Unique Protocol ID: IRB00036246
Title: Efficacy of Intrathecal Oxytocin to Speed Recovery After Hip Surgery
NCT03011307
Document Date: 1/03/2022

Section 1: Administrative information

1.a. Title

Statistical analysis plan for Efficacy of Intrathecal Oxytocin to Speed Recovery from Hip Surgery: a single-center randomized controlled trial to assess whether a single intrathecal injection of oxytocin at the time of total hip arthroplasty (TKA) hastens recovery from pain over the next 2 months.

1.b. Trial registration

Trial Registration: NCT03011307

2. SAP version

Version: 2.0 Date: December 15, 2021

3. Protocol Version

This document has been written based on information contained in the study protocol version 1.0, dated December 13, 2016.

4. SAP Revisions

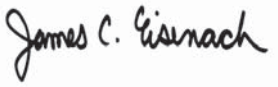
Protocol version	Updated SAP version no.	Section number changed	Description and reason for change	Date changed
1.0	2.0	Section 6	Details provided after data acquisition and before unblinding	Dec 15, 2021

5. Roles and Responsibilities – non-signatory names and contribution

James C. Eisenach, M.D.
Professor of Anesthesiology
Wake Forest School of Medicine
Winston-Salem, NC
Role: Chief investigator

Timothy T. Houle, M.D.
Associate Professor of Anesthesiology
Massachusetts General Hospital
Boston, MA
Role: Co-investigator and Senior Statistician

6.a. Roles and Responsibility – signature of the person writing the SAP



01/03/2022

James C. Eisenach, M.D.
Chief Investigator

Date

6.b. Roles and Responsibility – signature of senior statistician responsible

01/03/2022

Timothy T. Houle, Ph.D.
Co-Investigator & Senior Statistician

Date

Section 2: Introduction

7. Background and Rationale

Recovery from pain after surgery varies greatly among individuals, with some who are pain free within days and others who experience pain for years. Recovery from pain after cesarean delivery is more rapid than after other pelvic surgeries, and work in animals suggest this reflects actions of the hormone oxytocin in the spinal cord. This study is the first to test whether intrathecal injection of oxytocin at the time of surgery will speed recovery from pain after surgery.

8. Objectives

Research hypothesis: The null hypothesis is that there is no difference in modeled trajectory of change in pain intensity report after TKA between intrathecal injection of saline and of oxytocin. The alternative hypothesis is that there is a difference between the two groups.

Study objective: The primary objective is to determine the effectiveness of intrathecal oxytocin to speed recovery from pain, as measured by growth curve modeling, after TKA.

Secondary objectives are:

a. To determine the effectiveness of oxytocin to speed gain in activity, assessed by actigraphy and analyzed by growth curve modeling, after TKA
b. To determine the effectiveness of oxytocin to speed recovery from self-assessed disability, as measured by growth curve modeling, after TKA
c. To determine the effectiveness of oxytocin to hasten time to cessation of opioid use, as measured by growth curve modeling, after TKA

Section 3: Trial Methods

9. Trial design

The trial is a single center, randomized, parallel-group, placebo-controlled trial. Treatment allocation is a 1:1 ratio. Patients are randomized to either oxytocin or saline control.

10. Randomization

Participants were randomized into treatment arms in a 1:1 ratio using permuted blocks of sizes 4 to 8. Randomization assignments. The 'blockrand' package in R was used by the study statistician. Allocation concealment was maintained using a centralized pharmacy administration of study drug. Both study agents are designed to look identical to each other to maintain blinding.

11. Sample size

Enrolling N = 120 participants (n = 60 per study arm) provides sufficient power to examine the study's primary hypothesis. Assuming that each individual contributes ~ 60 daily measurements, with alpha = 0.05, power = .80, a modest level of within-subject correlation, and adequate model fits (SD of the residuals < 2.0), we will be able to detect difference in slopes of d = 0.09 units. Thus, if the saline group exhibited a mean pain score of 8.0 VAS units after surgery and decreased on average 0.38 points each day (i.e., expected full recovery at 21 days), sufficient power exists to detect an enhanced slope of 0.47 in the oxytocin group (i.e., expected full recovery in ~17 days). This difference is a clinically meaningful effect size that could substantially improve the life of patients.

12. Framework

Superiority testing will be used to examine differences between the treatment arms.

13a. Information on interim analyses

There were no planned interim analyses. An interim analysis was performed on Feb 8, 2019 after 49 subjects had completed the 2 month daily diary portion of the study. This was performed by the study statistician for the purpose of an NIH grant submission, with the study team remaining blinded to allocation assignments. Only descriptive statistics were performed and there was a decision prior to the analysis that early stopping would not be allowed based on the results.

13b. Any planned adjustment of the significance level due to interim analysis

None.

13c. Details of guidelines for stopping a trial early

There was no plan to stop the trial early aside from recommendations based on unanticipated serious adverse events by the DSMB.

14. Timing of final analysis

Final analysis of the oxytocin vs saline comparison will take place when all patients have completed the 12 month follow-up and data for the primary endpoint have been cleaned (anticipated January 2022). The first main report/publication of the primary endpoint will be prepared for submission, anticipated by May, 2022.

15. Timing of outcome assessments

The schedule of study procedures is provided in the protocol and summarized in the table below.

Phase	Ortho Clinic	V1 (COMPLETED 2-4 WEEKS PRIOR TO SURGERY)	14 DAYS Pre op-1 day pre op	DOS	POD 1	POD 2-POD 56	Post op Week 1-8	Post op V2 (8-10 weeks)	Monthly follow up 3 mos	Monthly follow up 4 mos	Monthly follow up 5 mos	Monthly follow up 6 mos	Monthly follow up 12 mos
Obtain/Confirm Informed Consent	X	X		X									
Obtain Medical History	X	X											
Urine Pregnancy Test if applicable				X									
Questionnaires		X						X					
Iowa Gambling		X						X					
Card Sorting		X						X					
Daily Diary			X	X	X	X							
Weekly Diary (WHODAS)							X						
Step tracker			X		X	X							
Verbal Pain				X	X								
Monthly Diary									X	X	X	X	X

Section 4: Statistical Principles

16. Level of statistical significance

All applicable statistical tests will be 2-sided and will be performed using a 5% significance level.

17. Adjusting for multiplicity

The primary analysis involves a single model-based inference that will be interpreted at the 5% threshold of statistical significance. Multiple secondary outcomes will be evaluated and each will be interpreted at the 5% level of significance with no adjustments made for multiplicity.

18. Confidence intervals

All confidence intervals presented will be 95% and two-sided.

19a. Definition of adherence to the intervention and how it is assessed

Compliance is assessed based on the number of daily diary entries after hospital discharge. It is defined as:

% compliance = (number of entries/number of days entries were to be completed)*100%

19b. How adherence will be presented

The number and % of participants completing more than 80% of the daily diaries after surgery will be provided by treatment group in a table. All available data will be included in the analysis from the intention to treat sample.

19c. Definition of protocol deviation

The following are pre-defined major protocol violations with a direct bearing on the primary outcome:

1) Surgery during the primary outcome assessment period (first 8 weeks after TKA surgery)

19d. Description of which protocol deviations will be summarized

Protocol deviations are classified prior to unblinding of treatment. The number (and percentage) of patients with major and minor protocol deviations will be summarized by treatment group with details of type of deviation provided. The patients that are included in the ITT analysis data set will be used as the denominator to calculate the percentages. No formal statistical testing will be undertaken.

20. Analysis population

The intention-to-treat population will include all randomized patients who received study drug.

Section 4. Trial Population

21. Screening data

Enrollment: The dates of recruitment, the number of patients screened, the number of patients recruited, the number of screened patients not recruited, and the reason for non-recruitment will be reported as part of the CONSORT diagram.

22. Eligibility criteria

Inclusion and exclusion criteria are provided in the protocol. The number of ineligible patients randomized, if any, will be reported, with reasons for ineligibility

23. Information to be reported in the CONSORT flow diagram

A CONSORT flow diagram (appendix A) will be used to summarize the number of patients who were:

- assessed for eligibility at screening
- eligible at screening
- ineligible at screening*
- eligible and randomized
- eligible but not randomized*
- received the randomized allocation
- did not receive the randomized allocation*
- lost to follow-up*
- discontinued the intervention*
- randomized and included in the primary analysis

*reasons will be provided for post randomization exclusions.

24a. Description of level of withdrawal

The level of consent withdrawal will be tabulated (classified as “consent to continue follow-up and data collection” “consent to continue data collection only”, “complete – no further follow-up or data collection”).

24b. Timing of withdrawal/lost to follow up data

This will be presented in CONSORT diagram format rather than as a table, with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage

24c. Reasons for withdrawal/lost to follow up data

The numbers (with reasons, if available) of losses to follow-up (drop-outs and withdrawals) over the course of the trial will be summarized by treatment arm.

25a. List of baseline characteristics to be summarized

Patients will be described with respect to age, sex, race, ethnicity, medications, number of painful conditions excepting the hip and their locations, questionnaire and cognitive game scores, baseline pain and activity from the daily diaries prior to surgery, attending surgeon, and surgical approach (anterior vs posterior).

25b. How baseline characteristics will be described

Categorical data will be summarized by numbers and percentages. Continuous data will be summarized by mean, SD and range if data are normal and median, IQR and range if data are skewed. Minimum and maximum values will also be presented for continuous data. Tests of statistical significance will not be undertaken for baseline characteristics; rather the clinical importance of any imbalance will be noted.

Section 5: Analysis

26. Outcomes and timings

Primary outcome: Worst pain intensity measured using a 0 to 10 numerical rating scale on a daily diary. Measurements will be obtained each evening from day 0 (surgery) to day 56 after surgery. The pain measurements will be evaluated in a statistical model using an intercept (i.e., immediate pain after surgery) and slope (i.e., rate of change of pain during the observation period).

Secondary outcomes: Daily measurements such as physical activity, as measured via actigraphy, disability, as measured by daily diaries and weekly WHODAS 2.0 scores, and opioid self-administration, as measured by daily diaries will be collected from day 0 (surgery) to day 56 after surgery.

Questionnaire and cognitive game data will be assessed at baseline and follow-up. Validated scoring systems from the individual items will be used to score each questionnaire or measure. For the purpose of the primary analysis publication, summary scores with measure of central tendency and variance will be provided for oxytocin and saline groups as mentioned in Item *25b*.

27a. Analysis methods to be used
and
27b. Adjustments for covariates

Primary analysis:

To examine how participants recover over the repeated measurement occasions, we will employ hierarchical linear modeling (HLM) [1]. The application of HLM to individual change data is well suited for the expected changes that occur after surgical injury. HLM is a polynomial regression with random effects where interaction terms are used to examine impacts on the change parameters (e.g., treatment group x intercept; treatment group x slope). Because of the nature of the expected change, the trajectories will be modeled using parameters consisting of an intercept and the slope estimated using the natural log of time (change in recovery). The intercept is coded to represent the initial level (first day after hospital discharge) of the outcome variable. The slope estimates the change over ln of time and represents the increase or decrease in the variable. Through the use of fixed and random effects, we will then be able to examine the influences on the changes in pain. In this way, we test efficacy of the test intervention, intrathecal oxytocin.

To examine the proposed hypothesis, we have *a priori* defined key fixed effects that will be included in the primary model, based on previous literature regarding their influence on recovery after surgery. This adjusted analysis is the primary analysis. Each of these predictors is described below:

Oxytocin: Randomized treatment assignment of either oxytocin or placebo will be a fundamental predictor of trajectory.

Patient characteristics: To adjust the associations for individual characteristics, age, sex, number of painful conditions beyond the operated hip, ongoing use of opioids at the time of recruitment, and PROMIS depression score will be entered into the model.

Preoperative study data: Degree of daily pain and its variance during the two weeks prior to surgery as assessed by daily diaries will be entered into the model.

These individual predictors sets will then be used to test the primary hypothesis that oxytocin alters time course of recovery after TKA. The primary statistical inference will be a likelihood ratio comparing two models estimated with and without oxytocin group assignment. Thus, the pain trajectory model comparing medication assignment (oxytocin versus placebo) adjusted for the covariates of characteristics and preoperative study data will be compared to a similar model without medication assignment using a Chi-squared test. This approach allows a single direct test of the primary hypothesis and will be interpreted using a two-tailed hypothesis test using $p < 0.05$. A statistically significant difference will be interpreted as evidence to support the primary hypothesis that pain trajectory differs between oxytocin and saline treatment.

Secondary analyses: Trajectories for daily activity, as measured via actigraphy, disability, as measured by daily diaries and weekly WHODAS 2.0 scores, and opioid self-administration, as measured by daily diaries, will be analyzed using a similar approach to that of the primary outcome, with one exception. For the activity analysis, preoperative data covariate included will not be pain, but rather degree of daily actigraphy activity and its variance during the two weeks prior to surgery as assessed by daily diaries. For disability analysis, preoperative data covariate included will not be pain, but rather degree of daily disability and its variance during the two weeks prior to surgery as assessed by daily diaries and baseline WHODAS 2.0 score. As noted, with the exception of the PROMIS depression score, other questionnaire and cognitive game scores will be described for the two groups but not analyzed or entered as covariates for the purpose of the primary analysis and publication. Changes in questionnaire data will be evaluated using the generalized linear model regressing follow-up score on baseline score and medication group assignment.

27c. Methods to be used for assumptions to be checked for statistical methods for primary and secondary analyses

The assumptions underlying parametric modeling will be evaluated using regression diagnostic procedures and histograms of residuals.

27d. Alternative methods to be used if distributional assumptions do not hold for primary and secondary analyses

If necessary, the primary outcome distribution will be log transformed to satisfy assumptions. For secondary outcomes, the generalized linear model allows the specification of an array of distributions (e.g., normal, binomial, gamma) with corresponding link functions (e.g., log).

27e. Planned sensitivity analyses

Three sensitivity analyses will be performed to evaluate the robustness of the trajectory analysis to the primary and secondary analyses. These are to include multiple imputation for missing data, to include monthly through 6 months and the 12-month daily diary data and to include a change point rather than log of time model, as previously described [2]. Two other sensitivity analyses will be

performed to evaluate the robustness of other causes of variability in trajectory in the primary and secondary analyses: race and surgeon.

27f. Planned subgroup analyses

No subgroup analyses are planned.

28. Missing data – reporting and assumptions/statistical methods to handle missing data

Multiple imputation will be used to account for embedded missing data in daily diary entries. The autocorrelation of each time-series for each individual along with baseline predictors will be used in conjunction with the MICE algorithm to estimate m = 20 imputations. These imputations will then be re-analyzed using the primary model with Rubin's rules used to pool the estimates of model parameters.

29. Details of any additional statistical analyses required e.g., complier-average causal effect analysis.

None

30. Harms

The number (and percentage) of patients experiencing each AE/SAE will be presented for each treatment arm categorized by severity. For each patient, only the maximum severity experienced of each type of AE will be displayed. The number (and percentage) of occurrences of each AE/SAE will also be presented for each treatment arm. No formal statistical testing will be undertaken

31. Statistical Software

The analysis will be carried out using R version 4.1.2

32. References

There are no non-standard statistical methods used, but references to methods used are listed below. As regards the Data Management Plan, data handling and cleaning were provided at Wake Forest School of Medicine and Massachusetts General Hospital, where data are secured. The Trial Master File is included in Investigator New Drug (IND) 107166, Food and Drug Administration. The Statistical Master File materials are housed at Wake Forest School of Medicine. Standard Operating Procedures followed when writing the SAP are from the guidelines published in JAMA in 2017 [3].

Cited References

1. Raudenbush, S.W. and A.S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. 2002, Thousand Oaks, CA: Sage Publications.
2. Houle, T.T., et al., *Day-to-day experience in resolution of pain after surgery.* Pain, 2017. **158**(11): p. 2147-2154.
3. Gamble, C., et al., *Guidelines for the Content of Statistical Analysis Plans in Clinical Trials.* Jama, 2017. **318**(23): p. 2337-2343.