Document Type:	Statistical Analysis Plan
Official Title:	A Double-Blind, Randomized, Crossover Study to Assess Menstrual Cramp Pain Associated with Primary Dysmenorrhea
NCT Number:	NCT03448536
Document Date:	2018-12-07

Reference Number: ID-SOP-6021
Supplement Version: 3.0

Statistical Analysis Plan Approval Form

Clinical Protocol Title: A Double-Blind, Randomized, Crossover Study to Assess Menstrual Cramp Pain Associated with Primary Dysmenorrhea	
Investigational Product: *(BAY No / INN / Trade Name):* BAY 117031 / Naproxen Sodium	
Trial IMPACT No.: 19737	**Phase: 4**
Version No./Date*(dd/mmm/yyyy)*: **07-Dec-2018**	

Signing this form constitutes agreement with the Statistical Analysis Plan or any Amendment.

APPROVED BY:

Role	Name	Signature	Date
Clinical Development Services Representative	PPD	PPD	12-DEC-2018
Trial Statistician			12/11/18
Medical Affairs Representative			12/12/18
CRO Statistician		PPD I approve this document 12 Dec 2018 19:14:54 +00:00 OnSign	

Bayer HealthCare

Clinical Study Protocol 19737

A Double-Blind, Randomized, Crossover Study to Assess Menstrual Cramp Pain Associated with Primary Dysmenorrhea

07DEC2018

Statistical Analysis Plan

Version 1.0

Prepared by:

PPD
3900 Paramount Parkway
Morrisville, NC 27560

TABLE OF CONTENTS

List of Abbreviations

AE	Adverse Event
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
CI	Confidence Interval
CMH	Cochran-Mantel-Haenszel
CSR	Clinical Study Report
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
(e)CRF	(electronic) Case Report Form
eDiary	Electronic Diary
ICH	International Conference on Harmonisation
IMP	Investigational Medicinal Product
ITT	Intent-to-Treat
IxRS	Interactive Voice/Web Response System
MedDRA	Medical Dictionary for Regulatory Activities
NRS	Numerical Rating Scale
NSAID	Nonsteroidal Anti-inflammatory Drug
OTC	Over the Counter
PID	Pain Intensity Differences
PP	Per Protocol
PT	Preferred Term
PVG	Pharmacovigilance
SAE	Serious Adverse Event
SD	Standard Deviation
SNR	Screening Number
SOC	System Organ Class
SPID	Summed Pain Intensity Difference
SPID0-12	Summed Pain Intensity Difference over 12 hours
SPID0-6	Summed Pain Intensity Difference over 6 hours
SPID6-12	Summed Pain Intensity Difference from 6 to 12 hours
TEAE	Treatment-Emergent Adverse Events
TOTPAR	Total Pain Relief
TOTPAR0-12	Total Pain Relief over 12 hours
TOTPAR0-6	Total Pain Relief over 6 hours
TOTPAR6-12	Total Pain Relief from 6 to 12 hours
WHODD	World Health Organization Drug Dictionary

1. Introduction

Primary dysmenorrhea (painful menstrual cramping that occurs during menses) is a common problem experienced by women in their reproductive years. In a study to determine the prevalence of dysmenorrhea in reproductive-age women, 36% reported always or often having pain with their menstrual periods and 90% of the women reported to have some degree of dysmenorrhea. This type of pain can interfere significantly with their daily activities. The pathophysiology of the condition is primarily prostaglandin driven. At the beginning of menses, prostaglandins are released and play a major role in inducing uterine contractions within the female reproductive system. These contractions lead to uterine ischemia which stimulates pain neurons resulting in pain. The severity of pain has been studied and correlated with the level of menstrual prostaglandin.

Naproxen sodium is a non-selective cyclooxygenase inhibitor and ultimately inhibits prostaglandins. Aleve® (naproxen sodium) is available over-the-counter and is indicated for temporary relief of minor aches and pains due to various pain states, including menstrual cramps. Tylenol® Extra Strength (acetaminophen) is also available over-the-counter and is indicated to temporarily relieve minor aches and pains due to premenstrual and menstrual cramps.

The purpose of this study is to compare the maximum single dose of Aleve® (two tablets, equivalent to 440 mg of naproxen sodium) to the maximum single dose of Tylenol Extra Strength (two tablets, equivalent to 1000 mg of acetaminophen) in the treatment of primary dysmenorrhea.

Patients with at least moderate pain due to primary dysmenorrhea who would typically take over-the-counter NSAIDs for pain relief will be enrolled in this study. The two menstrual cycle treatment design falls within established treatment guidelines. Patients who experience a treatment failure have the option of taking rescue medication for pain relief.

During the study, patients will be closely monitored for the occurrence of adverse events. Weighing between the potential risks of OTC analgesics associated with the study, and given the ability to mitigate risks through close monitoring, this study is considered clinically and ethically acceptable.

2. Objectives

The primary objective of this study is to evaluate the analgesic efficacy of a maximum single dose of two tablets of Aleve (2 x naproxen sodium 220 mg; total dose 440 mg) as

compared to two tablets of Tylenol Extra Strength (2 x acetaminophen 500 mg; total dose 1000 mg) for the treatment of menstrual cramping pain associated with primary dysmenorrhea.

The secondary objective of this study is to evaluate the safety and tolerability of naproxen sodium and acetaminophen.

3. Investigational Plan

3.1. Overall Study Design and Plan

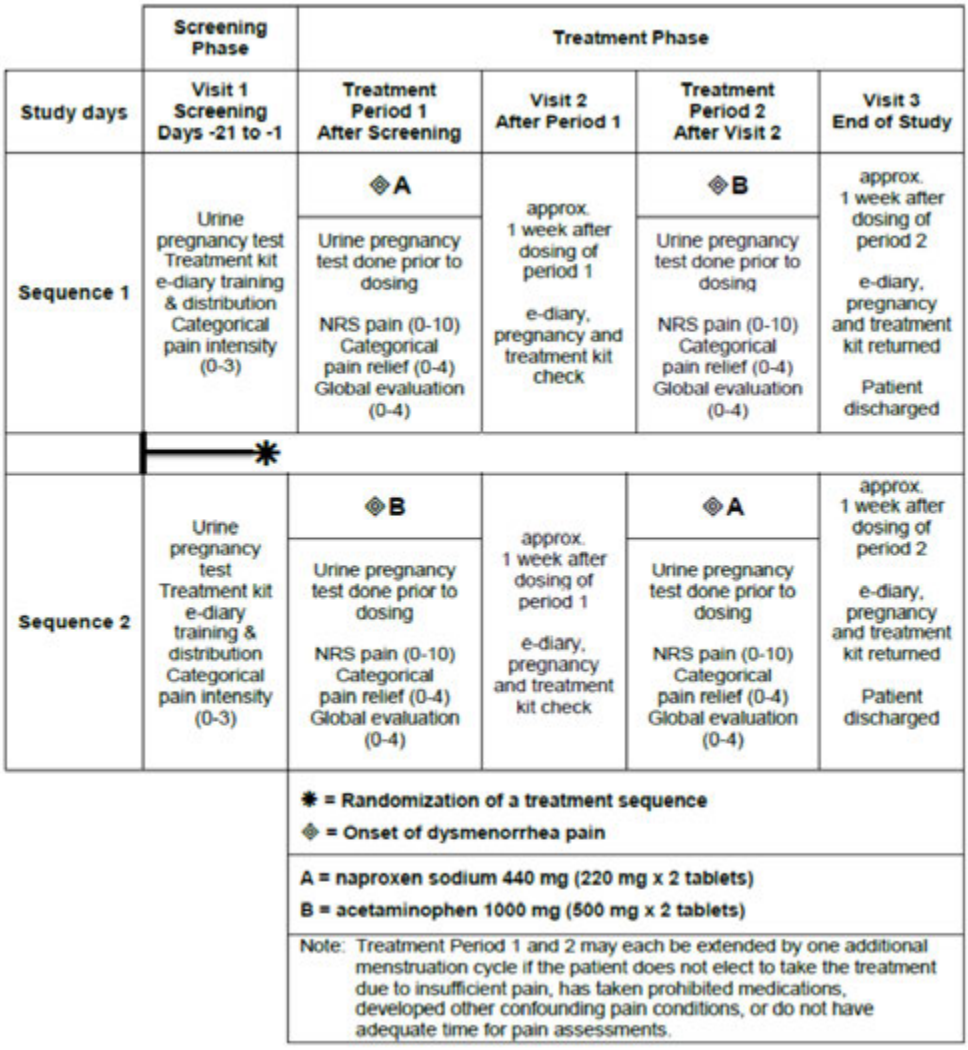
This is a multi-center, randomized, double-blind, two treatment cycle, crossover study in adolescent and adult female patients (15 to 35 years old) with primary dysmenorrhea. At the completion of the Screening phase, eligible patients will be randomized in a 1:1 crossover fashion (A:B or B:A) to a single dose of naproxen sodium 440 mg (two tablets of 220 mg [Treatment A]) and a single dose of acetaminophen 1000 mg (two tablets of 500 mg [Treatment B]).

The study consists of a 21 day Screening Phase and two Treatment Periods within the Treatment Phase. The study center may schedule screening testing on multiple days as needed, provided all screening tests are done within 21 days. Qualified patients will be randomized into one of two treatment sequences. Approximately 260 patients will be screened prior to the first Treatment period during the Treatment Phase. Approximately 200 patients will be randomized to a specific treatment sequence. Duration of trial participation will be 77 to 133 days as detailed in Figure 1. Design Overview . See Appendix 12.3 for complete schedule of events.

In order to be evaluated for this study, patients must have at least moderate menstrual pain during four of the past six menstrual cycles at screening and a Numerical Rating Scale (NRS $\geq$ 5) at the time of each treatment period. Each patient will enter their score on an eDiary prior to taking the medication for each period. If the patient's score is < 5 or don't have enough time to complete the eDiary, they will be instructed to contact the site and may be granted a maximum of one month's extension for each treatment period. If the score is $\geq$ 5, and they have a negative pregnancy test, they will be instructed to take the medication for each treatment period. Note that concomitant pain therapies must be discontinued at least 3 days prior to expected menses. If a patient withdraws from the study, they will not be replaced. Once a patient takes their study dose, they will enter the time in the eDiary, then will be prompted to enter in NRS scores (0-10). At subsequent timepoints they will be prompted to enter NRS scores and Categorical Pain Relief scores (0-4). The eDiary will be used to collect data throughout treatment phase as detailed in schedule of events.

Patients may also take a rescue medication, recommended at no earlier than 2 hours post-dose. The patient will record the time of rescue medication in the diary and complete the Global Evaluation at this time. If the patient does not take rescue medication throughout the 12 hours, they will be prompted to complete the Global Evaluation at the 12 hour timepoint.

Figure 1. Design Overview



3.2. Study Endpoints

3.2.1. Primary Efficacy Endpoint

The primary efficacy endpoint, total pain relief scores (TOTPAR) over 12 hours (TOTPAR0-12), will be calculated by multiplying the pain relief score at each post-

dose time point by the duration (in hours) since the preceding time point and then summing these values.

3.2.2. Secondary Efficacy Endpoints

The secondary efficacy endpoints include

- Summed Pain Intensity Difference (SPID) over the 12-hour study period (SPID0-12) using 0-10 NRS
- SPID over 0-6 hours (SPID0-6)
- SPID 6-12 hours (SPID6-12)
- TOTPAR over 0-6 hours (TOTPAR0-6)
- TOTPAR 6-12 hours (TOTPAR6-12)

The above mentioned endpoints will be calculated in similar manner as primary on indicated scale and timepoints. Other endpoints will include:

- Time to first intake of rescue medication
- Pain Intensity Difference (PID) scores derived from the NRS at each evaluation
- Global Evaluation at 12 hours post-dose or immediately before first intake of rescue medication

3.2.3. Safety Endpoints

Safety assessments will include the following:

- Overall incidence of treatment-emergent AEs (TEAEs), serious AEs (SAEs) will be summarized including breakout by intensity and relationship to investigational medicinal product (IMP)
- Quantitative vital signs (systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, and body temperature) and laboratory tests (urine drug screen and urine pregnancy test) will be summarized, including observed and change from baseline.

3.3. Treatments

The study center will dispense a blinded treatment sequence (A/B or B/A) after successfully completing screening procedures. The overencapsulated tablets (IMP) for each treatment sequence will be dispensed using an Interactive Voice/Web Response System (IxRS) to maintain a balanced stratification amongst the study centers participating in the study.

4. General Statistical Considerations

All statistical analyses will be conducted using SAS Version 9.3 or higher (SAS Institute, Cary, North Carolina).

All statistical tests will be two-sided and will be performed at the 5% level of significance leading to 95% (2-sided) confidence intervals (CIs), unless otherwise stated. P-values will be rounded to 3 decimal places. If a p-value is less than 0.001 it will be reported as "<0.001." If a p-value is greater than 0.999 it will be reported as ">0.999."

Continuous data will be summarized by treatment group using descriptive statistics (n, mean, median, standard deviation [SD], minimum, and maximum). Categorical data will be summarized by treatment group using frequency tables (frequencies and percentages). Mean and median will be presented to one decimal place beyond the precision with which the data was captured. SD will be presented to two decimal places beyond the precision with which the data was captured. Minimum and maximum will be presented to the precision with which the data was captured. When count data are presented, the percentage will be suppressed when the count is zero in order to draw attention to the non-zero counts. A row denoted "Missing" will be included in count tabulations where specified on the shells to account for dropouts and missing values. The denominator for all percentages will be the number of patients in that treatment group within the analysis set of interest, unless otherwise specified.

All study-related raw data that support the corresponding tables and figures will be presented in data listings. Additional data listings may be generated as needed. All table, listing and figure (TLF) shells will appear in landscape format employing Courier New 9 point font. Unless otherwise noted, all tables will summarize patient results by treatment group sorted in the following order: Naproxen Sodium, Acetaminophen. As this is a crossover study, we expect that individual patients will be included in both treatment groups. Unless otherwise noted, all data listings will be sorted by treatment group and patient identification number defined as investigator identification number concatenated with the subject number.

The baseline value for an assessment will be defined as the last non-missing measurement including unscheduled assessments before or on the day of treatment with IMP for each treatment period. Change from baseline is defined as the post-baseline value minus the baseline value for the given assessment.

4.1. Sample Size

The proposed design of the study is to detect the treatment difference of 5.9 in TOTPAR0-12 with 90% power and a type I error of 0.05. The required sample size to meet these

design criteria are 154 patients in the PP set. Assuming a 23% drop out rate, approximately 200 patients will need to be randomized.

4.2. Randomization and Blinding

Patients will be randomized in a 1:1 fashion into one of two blinded treatment sequences:

- Naproxen sodium 440 mg (two tablets of naproxen sodium 220 mg) for Treatment Period 1 and acetaminophen 1000 mg (two tablets of acetaminophen 500 mg) for Treatment Period 2 (treatment sequence A/B);
- Acetaminophen 1000 mg (two tablets of acetaminophen 500 mg) for Treatment Period 1 and naproxen sodium 440 mg (two tablets of naproxen sodium 220 mg) for Treatment Period 2 (treatment sequence B/A).

Patients enrolled in the study, investigators and their staff involved in protocol procedures or data collection analysis will be blinded to the identity of the treatment sequence. The study monitor will conduct product accountability during and after database lock. To preserve the blinding, all investigational products will be over encapsulated and prepackaged according to the randomization schedule and managed using IV/WRS.

4.3. Analysis Set

Three populations will be identified in this study. Patients' inclusion status in each of the analysis sets will be presented in a data listing.

4.3.1. Intent-to-Treat (ITT)

The ITT will consist of all patients who were randomized and provided at least one measure of an efficacy parameter after the first dose of IMP. ITT set will be used as the sensitivity analysis for selected parameters. All analyses using the ITT will group participants according to randomized treatment.

4.3.2. Per Protocol (PP)

The PP will consist of all patients in ITT who do not have major protocol violations. The PP set will be used as the primary analysis for the efficacy parameters. Protocol deviations, including assessment of significance, will be identified prior to database lock and study unblinding. All analyses using the PP will group participants according to treatment sequence actually received. The following incidences will be considered as a protocol violation and excluded from the PP analysis if any one of the following occurred during the treatment period:

- Patient vomited within 30 minutes of ingesting the study medication;

- Patient took rescue medication before the 2-hour evaluation;
- Patient does not have a baseline (predose) pain evaluation;
- Patient dosed with study medication despite having a NRS <5;
- Patient took a prohibited concomitant medication preceding or during the study Treatment Period;
- Patient did not start menstrual flow within 48 hours of dosing;
- Patient failed to continue to meet the inclusion/exclusion criteria prior to the Treatment Period;
- Patient had more than 50% evaluations that were missed during the 0 through 12-hour evaluation period in those who did not use rescue medication;
- Patient missed more than 2 evaluations from hours 6-12 in those who did not use rescue medication.
- Patient dosed more than 30 minutes before or more than 2 hours after the predose pain evaluation

Patients may be invalid for one period but valid for another period. Therefore, invalidity for either period does not make them invalid for both. Protocol violations which lead to exclusion from the per-protocol set will be summarized and listed.

4.3.3. Safety

The Safety set is comprised of all patients who received at least one dose of IMP. The Safety set is usually used to tabulate all of the safety information for a study, such as adverse events, treatment compliance, laboratory results, and vital signs.

4.4. Assessment Windows

4.4.1. Treatment Periods

For classifying AEs, the treatment periods will be defined as:

- Treatment Period 1: The onset date of the AE is on or after the start date of study medication in treatment period 1 but before the start date of the study medication in treatment period 2.
- Treatment Period 2: The onset date of the AE is on or after the start date of study medication in treatment period 2.

Partial dates will be imputed for AE start dates in order to determine the treatment period to which the event will be assigned.

- If the onset date is completely missing, onset date is set to the date of first dose and will be considered TEAE for both treatment periods.
- If the year is present and the month is missing, then the month is set to January. If the year is same as the year of the date of first dose and the AE end date is not prior to the date of first dose, then onset date is set to the date of first dose.
- If the month and year are present and the day is missing, then the day is set to the 1st day of month. If the month and year are same as the month and year of the date of first dose and the AE end date is not prior to the date of first dose, then onset date is set to the date of first dose.
- If the AE end date is present, then the imputed start date will be no later than the end date.

Partial dates will be imputed for medication start dates in order to determine if a medication is prior or concomitant.

- Partial start dates of prior and concomitant medications will be assumed to be the earliest possible date consistent with the partial date.
- If the year is present and the month is missing, then the month is set to January.
- If the month and year are present and the day is missing, then the day is set to the 1st day of month.
- If the year is missing, then the year will be assumed to be the year part of the patient's informed consent date.

Partial or missing medication stop dates for medications that are not ongoing will be handled as follows:

- Partial stop dates of prior and concomitant medications will be assumed to be the latest possible date consistent with the partial date.
- If the year is present and the month is missing, then the month is set to December.
- If the month and year are present and the day is missing, then the day is set to the last day of month.
- For patients who are treated, if the year is missing and the month or day are not missing, then the year will be assumed to be the year part of the patient's last recorded study visit date.
- For patients who are not treated, if the year is missing and the month or day is not missing, then the year will be assumed to be the year part of the patient's discontinuation date.
- If the complete stop date is missing, the stop date will be considered to be either the patient's last recorded visit date for treated patients or the patient's discontinuation date for non-treated patients.

5. Subject Disposition

5.1. Study Sets

The number of patients who were enrolled and the number of patients within each analysis set (ITT, PP, and Safety) will be summarized for each treatment group for all randomized patients. A by-patient listing indicating the patient's inclusion in each analysis set and reason(s) that the patient is excluded from an analysis set will be presented.

5.2. Disposition

The counts and percentages of patients who complete or discontinue from the study treatment will be presented based on the number of patients in each treatment group and overall for all patients. The count and percentage of patients who complete or discontinue from the study will be summarized in a similar manner. Reasons for discontinuation of treatment for patients who do not complete the study treatment and reasons for not completing the study will be summarized for each treatment group as Treatment Disposition and Patient Disposition, respectively. Unless noted all percentages will be based on the number of patients screened. Patient disposition data will be listed as well.

5.3. Protocol Deviations and Violations

Protocol deviations will be tracked by the clinical team on an on-going basis. Specific criteria for what constitutes a significant protocol deviation will be determined by the clinical team. A blinded review of the deviation log collected by the clinical group will be conducted prior to database lock. All deviations will be listed.

Protocol violations, which disqualify a patient from the per-protocol set, will also be tracked on an on-going basis. These include but are not necessarily limited to the events listed in Section 4.3.2. This section also describes analyses related to this data..

6. Demographics and Baseline Characteristics

6.1. Demographics

Demographic information and baseline characteristics collected at screening will be summarized for the ITT, PP, and Safety sets. Continuous variables, including age (years), baseline weight (kg), baseline height (cm), and baseline body mass index (BMI) (kg/m^2) will be summarized using descriptive statistics for each treatment sequence and overall. Categorical variables including sex (Male, Female), ethnicity (Hispanic or Latino, Not Hispanic or Latino), race (White, Black or African American, Asian, American Indian or

Alaska Native, Native Hawaiian or Other Pacific Islander, and Other), native language (English, Spanish, Chinese, Tagalog, Vietnamese, Arabic, French, Korean, Russian, German, Other), and present method of birth control (Abstinence and Double Barrier Method, Double Barrier Method, Permanent Sterilization, Oral Contraceptives) will be summarized by reporting the number and percentage of patients in each category for each treatment group and overall. Percentages will be based on the total number of patients in the ITT set. All demographic and baseline characteristics will be listed for all patients in the ITT set.

Age will be calculated as (date of Screening visit – date of birth)/365.25.

BMI is calculated as (body weight in kilograms) / (height in meters)^2.

6.2. Alcohol, Tobacco and Drug Usage

Descriptive statistics will be provided for subject alcohol history (<3 drinks per day, ≥3 drinks per day), tobacco usage (Current, Previous, Never), and whether drug use history was discussed with the subject. Percentages will be based on the total number of patients in the safety set..

6.3. Medical History

6.3.1. General Medical History

Medical history will be collected in the Electronic Case Report Form (eCRF) at the screening and baseline visits and verbatim terms will be coded classified on the basis of Medical Dictionary for Regulatory Activities (MedDRA) terminology, which will be updated whenever possible throughout the life of the study. Medical history will be summarized by the number and percentage of patients with any medical history reported by coded system organ class, preferred term, and treatment group for the safety set. A subject listing will also be included with start date and end date or ongoing status.

6.4. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria deviations noted in the eCRF will be presented for the set of randomized patients in a data listing.

7. Treatments and Medications

7.1. Prior and Concomitant Medications

Any medicinal product, prescribed or over-the-counter (OTC), including herbal products, vitamins, and minerals is considered a concomitant medication. Any medications used during the study will be coded with the World Health Organization Drug Dictionary (WHODD), which will be updated whenever available throughout the life of the study.

Prior/concomitant medication use will be recorded in the eCRF beginning at Screening visit before the first dose of study medication and will be recorded at the subsequent visits as well. Any changes in prior or concomitant medications will also be recorded in the subject's eCRF.

Summary of prior and concomitant medications by treatment group and preferred term will be provided. All prior and concomitant medications will be presented in data listings. Prior and concomitant medications will be summarized for the safety set.

7.2. Study Treatments

In each treatment period, the patient will dose with either naproxen sodium 440mg (220mg x 2 tablets) or acetaminophen 1000mg (500mg x 2 tablets). Each treatment phase of the study is single-dose

7.2.1. Exposure and Treatment Compliance

The distribution of the study medication will be supervised by a member of the Investigator's team. Patients will self-administer IMP at their discretion and document the dosing in the e-diary system. The study center will monitor used and unused IMP for compliance. Any discrepancies between actual and expected amount of returned study medication must be discussed with the patient at the time of the visit, and any explanation must be documented in the source records. Total exposure and percent compliance will be summarized for the Safety and ITT sets, and a by-patient listing will be included on the Safety set. A listing of patients who took rescue medication will also be included.

7.2.2. Treatment Review

A treatment review listing will be provided for each treatment in assigned sequence, including the following information: date/time of last menses, reason for not taking study medication, alcohol use during treatment period, caffeine use during treatment period, and vomiting within 30 minutes of dosing.

8. Efficacy Analysis

Efficacy data will be analyzed utilizing the PP set unless noted otherwise. Data imputation may occur for the following:

If a patient takes rescue medication during the 12 hour assessment period, the subsequent pain intensity scores (NRS) following the intake of the rescue medication will be imputed by either the baseline score (pre-dose) or the score recorded immediately prior to taking rescue medication, whichever is worse. Subsequent pain relief scores following the intake of the rescue medication will be imputed by "none relief" (0).

For those patients who are not rescued during the 12 hour assessment period, the 30 and 60 minute assessments will be considered as missing in the event that pain intensity assessment is off-schedule by more than 10 minutes. Subsequent to the 1 hour reading values will be considered as missing in the event that the pain intensity assessment is off-schedule by more than 15 minutes. Any missing values during 12 hour assessment period will be imputed using linear interpolation (trapezoid rule) or extrapolation (last observation carried forward) approach depending on availability of data surrounding the missing value. If a missing value occurs at the 30 minute assessment, a value of "none relief" (0) will be imputed.

8.1. Primary Efficacy Endpoint

The primary efficacy endpoint is total pain relief scores (TOTPAR) over 12 hours. It will be calculated by multiplying the pain relief score at each post-dose time point by the duration (in hours) since the preceding time point and then summing these values.

$$(1)\ TOTPAR0-12 \quad = \sum_{i=0.5}^{12} f_i\,(t_i - t_{i-1})$$

Where f_i is the pain relief score at post-baseline time *i*, and ($t_i - t_{i-1}$) represents the duration in hours since the last timepoint, with t_0 representing the baseline timepoint. The pain relief score is a categorical score from 0-4 measuring relief from starting menstrual cramp pain where 0=none, 1=a little, 2=some, 3= a lot, 4= complete. The patient completes this at 30 minutes, 60 minutes, and hourly from 2-12 hours post-dose. All post-dose time point assessments have an allowable window of ±10 minutes for the 30 and 60 minute assessments, and a window of ±15 minutes for each subsequent timepoint.

8.1.1. Primary Analysis

The primary variable of TOTPAR0-12, as defined in (1) above, will be analyzed using a mixed analysis of covariance (ANCOVA) model with fixed effects for baseline TOTPAR0-12, sequence, treatment and period and will include a random effect for subject within sequence for the crossover study design. If 'sequence' is not significant in the initial model, then a reduced model with 'sequence' removed from the fixed effects will be used. The estimated treatment difference between the two treatment groups will be presented together from the reduced model with the two-sided 95% confidence interval for the difference and the p-value. The study will be considered positive if the p-value for the treatment difference is less than 0.05. However, if 'sequence' is significant (p-value ≤ .10) in the initial model, then further exploratory analyses will be conducted to investigate any potential bias due to carry-over effects.

Example SAS code for the primary analysis (based on Yarandi, 2004)

```
PROC MIXED DATA=DATA1;
CLASS SEQUENCE SUBJECT PERIOD TRTA;
MODEL TOTPAR= SEQUENCE BASELINE TRTA PERIOD;
RANDOM SUBJECT (SEQUENCE);
LSMEANS TREATMENT / PDIFF CL E;
RUN;
```

where *TOTPAR* is TOTPAR0-12 as defined as in the equation above, TRTA is actual treatment corresponding to the TOTPAR0-12, BASELINE is the baseline NRS score for that period, SEQUENCE is the treatment sequence a patient was randomized to (B/A or A/B), and PERIOD is the treatment period (either 1 or 2).

8.1.2. Sensitivity Analysis

In addition to the primary analysis, the same ANCOVA analysis will be performed on TOTPAR0-12 for the ITT set. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value

8.2. Secondary Efficacy Endpoints

8.2.1. Summed Pain Intensity Difference (SPID) over the 12-hour study period (SPID0-12)

The NRS is a numeric scale from 0-10 which measures pain at Baseline (predose), 30 minutes, 60 minutes, and hourly from 2-12 hours post-dose. All post-dose time point assessments have an allowable window of ±10 minutes for the 30 and 60 minute

assessments, and a window of ±15 minutes for each subsequent timepoint. The baseline dose must be taken no more than 30 minutes before or 2 hours after the predose measurement.

Pain intensity differences (PID) will be derived by subtracting the pain intensity (using NRS 0-10) at the baseline time point from the post dose intensity score (baseline score – post-baseline score). A positive difference is indicative of improvement. Summed Pain Intensity Difference will be calculated as follows:

$$(2)\ SPID0-12 = \sum_{i=0.5}^{12} (f_0 - f_i)(t_i - t_{i-1})$$

Where f_i is the NRS score at post-baseline time i, f_0 represents the baseline NRS score, and $(t_i - t_{i-1})$ represents the duration in hours since the last timepoint, with t_0 representing the baseline timepoint.

The SPID0-12 as defined in equation (2) above, will be analyzed using an ANCOVA model similar to the primary analysis. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value.

8.2.2. SPID over 0-6 hours (SPID0-6)

Summed Pain Intensity Difference from 0-6 hours will be calculated as follows:

$$(3)\ SPID0-6 = \sum_{i=0.5}^{6} (f_0 - f_i)(t_i - t_{i-1})$$

Where f_i is the NRS score at post-baseline time i, f_0 represents the baseline NRS score, and$(t_i - t_{i-1})$ represents the duration in hours since the last timepoint, with t_0 representing the baseline timepoint.

The SPID0-6 as defined in equation (3) above, will be analyzed using an ANCOVA model similar to the primary analysis. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value.

8.2.3. SPID 6-12 hours (SPID6-12)

Summed Pain Intensity Difference from 6-12 hours will be calculated as follows:

$$(4)\ SPID6-12 = \sum_{i=6}^{12} (f_0 - f_i)(t_i - t_{i-1})$$

Where f_i is the NRS score at post-baseline time i, f_0 represents the baseline NRS score, and$(t_i - t_{i-1})$ represents the duration in hours since the last timepoint, with t_0 representing the baseline timepoint.

The SPID6-12 as defined in equation (3) above, will be analyzed using an ANCOVA model similar to the primary analysis. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value.

8.2.4. TOTPAR over 0-6 hours (TOTPAR0-6)

Total Pain Relief Scores from 0-6 hours will be calculated as follows:

$$(5)\ TOTPAR0-6 = \sum_{i=0.5}^{6} f_i\,(t_i - t_{i-1})$$

Where f_i is the pain relief score at post-baseline time i, and ($t_i - t_{i-1}$) represents the duration since the last timepoint, with t_0 representing the baseline timepoint.

The TOTPAR0-6 as defined in equation (5) above, will be analyzed using an ANCOVA model similar to the primary analysis.. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value.

8.2.5. TOTPAR 6-12 hours (TOTPAR6-12)

Total Pain Relief Scores from 6-12 hours will be calculated as follows:

$$(6)\ TOTPAR6-12 = \sum_{i=6}^{12} f_i\,(t_i - t_{i-1})$$

Where f_i is the pain relief score at post-baseline time i, and ($t_i - t_{i-1}$) represents the duration since the last timepoint, with t_0 representing the baseline timepoint.

The TOTPAR6-12 as defined in equation (6) above, will be analyzed using an ANCOVA model similar to the primary analysis. The estimated treatment difference between the two treatment groups will be presented together with the two-sided 95% confidence interval for the difference and the p-value.

8.2.6. Time to First Intake of Rescue Medication

The time to first intake of rescue medication is calculated as the number of hours between the time of first dose of study medication and the time of rescue medication. Patients who take rescue medication in the first 2 hours will be assessed for ITT analysis set only. If a patient does not take a rescue medication during the treatment period, she will be censored at the time of last assessment. Patients who take rescue medication prior to IMP will be considered a “failure” at the time of IMP dose.

Kaplan-Meier (KM) estimates of rescue medication use and their 95% CIs will be presented at each post-dose timepoint by treatment group. The log-rank test will be used to test the difference in KM curves between the two treatments, with p-values presented. The median time to first intake of rescue medication and 95% CI for each treatment group will also be summarized. The plots will be based on hours to rescue medication. Pain Intensity Difference (PID) Scores at Each Evaluation

Pain Intensity Difference (PID) scores will be evaluated at each post-dose timepoint. Pain Intensity Difference scores will be summarized descriptively by treatment group and plotted over time.

8.2.7. Pain Relief Scores at Each Evaluation

Pain Relief scores will also be evaluated at each post-dose timepoint. Pain Relief scores will be summarized numerically and categorically by treatment group; scores will also be plotted over time.

8.2.8. Global Evaluation at 12 Hours Post-Dose or Immediately Before First Intake of Rescue Medication

The Global Evaluation is a question that rates the IMP as a pain reliever either at 12 hours post-dose or immediately at the first intake of rescue medication.

Overall, I would rate the effectiveness of this study medication in relieving my menstrual cramp pain as…
• 0 = poor
• 1 = fair
• 2 = good
• 3 = very good
• 4 = excellent

Results for this variable will be summarized descriptively and analyzed using the Cochran-Mantel-Haenszel (CMH) method with modified ridit scores. P-values will be displayed.

Example SAS code for the this analysis

```
PROC FREQ DATA=DATA1;
TABLES TRTA*GESCORE/CMH SCORES=MODRIDIT;
RUN;
```

9. Safety Analysis

Safety parameters include monitoring of TEAEs, physical examination findings, vital sign measurements, and laboratory examinations. All safety summaries and analyses will be summarized by actual treatment group within the Safety set.

9.1. Adverse Events

All AEs that occur during the study will be recorded and classified on the basis of Medical Dictionary for Regulatory Activities (MedDRA) terminology version 20.1. Treatment-emergent AEs (TEAEs) are defined as AEs with an onset on or after the date of the first IMP dose. Adverse events noted prior to the first IMP administration that worsen after baseline will also be reported as TEAEs and included in the summaries

All information pertaining to an AE noted during the study will be listed by subject, detailing verbatim term, preferred term, system organ class (SOC), onset date, resolution date, intensity, seriousness, action taken, outcome, drug relatedness, and relatedness to protocol-required procedures. The event onset will also be shown relative to IMP initiation (in number of days). Listings will be provided for all TEAEs.

TEAEs will be presented on the Safety set and displayed by treatment period and group unless noted.

9.1.1. Incidence of TEAEs

Adverse events will be deemed treatment emergent if the onset date is on or after the date of first treatment, Day 1 in each period. Events with onset date during the screening period are not collected for this study. An overall summary of incidence of TEAEs (patient/event counts) including drug relation, serious adverse events (SAE), TEAEs leading to drug discontinuation and fatal AEs will be presented.

The number and percentage of patients with at least one TEAE and the number of events overall will be presented by SOC and PT.

9.1.2. Relationship of Adverse Events to IMP

A summary of TEAEs by relationship to IMP will be presented in a table by total number of TEAE and incidence of occurrence. The investigator will provide an assessment of the relationship of the event to the IMP as Related or Not Related. If a subject reports multiple occurrences of the same TEAE, only the most closely related occurrence will be presented in the incidence count. Treatment-emergent AEs that are missing a relationship will be presented in the summary table as "Related" but will be presented in the data listing with a missing relationship. Percentages will be calculated based on the number of patients in the safety set.

The TEAE data will be categorized and presented by SOC, PT, and relationship in a manner similar to that described in Section 9.1.1.

9.1.3. Intensity of Adverse Event

A summary of TEAEs by intensity will be presented in a table by total number of TEAE and incidence of occurrence. The intensity that will be presented represents the most extreme intensity captured on the eCRF page. The possible intensities are "Mild," "Moderate," and "Severe." In the TEAE intensity table, if a subject reported multiple occurrences of the same TEAE, only the most severe will be presented in the incidence count. Treatment-emergent AEs that are missing intensity will be presented in tables as "Severe" but will be presented in the data listing with a missing intensity. Percentages will be calculated based on the number of patients in the safety set.

The TEAE data will be categorized and presented by SOC, PT, and intensity in a manner similar to that described in Section 9.1.1.

9.1.4. Serious Adverse Events

The treatment-emergent SAE data will be categorized and presented by SOC and PT in a manner similar to that described in Section 9.1.1. At each level of subject summarization, a subject is counted once for the incidence if the subject reported one or more events.

9.1.5. Deaths, Adverse Events Leading to Treatment Discontinuation/Study Discontinuation

A listing of all Serious TEAEs (including deaths) and/or AEs leading to treatment or study discontinuation will be presented. Additionally, a separate listing of any deaths will be presented. Deaths and discontinuations will be assessed based on data recorded in the AE eCRF page. Subject deaths will be identified as AEs where the outcome is "Fatal". Adverse events leading to treatment discontinuation will be identified as AEs where the action taken with IMP is "Drug Withdrawn". Adverse events leading to study discontinuation will be identified as AEs where the caused study discontinuation field is marked as "Yes".

9.2. Clinical Laboratory Evaluations

9.2.1. Urine Drug Screen

A screen for drugs of abuse will be performed at Screening (Visit 1) and Visit 2. A by-patient list of urine drug screen will be presented. Patients will be tested for the following drugs: Methamphetamines, Amphetamines, cannabinoids, cocaine, opiates, benzodiazepines, barbiturates, Methylenedioxymethamphetamine, methadone, Oxycodone, Adulterants.

9.2.2. Urine Pregnancy Test

At screening, and prior to dosing all patients must perform a urine pregnancy test and verify that the results are negative (not pregnant). Any patient who has a positive pregnancy test or is not sure of the results must contact the study center immediately. Patients with a verified positive pregnancy test will be withdrawn from the study. Pregnancy test results will be listed by patient.

9.3. Vital Sign Measurements

The vital sign summary and analysis will be based on the recordings of the variables blood pressure (systolic and diastolic [mm Hg]), oral body temperature (F), heart rate (beats/min), and respiratory rate (breaths/min). Notable abnormal values will be flagged in the data listings. The conversion between conventional unit to SI unit for temperature and weight and definitions for notable abnormal values are listed in the appendix (Section 12.1**Error! Reference source not found.**, Table 1). Observed values at Screening and Visits 1 and 2 and change from Baseline values to Visits 1 and 2 for vital sign measurements will be summarized descriptively by treatment group.

All summaries will be done for the Safety set and all vital sign data will be listed.

9.4. Physical Examination

The physical examination (by means of inspection, palpation, auscultation) will be performed by a physician at the study site covering at least the organs of the cardiovascular, respiratory, and abdominal system.

Abnormal physical examination findings are recorded either as medical history or as adverse events. Physical examination results will be listed.

10. Interim Analysis

There is no interim analysis scheduled for this study.

11. References

Kenward MG, Roger JH. The use of baseline covariates in crossover studies. *Biostatistics*, Jan 2010; 1-17.

Choi K, Hong T, Lee J. On comparison of SAS Codes with GLM and MIXED for the crossover studies with QT interval data. Transl Clin Pharamcol, 2014; 78-82.

Yarandi, HN. Paper SD04 – Crossover Designs and Proc Mixed in SAS. http://analytics.ncsu.edu/sesug/2004/SD04-Yarandi.pdf

12. Appendices

12.1. Notable Abnormal Vital Signs

Table 1 Notable Abnormal Vital Signs

Parameter	Actual Value	Change from Baseline
Pulse	< 40 bpm > 120 bpm	>= 15 bpm increase from baseline >= 15 bpm decrease from baseline
Systolic blood pressure	< 80 mmHg > 180 mmHg	>= 20 mmHg increase from baseline >= 20 mmHg decrease from baseline
Diastolic blood pressure	< 50 mmHg > 105 mmHg	>= 15 mmHg increase from baseline >= 15 mmHg decrease from baseline
Temperature	NA	>= 2 F increase from baseline >= 2 F decrease from baseline

12.2. Table of Contents: Summary Tables, Listings and Figures

Type	Table Number	TLF Title
Table	14.1-1	Summary of Patient Disposition *All Patients*
Table	14.1-2	Summary of Analysis Sets *All Randomized Patients*
Table	14.1-3	Summary of Exclusions from Per-Protocol Set *All Randomized Patients*
Table	14.1-4.1	Summary of Demographics and Baseline Characteristics *Intent-to-Treat*
Table	14.1-4.2	Summary of Demographics and Baseline Characteristics *Per-Protocol*
Table	14.1-4.3	Summary of Demographics and Baseline Characteristics *Safety Set*
Table	14.1-5	Summary of Medical History *Safety Set*
Table	14.1-6	Summary of History of Substance Use *Safety Set*
Table	14.1-7	Summary of Prior Medications *Safety Set*
Table	14.2-1.1	Summary of Analysis of Sum of Total Pain Relief over 0-12 hours (TOTPAR0-12) *Per-Protocol*

Table	14.2-10.2	Summary of Global Evaluation Scores *Intent-to-Treat*
Table	14.3-1	Summary of Treatment Exposure and Compliance *Safety Set*
Table	14.3-2	Summary of Treatment Exposure and Compliance *Intent-to-Treat*
Table	14.3-3	Summary of Concomitant Medications *Safety Set*
Table	14.3-4	Summary of Change from Baseline in Vital Signs *Safety Set*
Table	14.3.1-1	Overall Summary of Treatment Emergent Adverse Events *Safety Set*
Table	14.3.1-2	Summary of Frequent Treatment Emergent Adverse Events with 2% or Greater Incidence *Safety Set*
Table	14.3.1-3	Summary of Treatment Emergent Adverse Events by System Organ Class and Preferred Term *Safety Set*
Table	14.3.1-4	Summary of Treatment Emergent Adverse Events by Intensity, System Organ Class and Preferred Term *Safety Set*
Table	14.3.1-5	Summary of Treatment Emergent Adverse Events by Relationship to IMP, System Organ Class and Preferred Term *Safety Set*
Table	14.3.1-6	Summary of Serious Treatment Emergent Adverse Events by System Organ Class and Preferred Term *Safety Set*
Listing	16.1-7	Listing of Randomized Allocation to Treatment *All Randomized Patients*
Listing	16.2-1.1	Listing of Study Completion *All Patients*
Listing	16.2-2.1	Listing of Protocol Deviations *All Randomized Patients*
Listing	16.2-2.2	Listing of Patient Populations and Reasons Excluded from Analysis Sets *All Randomized Patients*
Listing	16.2-4.1	Listing of Demographics and Baseline Characteristics *Safety Set*
Listing	16.2-4.2	Listing of Inclusion and Exclusion Criteria *All Patients*
Listing	16.2-4.3	Listing of Prior and Concomitant Medications *Safety Set*
Listing	16.2-4.4	Listing of Medical History *Safety Set*
Listing	16.2-5.1	Listing of IMP Administration Patient Diary *Safety Set*
Listing	16.2-5.2	Listing of Rescue Medication Administration *Safety Set*
Listing	16.2-5.3	Listing of Treatment Review *Safety Set*
Listing	16.2-7.1	Listing of Treatment-Emergent Adverse Events *Safety Set*
Listing	16.2-7.2	Listing of Serious Adverse Events *Safety Set*
Listing	16.2-7.3	Listing of Adverse Events Leading to Treatment or Study Discontinuation *Safety Set*

Listing	16.2-7.4	Listing of Deaths *Safety Set*
Listing	16.2-8.1	Listing of Laboratory Results - Urine Drug Screen *Safety Set*
Listing	16.2-8.2	Listing of Laboratory Results - Pregnancy Testing Results *Safety Set*
Listing	16.2-8.3	Listing of Physical Exam Results at Screening *Safety Set*
Listing	16.2-9.1	Listing of Vital Sign Results *Safety Set*
Listing	16.2-10.1	Listing of Time to Rescue Medication Safety Set
Listing	16.2-10.2	Listing of NRS and PID Scores *Safety Set*
Listing	16.2-10.3	Listing of Pain Relief Scores *Safety Set*
Listing	16.2-10.4	Listing of Global Evaluation Scores *Safety Set*
Figure	14.2-1.1	Kaplan-Meier Plot of Time to Rescue Medication by Treatment *Per-Protocol*
Figure	14.2-1.2	Kaplan-Meier Plot of Time to Rescue Medication by Treatment *Intent-to-Treat*
Figure	14.2-2.1	Line Graph of Average Pain Relief Scores over Time by Treatment *Per-Protocol*
Figure	14.2-2.2	Line Graph of Average Pain Relief Scores over Time by Treatment *Intent-to-Treat*
Figure	14.2-3.1	Line Graph of Average NRS Scores over Time by Treatment *Per-Protocol*
Figure	14.2-3.2	Line Graph of Average NRS Scores over Time by Treatment *Intent-to-Treat*
Figure	14.2-4.1	Line Graph of Average Pain Intensity Difference Scores over Time by Treatment *Per-Protocol*
Figure	14.2-4.2	Line Graph of Average Pain Intensity Difference Scores over Time by Treatment *Intent-to-Treat*

12.3. Schedule of Study Procedures

Study Procedure	Screening Phase	Treatment Phase			
	Screening Visit 1	Treatment Period 1	Visit 2	Treatment Period 2	Visit 3 End of Study
	Days -21 to -1	After Screening	After Period 1	After Visit 2	After Period 2
Signed Informed Consent	X				
Inclusion/Exclusion Criteria Review	X				
Patient Demographics	X				
Medical History	X				
Prior and Concomitant / Medication History	X		X		X
History of drug, alcohol and tobacco use	X				
Body weight, height, and BMI	X				
Physical examination (general routine)	X				
Urine drug screen	X		X		
Vital signs (incl. temperature)[a]	X		X		X
Urine pregnancy test	X	X		X	
Categorical Pain Intensity 0-3 Scale	X				
e-diary distribution/training	X				
Randomization/Kit assignment	X[b]				
IMP administration		X[c]		X[c]	
0-10 Point Pain Intensity NRS		X		X	
0-4 Categorical – pain relief		X		X	
12 hour post-dose Global Evaluation[d]		X		X	
IMP compliance			X		X
e-diary/treatment kit return/review			X		X
Adverse events		X	X	X	X

[a] Vital signs (blood pressure, respiratory rate, heart rate and body temperature after 5 minutes of rest in a sitting position).
[b] Randomization to treatment occurs only for eligible patients.
[c] Urine pregnancy test must be performed and the results must be negative (not pregnant) before dosing.
[d] If the patient decides to take rescue medication prior to the 12 hours, the patient will be instructed to complete the Global Evaluation prior to taking the rescue medication.