \sigma = 3)$

Intercept $\sim Normal\ (\mu = 0, \sigma = 10)$

All other covariates $\sim Normal\ (\mu = 0, \sigma = 2)$

Standard deviation of random effect (hospital) $\sim half - noraml\ (\mu = 0, \sigma = 1)$

6.5.2 Bayesian Sensitivity Analysis

The primary outcome will be analyzed using skeptical and enthusiastic priors. The skeptical prior will be centered at a mean difference of 2 and the enthusiastic prior at a mean difference of -2. In both cases $\sigma = 3$. All other priors will be set in a similar manner to the primary analysis.

Skeptical trt effect $\sim Normal\ (\mu = 2,\ \sigma = 3)$

Enthusiastic trt effect $\sim Normal\ (\mu = -2, \sigma = 3)$

6.6 Missing Data

6.6.1 In-hospital data

Analysis will be completed on the ITT principle, so data will be analyzed as randomized. No imputation for missing outcome data is planned, so the analysis will be based on the Full Analysis Set.

6.6.2 Follow-Up data

For follow-up measures obtained under provision of consent, the study team will continue its commitment to data quality. When missing data are present, we will investigate the missing data pattern to see if the dropouts are independent of observed data. An exploratory approach to examine missingness mechanisms will be explored using logistic regression to identify potential participant characteristics associated with presence of missing data. In the event of large amounts of missingness for a given outcome or influential covariate, according to an agreed upon pre-specified threshold (e.g., more than 5%), we will explore multiple imputation analyses[3,4]. We will also examine rate of missing data for all variables, site, or intervention arm. For data missing-at-random (MAR) or missing-not-at-random (MNAR), we will use selection models such as MNAR Dale or Diggle- Kenward model.22 These models often require strong assumptions about the dropout mechanism, which are unverifiable using the observed data. We will conduct sensitivity analyses to investigate the sensitivity of our conclusion to possible violation of such assumptions. For covariates for which the MAR assumption is justified, multiple imputation using sequential regression techniques will be implemented. All models will be fitted to each imputed data set and results combined using Rubin's rules that account for variability across imputations[3].

6.7 Statistical Software

SAS and R

6.8 List of Potential Displays

- Figure 1: Consort Diagram
- Table 1: Maternal Demographic Characteristics of Enrolled Infants Participants
- Table 2: Demographic Characteristics of Enrolled Infant Participants
- Table 3: In-Hospital Outcomes
- Table 4: Adverse and Serious Adverse Events
- Table 5: Protocol deviation/protocol violation Summary
- Table 6: In Hospital Status
- Table 7: Bayley scores represented using cutoffs specified in section 2.2.1

- Additional tables to report model results as specified in section 6.4.1 may be provided

6.9 Reporting

Unless required 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-values will be reported to 3 decimal places if > 0.001. If it is less than 0.001 then report '<0.001'.

7 References

1. E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials (fda.gov)
2. Liu, J., Bann, C., Lester, B., Tronick, E., Das, A., Lagasse, L., ... & Bada, H. (2010). Neonatal neurobehavior predicts medical and behavioral outcome. Pediatrics, 125(1), e90-e98.
3. Jones HE, Seashore C, Johnson E, Horton E, O'Grady KE, Andringa K, et al. Psychometric assessment of the Neonatal Abstinence Scoring System and the MOTHER NAS Scale. Am J Addict 2016, 25(5): 370-373.
4. 21. Mehta A, Forbes KD, Kuppala VS. Neonatal Abstinence Syndrome Management From Prenatal Counseling to Postdischarge Follow-up Care: Results of a National Survey. Hosp Pediatr 2013, 3(4): 317-323.