PRODUCT: LUCEMYRA (Lofexidine)
STATISTICAL ANALYSIS PLAN PROTOCOL NUMBER: USWM-LX1-3003-1 / 02

SPONSOR:

USWM, LLC (dba US WorldMeds)
4441 Springdale Rd
Louisville, KY 40241

TITLE:

A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety, and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)

DOCUMENT DATE: 22JAN2015
NDA NUMBER: 209229
NCT NUMBER: NCT01863186

16.1.9 DOCUMENTATION OF STATISTICAL METHODS

Table 1: Statistical Analysis Plans

Statistical Analysis Plans Included in Appendix 16.1.9
16.1.9.1 Statistical Analysis Plan: Clinical Study Report- Version 1.0- 18 June 2014
16.1.9.2 Statistical Analysis Plan: Clinical Study Report- Addendum 1- 24 December 2014
16.1.9.3 Statistical Analysis Plan: Clinical Study Report- Addendum 2- 22 January 2015

Protocol Number: USWM-LX1-3003-1
Statistical Analysis Plan
Final Version: 1.0
Issue Date:18-JUN-14

STATISTICAL ANALYSIS PLAN

USWM-LX1-3003-1

A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)

Protocol Ref: USWM-LX1-3003-1

Final Version: 1.0

Date effective: 18-JUN-14

SUMMARY OF SECTIONS

1. INTRODUCTION

This document details the statistical analysis of the data that will be performed by for the US WorldMeds, LLC study USWM-LX1-3003-1.

The proposed analysis is based on the contents of the Final Version of the protocol (dated 10-Aug-2012) and amendments 1 and 2 (dated 08-Mar-13 and 06-May-13 respectively).

The Clinical Study Report will be written by following the guidelines in the ICH E3 document.

The primary objective of this study is to investigate the efficacy, safety, and dose-response of lofexidine (2.4 mg or 3.2 mg per day) in reducing withdrawal signs and symptoms and facilitating completion of detoxification/extending treatment retention in subjects undergoing detoxification from short-acting opioids in a double-blind inpatient setting (Days 1-7) followed by an open-label inpatient/outpatient setting (Days 8-14).

It is hypothesized that a daily dose of either 2.4 mg or 3.2 mg lofexidine will achieve greater efficacy than placebo with respect to overall symptom relief over the first 7 days of opioid withdrawal (primary endpoint) and will increase the likelihood of subjects completing Days 1-7 of treatment (secondary endpoint). Further, safety measures in the 2.4 mg and 3.2 mg total daily dose groups will be compared descriptively to assess whether the lower dose results in fewer and less severe adverse events than does the higher dose.

2. SAMPLE SIZE

Treatment effect and subject variability with respect to SOWS-Gossop scores were estimated from the prior Phase 3 study (USWM-LX1-3002), using the random coefficients model planned for the present study and estimating the treatment effect of 3.2 mg lofexidine versus placebo with respect to area under the curve (AUC) based on the SOWS-Gossop scores from Days 1 through 7 (AUC(1-7)). It was assumed that the treatment effect of the 2.4 mg lofexidine dose will be three-fourths the treatment effect of the 3.2 mg dose. The table below shows the power to find a significant treatment effect for the comparisons of the two lofexidine treatments versus placebo, assuming a sample size allocation ratio of 3:3:2 (3.2 mg lofexidine : 2.4 mg lofexidine : placebo) and accounting for the sequential testing approach described in Section 6.1. The power to find a statistically significant effect of the 3.2 mg lofexidine dose with respect to AUC (1-7) is in excess of 90% with the planned total sample size of 600.

Total Sample Size	Power (%) with Respect to AUC (1-7)	
	3.2 mg Lofexidine versus Placebo	2.4 mg Lofexidine versus Placebo
600	94.6	72.6
550	92.6	67.6
500	90.1	62.2
450	86.8	55.7

3. RANDOMIZATION

A prospective subject who meets all of the study inclusion criteria and does not meet any of the exclusion criteria may be randomized into the study. The final eligibility criterion is a score of at least 2 on the OOWS-Handelsman at Baseline.

A stratified randomization procedure will be used to separately allocate male and female subjects in 1 of the 3 treatment groups: 2.4 mg lofexidine, 3.2 mg lofexidine, or placebo.

Randomization will be implemented centrally, that is, take place across all investigational sites using an Interactive Web Response System (IWRS) managed by Once eligibility criteria are confirmed at the Baseline visit, site personnel will access the IWRS and complete randomization procedures. The system will assign a unique kit number for drug dispensation according to the randomization scheme. The subject number assigned at screening should be used throughout the study in all source documents and eCRFs.

4. INTERIM ANALYSIS

No interim analysis is planned.

5. STATISTICAL METHODS

5.1 Continuous

Continuous variables and ordered categorical variables not subject to censoring will be summarized with the number of non-missing observations, mean, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum. Continuous variables subject to censoring (e.g., time to removal from study treatment) will be summarized by the number of subjects with data, 25th percentile, median, 75th percentile derived from Kaplan-Meier estimates of probabilities.

For summary statistics, means, medians and percentiles will be displayed to one more decimal place than the raw data, dispersion statistics (e.g. standard deviation) will have two more decimal places, and the minimum and maximum will be displayed to the same number of decimal places as reported in the raw data.

5.2 Categorical

Unordered categorical and ordered categorical (depending on the number of categories) variables will be presented in contingency tables with cell frequencies and percentages for the number of non-missing observations and frequencies for the number of missing observations apart from disposition of subjects, concomitant medications, measurement of treatment compliance and adverse events where percentages will be presented for the population.

6. ANALYSIS PLAN

6.1 General

The treatment groups will be labelled as 2.4 mg Lofexidine, 3.2 mg Lofexidine and Placebo within the statistical output. In addition the 2.4 mg Lofexidine and 3.2 mg Lofexidine doses are

combined to form a single Lofexidine treatment group for presentation of safety data. Subject data in Appendix 16.2 listings will be ordered by treatment group, site and subject.

All statistical tests for efficacy will be 2-sided at the α=0.05 significance level.

All comparisons between the treatments will be reported with 95% confidence intervals for the difference. P-values will be rounded to four decimal places.

Only scheduled post-baseline laboratory and vital signs values will be tabulated. Post-baseline repeat/unscheduled assessments will be disregarded, although they will be listed in the relevant appendices to the report; in particular all clinically significant values will be noted.

All calculations and figures will be produced using SAS Version 9.3[3] or higher.

All summaries and analyses documented below will be presented in the final integrated statistical/clinical report and tables that will be based on the E3 guidelines published by ICH. However, it is noted here that no analysis plan prepared in advance of the data can be absolutely definitive and so the final report may contain additional tables or statistical tests if warranted by the data obtained. The justification for any such additional analyses will be fully documented in the final report.

6.2 Derived Data

- Definition of baseline

For all variables, baseline is defined as the last non-missing value prior to the receiving the first dose of study medication.

All tabulations involving change from baseline data will only include subjects with cohort data i.e. with data at baseline and at follow-up. Change from baseline will be calculated as follow-up values minus baseline value.

- Incomplete dates

All incomplete dates will be included in the clinical database as they were entered in the eCRF. Thereafter for calculation purposes, the incomplete dates will be completed using pre-defined rules. If a day or month is not recorded it will be replaced by the first day of the month or January respectively, provided this does not contradict any other dates recorded. For missing adverse events and medications dates/times during the trial, the worst-case date will be used (e.g. the end of the month for a stop date and 23:59 for the stop time, the date/time of initial dose for start of AE i.e. all events with missing start dates will be assumed to be treatment emergent).

- Ambiguous values

In the case where a variable is recorded as “>x”, “≥x”, “<x” or “≤x”, then for analysis purposes a value of x will be taken. Where a range of values is quoted the midpoint of the range will be taken.

- Questionnaire Data

In the event that component items used to calculate total scores for questionnaire data e.g. SOWS Gossop score are missing then the total scores will also be set to missing.

- Site pooling for statistical analyses

For the purposes of statistical analyses models where site is included within the model as a fixed effect, sites randomizing < 6 subjects will be pooled from smallest to largest until the pooled site subsequently created has at least 6 subjects. In the event that the pooling of sites with less than 6 subjects does not result in a site with > 6 subjects, then that pooled site with < 6 subjects will also be pooled with the next largest site with > 6 subjects. To break ties if more than one site has the same number of subjects, the site identifier will be used as a secondary sort key.

6.3 Analysis populations

The **Enrolled Population** includes all subjects screened into the study irrespective of whether they received the study medication.

Three principal analysis populations are defined as follows:

- **Intent-to-treat (ITT)**, consisting of all randomized subjects. Subjects will be assigned for analysis according to the group to which they are randomized.
- **Modified intent-to-treat (mITT)**, consisting of all subjects in the ITT group who received at least one dose of study medication. Subjects will be assigned for analysis according to the group to which they are randomized.
- **Safety** consisting of all subjects who received at least one dose of study medication. Subjects will be assigned for analysis according to treatment received.

The principal analysis population for the analyses of demographics and baseline characteristics and efficacy will be the mITT population. Sensitivity analyses of the completion status endpoint will be carried out using the ITT population.

Safety summaries will be provided for the Safety population.

Data recorded on subjects who are in the ITT but not in the mITT or Safety populations will be included in data listings.

6.4 Protocol Deviations

All protocol deviations will be assessed and documented on a case-by-case basis prior to the database lock. A listing of protocol deviations will be provided within Appendix 16.2.

6.5 Data Summaries

The data will be summarized in tabular form by treatment group (2.4 mg Lofexidine, 3.2 mg Lofexidine and Placebo) and overall. In addition safety data will be summarized by treatment

group where the 2.4 mg Lofexidine and 3.2 mg Lofexidine doses are combined to form a single Lofexidine treatment group.

For physical examination, when calculating the percentage reporting each category, the "Not Done" category will not be included in the denominator.

Subject demographics and baseline characteristics will be presented for the mITT population.

All safety data and protocol deviations will be presented for the safety population.

Subject disposition will be summarized using the enrolled population.

All efficacy variables will be assessed using the modified ITT population.

Graphical presentations of the data will also be provided where appropriate.

6.6 Disposition of Subjects

The following will be summarized:

- The number of subjects who enrolled into the study,
- The number of enrolled subjects that failed to be randomized along with the reasons,
- The number of enrolled subjects randomized, and
- The number of randomized subjects that received at least one dose of study medication,
- The number of randomized subjects that complete the double-blind and open-label phases.

Reason for and timing of withdrawal will be summarized separately for the double-blind phase and the open-label phase. The reasons for withdrawal during the double-blind phase are classified as follows:

- Lack of efficacy (i.e. Intolerable withdrawal symptoms) AE related to the study drug
- AE related to study drug
- Other reasons
    - AE unrelated to study drug or withdrawal (e.g. Concomitant illness)
    - There is evidence of contraband drug use while participating in the study
    - The subject requires therapy with an exclusionary drug
    - Lack of compliance with protocol and/or unit procedures
    - Other (e.g. Subject withdrew consent due to personal reasons)

The reasons for withdrawal during the open-label phase are classified as follows:

- Lack of efficacy (i.e. Intolerable withdrawal symptoms) AE related to the study drug
- AE related to study drug
- Other reasons
    - AE unrelated to study drug or withdrawal (e.g. Concomitant illness)
    - The subject requires therapy with an exclusionary drug
    - Lack of compliance with protocol and/or unit procedures
    - The subject is Lost to Follow Up
    - Other (e.g. Subject withdrew consent due to personal reasons)

For those subjects entering the open-label phase, the following will be tabulated

- Which setting was the subject seen (inpatient, outpatient)
- Primary reason for inpatient- setting
    - PI Clinical Discretion - Not AE
    - Standard of Care
    - Treatment of AE (but subject remains on study medication)
    - Indigent
    - Other
- Primary reason for outpatient- setting
    - Subject Preference
    - Appropriate per PI Clinical Discretion
    - Standard of Care
    - Space Limitations at Site (Bed Availability)
    - Other

6.7 Baseline Comparability

6.7.1 Study Population

Subject demographics and baseline characteristics will be presented for all subjects within the mITT population.

6.7.2 Variables Considered

Standard continuous or categorical variable summaries will be presented for the following variables:

Demography

- Site number
- Age at screening visit as recorded on the study database (years)
- Gender (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Other). If more than one race is selected than categories will be created for the mixed race.
- Height (cm)
- Bodyweight (kg)
- BMI (kg/m^2) calculated as weight (kg) divided by $(height\ (m))^2$ rounded to one decimal place.

Medical History

Separate tabulations will be produced for previous and ongoing conditions with all conditions coded using MedDRA Version 16.0 (primary system organ class and preferred term).

Smoking & Alcohol History

- Ever smoked cigarettes (Yes, No). If yes,
    - Number of years used
    - Average number of times used/day
    - Whether currently smoking (Yes, No)
- Ever used other tobacco products (Yes, No). If yes, whether used
    - Cigar
    - Chew
    - Snuff
    - Pipe

Alcohol History

- Was Alcohol ever used? (Yes, No). If yes,
    - Number of years used alcohol was used (Yes, No)
    - Currently using (Yes, No)
    - Average number of times consumed/day

Opioids Abuse

(If for any subject, information is recorded within the obsolete opioids abuse screen and the concomitant medication screen, data recorded on the concomitant medication screen will take precedence)

- Primary Opioid of Abuse
- Duration of substance abuse (years). If a subject reports more than one drug the maximum duration reported will be used.
- Time since last use of an opioid to date of informed consent (days)
- Amount consumed per day (mg). If more than one drug is reported the cumulative mg use will be used.
- Average number times taken per day. If more than one drug is reported the cumulative value will be taken.

Infectious Disease Panel and Syphilis Tests at Screening

- Syphilis Antibody (Positive, Negative)
- HBsAg (Reactive, Non-reactive)
- Anti-HBc, Total (Reactive, Non-reactive)
- Hep B Surface Ab (IU/L)
- Anti-HCV (Reactive, Non-reactive)

Urine Drug Screen at Screening

- Result, positive for:
    - Amphetamines (Yes, No)

- Methamphetamines (Yes, No)
- Cocaine (Yes, No)
- Barbiturates (Yes, No)
- Opiates (Yes, No)
- Benzodiazepines (Yes, No)
- Cannabinoids (Yes, No)
- Methadone (Yes, No)
- Buprenorphine (Yes, No)

Prior medications

Verbatim terms (as recorded on the CRFs) of medications that ceased prior to the time of the initial dose of study medication will be mapped to Anatomical Therapeutic Chemical (ATC) Level 2 and Drug Reference Names using the World Health Organization (WHO) dictionary (Version 3.13 Enhanced). Prior medications will be listed in Appendix 16.2.

6.8 Measurement of Treatment Compliance

Each of the inpatient doses during Days 1-7 will be observed by the site staff. Following administration of the oral study medication, hand and mouth checks will be performed to ensure that the dose is swallowed. In an inpatient setting during Days 8-14, each dose will be observed by study staff and hand and mouth checks will be performed. In an outpatient setting during Days 8-14, self-dosing compliance will be evaluated by pill count and whether the dose has changed from the prior day and if so the reason for the change. Subjects will be instructed to call the physician's office before taking the next dose of study medication if they notice any marked dizziness, especially when standing from a sitting or lying position. The physician will determine if the next dose should be delayed, skipped, or the subject should be seen. Any change in physician prescribed dosing will be noted in the study file and confirmed also by pill count and subject report at the next visit. Details of compliance will be tabulated.

6.9 Primary Endpoint

The primary efficacy endpoint is AUC based on SOWS-Gossop scores from Days 1 through 7.

The study null hypothesis H_0 and alternative hypothesis H_1 are as follows:

H_0: There is no difference in the mean SOWS-Gossop AUC (1-7) between either of the 2.4 mg Lofexidine and 3.2 mg Lofexidine doses and placebo.

H_1: There is a difference in the mean SOWS-Gossop AUC (1-7) between either of the 2.4 mg Lofexidine and 3.2 mg Lofexidine doses and placebo.

The null hypothesis H_0 will be rejected in favor of H_1 if there is evidence at the $\alpha = 5\%$ significance level using a 2-sided test. Furthermore if H_0 is rejected in favor of H_1 and the mean

SOWS-Gossop AUC (1-7) is less for the Lofexidine dose in comparison to placebo then the Lofexidine dose for which there is evidence of a difference will be declared superior to placebo in the treatment of opioid withdrawal in adult patients with opioid dependence.

6.9.1 Principal Analysis

A pattern-mixture approach[5], with subjects stratified by disposition within Days 1-7, will be used in assessing SOWS-Gossop AUC (1-7). The 4 disposition strata are as follows:

- Subjects who complete Days 1-7 of the study (i.e. receives at least one dose of study medication on Day 7 and completes the 3.5-hour post-dose SOWS-Gossop assessment on Day 7);
- Subjects who discontinue due to lack of efficacy (including adverse events related to opioid withdrawal);
- Subjects who discontinue due to adverse events related to intolerability or toxicity to study drug; and
- Subjects who discontinue for other reasons.

The inference is comprised of several steps. The first two rely on standard statistical software and a commonly-used type of statistical model. The remaining three rely on matrix manipulations. Additionally, the last step appeals to the statistical delta method (propagation of errors).

1. Transformation. Because of the inherent skewness in the SOWS-Gossop scores the raw data will first be transformed to the natural logarithm of the score plus 1.0.
2. Modelling. A linear mixed effects repeated-measures model will be constructed for the SOWS-Gossop score. The fixed effects will be disposition stratum, treatment and time, and their respective two-way and three-way interactions, as well as a main effect of gender (randomization stratification factor) and baseline (pre-dose) SOWS-Gossop score as a one degree of freedom covariate. For each combination of treatment and stratum, the time course will be modelled as a linear change point model, allowing for one slope between Days 1 and 2 and a possibly different slope from Days 2 through 7. The model will parameterized to ensure that the predictions from the two line segments agree at Day 2. Subjects will be treated as a random effect, and the slope and intercept parameters will be treated as random coefficients. Modelling the time course, rather than using each time point as a discrete level of a model factor, allows estimation of group means through Day 7 even in the non-completer strata. The choice of Day 2 as the change point is based on prior lofexidine studies, in which SOWS-Gossop mean scores in the placebo group increased from Day 1 to Day 2 and then decreased through Day 7. The modelling allows a different time course for each of the lofexidine dose groups.
3. Point estimates of AUC (1-7). For each combination of disposition stratum, treatment and study day, the estimated least squares means SOWS-Gossop (on the log scale) for males and females will be combined using a weighted average, the weights being the relative proportions of males and females in the mITT population for each of the 12 stratum. The results will then be transformed back to the original SOWS-Gossop scale of measurement, from which point estimates of AUC (1-7) for each combination of stratum and treatment will be computed using the linear trapezoidal rule. It should be

noted that these calculations will not be derived from AUC(1-7) computed within individual subjects, many of whom will have incomplete data for the later study days in Days 1-7. In particular, there will be no imputation of data for individual subjects.

4. Estimates of treatment effect. For each active treatment, the treatment effect (lofexidine 3.2 mg versus placebo or lofexidine 2.4 mg versus placebo) will be estimated with respect to AUC (1-7) within each disposition stratum. Then a weighted average of treatment effects will be computed, where the weights are the relative proportions of subjects in the 4 strata.
5. Interval estimates of AUC (1-7) and hypothesis testing. The covariance matrix of the mixed model's estimated fixed effects will be used with standard linear model methods to derive the covariance matrix of the weighted average of male and female log-transformed SOWS-Gossop scores. The multivariate delta method (Ku, 1966; NIST/SEMATECH e-Handbook) will be used to derive the covariance matrix of the back-transformed SOWS-Gossop time course. Standard linear model methods will then be used to derive the covariance matrix of the weighted average of AUC (1-7) treatment effects, from which confidence intervals and p-values will be derived.

6.9.2 Sensitivity Analysis

The following sensitivity analyses will be conducted:

- The $\log_e$ (SOWS-Gossop score +1) from day 1 to day 7 will be analyzed using a mixed model repeated measures (MMRM) model.

  The model will include fixed effects for treatment group, baseline $\log_e$(SOWS-Gossop score +1), gender, visit (day 1, day 2 ,day 3, day 4,day 5, day 6 and day 7) and site (using the pooling algorithm detailed in section 6.2). In this model the treatment group term in the model represents the effect of treatment over days 1 to 7. An unstructured covariance model will be used. However, in the event that this model does not converge then the following covariance models will be used in order until the model converges: Toeplitz, first-order autoregressive and compound symmetry. Analysis will be performed using PROC MIXED in SAS and the resultant F-tests will be based on Kenward-Roger's adjusted degrees of freedom.

- The above MMRM analysis assumes that the missing data are missing at random (MAR). This analysis provides an unbiased estimate of the treatment effect that would have been observed had all subjects continued on treatment until day 7. It therefore assumes that the response for withdrawn subjects will follow the trajectory of the respective treatment after discontinuation. In order to assess the robustness of the results to the MAR assumption a sensitivity analysis will be conducted as follows under the assumption that the data are missing not at random (MNAR). Specifically, a pattern-mixture model will be implemented, where it will be assumed that Lofexidine withdrawals will have a trajectory comparable to placebo post withdrawal. This analysis will therefore provide a stress test of the Missing at Random assumption of the conventional mixed model employed in the MMRM analysis. The pattern-mixture model will be implemented using multiple imputations (Ratitch[6], 2011). This method is detailed in the section entitled "Pattern-Mixture Model with Control-Based Pattern Imputation" (Ratitch[7], 2013). Briefly, the initial step is to impute the relatively rare,

non-monotone missing data via the MCMC Option of SAS® PROC MI. This imputation model will include fixed effects for treatment, $\log_e$ (SOWS-Gossop score +1) at each of the visits, gender and site. The next step applies to the monotone missing data. The values for each pattern will be imputed using the chained equation method by SAS® PROC MI option MONOTONE REG. This MNAR imputation employs only the appropriate data for each pattern. That is, the missing data for Lofexidine treatment arm at time t will be imputed using only the data on placebo subjects up to time t and the data for Lofexidine subjects who have data up to time t-1 but are missing data at time t. The model for imputing missing data at time t will include fixed effect terms for $\log_e$ (SOWS-Gossop score +1) from the baseline visit through to the t-1 visit, gender and site. Note that that a term for treatment group is not included in this imputation model as the imputation is based on placebo subjects with outcome data at time t. This process continues until the missing data are imputed for all time points. The imputed data sets will be analyzed with the same MMRM model employed in the sensitivity analysis described in the 1st bullet and then summarized using PROC MIANALYZE.

- In a further sensitivity analysis a "tipping point analysis" (Ratitch[7], 2013) will be conducted in order to understand the "tipping point" δ at which the conclusions change from being statistically significant at the α=5% level using a 2-sided test in favor of Lofexidine to not statistically significant. This sensitivity analysis as per the sensitivity analysis described above begins with an imputation of the non-monotone missing data using the MCMC option of Proc MI, based on standard MAR assumptions. Then all monotone missing values are imputed using standard MAR assumptions, except in the imputation model in this instance will also include a covariate for treatment. This is the start of the tipping point analysis (delta = 0). Delta is the adjustment that is added to the $\log_e$(SOWS-Gossop score +1) at each visit after dropout during imputation. This increment is only added to the missing visits for the Lofexidine group. Delta is gradually increased in increments of 1 until the point at which the comparison for the Lofexidine group to placebo loses statistical significance. (An increment of 1 is selected because the SOWS-Gossop scale ranges from its lowest to highest score by consecutive integers.) Effectively this analysis provides the amount by which the Lofexidine subjects who discontinued early would need to be worse at each visit for the null hypothesis of no treatment difference to no longer be rejected. The imputed datasets are analyzed with the same MMRM model used in the primary analysis. The results are analyzed using PROC MIANALYSE.

- The studentized residuals will be assessed for normality using a histogram and quantile-quantile plot. Other diagnostic tests may be performed following the inspection of data and potential outliers. If departure from normality or outliers are detected then non-parametric analysis may be conducted. Full details will be provided within the final SAP prior to database lock. These diagnostics will be performed for all ANCOVA and MMRM analyses specified within this SAP.

6.9.3 Exploratory Analysis

The 1st MMRM sensitivity analyses of the primary efficacy endpoint will be performed by the following key baseline characteristics: Age, gender, race (white or non-white) and baseline

$\log_e$(SOWS-Gossop score+1). For each key baseline characteristic, the main effect and treatment-by-subgroup interaction terms will be added to the linear mixed effects repeated-measures model used in the primary endpoint analysis. Any interactions that are statistically significant at the 10% level may have their nature described. This will be determined after the blind is broken.

6.10 Secondary Endpoint

The secondary efficacy endpoint is the completion status (i.e., whether a subject receives at least one dose of study medication on Day 7 and completes the 3.5-hour post-dose SOWS-Gossop assessment on Day 7). The proportion of subjects in each treatment arm who receive at least one dose of study medication on Day 7 and complete the 3.5-hour post-dose SOWS-Gossop assessment on Day 7 will be analyzed using a logistic regression model including fixed effects for treatment group, gender and site using the mITT population. One test will compare the 3.2 mg lofexidine group to placebo; a second will compare the 2.4 mg Lofexidine group to placebo. A sensitivity analysis will be performed on this endpoint using the ITT population.

6.11 Tertiary/Exploratory Endpoints

SOWS-Gossop on Days 1, 2, 3, 4, 5, 6, and 7

An MMRM model will be used to test for a difference in the $\log_e$ (SOWS-Gossop score +1) at each of days 1, 2, 3, 4, 5, 6, and 7 between the 3.2 mg Lofexidine group and placebo and between the 2.4 mg Lofexidine group and placebo. The model will include fixed effects for treatment group, baseline $\log_e$ (SOWS-Gossop score +1), gender, visit (day 1, day 2, day 3, day 4, day 5, day 6 and day 7), visit by treatment group interaction and site (using the pooling algorithm detailed in section 6.2). An unstructured covariance model will be used. However, in the event that this model does not converge then the following covariance matrix will be used in order until the model converges: Toeplitz, first-order autoregressive and compound symmetry. Analysis will be performed using PROC MIXED in SAS and the resultant F-tests will be based on Kenward-Roger's adjusted degrees of freedom.

OOWS-Handelsman, MCGI (Subject and Rater), VAS-E, and COWS

The same methods used for SOWS-Gossop detailed above on days 1, 2, 3, 4, 5, 6, and 7 will be used for OOWS-Handelsman, MCGI (Subject and Rater), VAS-E, and COWS scores on days 1, 2, 3, 4, 5, 6, and 7. Depending on the distribution of these outcomes data transformations maybe used to normalise the data. Any transformation will be specified within the SAP prior to database lock.

Mean treatment profiles will be presented for the following:

- SOWS-Gossop
- OOWS-Handelsman
- MCGI (Subject and Rater)
- VAS-E

- COWS

Retention Analysis (Time to Removal from Study Treatment)

Time to removal from study treatment is defined as the last study day on which the subject received treatment. The time for subjects who complete some but not all treatment on Day 7 will be Day 7, uncensored. Subjects who complete all treatments on Day 7 will be censored at Day 7. This endpoint will be summarized descriptively for each combination of treatment group and gender with Kaplan-Meier curves and tabulations of the number and percentage of subjects newly removed from receiving study treatment on each of study Days 1 to 7. Each Lofexidine dose group will be compared inferentially to placebo with a Cox proportional hazards regression model of time to removal from study, including covariates for treatment, gender and site. If the Cox proportional hazards model does not converge, then the covariates gender and site will be dropped from the model and the Cox proportional hazards regression model will be stratified by gender and site instead. The estimated hazards ratio will be reported as a descriptive measure along with the associated 95% confidence interval.

Concomitant Medication Analysis

For each of Days 1 to 7, each subject's number of concomitant medication taken will be treated as a continuous variable. Descriptive statistics will be provided on the as-observed data on each study day and also overall study days.

Status of Detoxification on Day 7 or Early Termination as Assessed by the Site Investigator

Descriptive statistics (numbers and percentages) on the status of detoxification (successful/unsuccessful) as assessed by the Site Investigator at Day 7 or early termination will be presented by treatment group overall and by gender within treatment group.

COWS AUC (1-7)

COWS AUC (1-7) will also be analyzed using a modelling approach analogous to the modelling of SOWS-Gossop AUC (1-7). While the functional form of the time course for SOWS-Gossop is based on historical data, the functional form of the time course for COWS will be based on the data from this study.

In addition, COWS will be analyzed using a mixed model repeated measures (MMRM) model in the same way as detailed for the sensitivity analysis of the primary endpoint with baseline COWS (based on same functional form for primary analysis) as a covariate.

Assessment of effectiveness

Summary statistics will be provided by treatment day (overall and by gender) over days 8-14 (Open Label Phase):

- SOWS-Gossop
- OOWS-Handelsman
- MCGI (Subject and Rater)
- VAS-E

- COWS

In addition, the following will be summarized overall and by gender over Days 8-14:

- Number and proportion of subjects successfully completing detoxification and the number of days required to complete detoxification as assessed by the Site Investigator.

6.12 Multiplicity

For the primary and secondary endpoints, the treatment comparisons subject to control of the false positive rate will be 3.2 mg lofexidine vs. placebo and 2.4 mg lofexidine vs. placebo. Comparisons of 3.2 mg vs. 2.4 mg will be descriptive. The familywise error rate (FWE) for the collection of primary and secondary endpoint comparisons will be controlled at the 0.05 level, two-sided, by a sequential testing strategy in which hypotheses are tested in the following order, each at the 0.05 level, using a two-sided test.

1. Primary: AUC(1-7), 3.2 mg lofexidine versus placebo;
2. Primary: AUC(1-7), 2.4 mg lofexidine versus placebo;
3. Secondary: Completion rate, 3.2 mg lofexidine versus placebo; and
4. Secondary: Completion rate, 2.4 mg lofexidine versus placebo.

The tertiary/exploratory endpoints will be tested without multiplicity adjustment.

6.13 Safety analysis

All subjects who receive at least dose one of study medication will be included in the analysis of safety.

Safety measures will be summarized for the following subject cohorts for the open label phase:

- All treated subjects;
- Treated subjects without urinary evidence of illicit drug use; and
- Treated subjects with urinary evidence of illicit drug use (may be further subdivided by type of illicit drug used).

Safety measures will only be summarized for all treated subjects for the double blind phase.

Extent of exposure

Extent of exposure will be described by whether the subject took the trial medication and the number of days exposure to double-blind study medication (last date of double-blind medication minus first date of double-blind medication + 1). Average daily dose of double-blind study medication will also be calculated. Exposure to open-label medication will also be calculated as last date of open-label medication – first date of open-label medication + 1. No allowance will be made for breaks in therapy in the exposure calculations. Average daily dose of open-label study medication will also be calculated. If the date of last dosing is completely missing for the trial medication then the date of last dosing will be taken for analysis purposes

as the date the medication was last dispensed. If only the month of the last dose is recorded, the first day of the month will be assumed as the last dosing date.

Adverse Events During the Double-Blind Phase of the Study

All AEs will be listed and tabulated by severity, relationship to study medication, primary system organ class and preferred term according to MedDRA Version 16.0. In counting the number of events reported, a continuous event, i.e., an event reported more than once and which did not cease, will be counted only once with the worst recorded severity; non-continuous AEs reported several times by the same subject will be counted as multiple events. Events present immediately prior to the first dose of study medication that do not worsen in severity, will not be regarded as treatment emergent. Events with start dates more than 30 days after the administration of the last dose of study medication will not be considered treatment emergent and will be listed separately. In deriving the tabulation relating to preferred term reporting, the severity of a recurrent AE will be taken to be the most severe and the relationship to study medication as the highest probable.

The following will be summarized and presented for the overall population and by treatment group:

- The number and percentage of subjects experiencing at least 1 treatment emergent TEAE by MedDRA preferred term and SOC. A chi-squared test will be used to test for a difference in the percentage of subjects experiencing at least 1 TEAE between each of the 2 Lofexidine doses and placebo.
- The number and percentage of subjects experiencing an opioid withdrawal related TEAE by MedDRA preferred term and SOC.
- The number and percentage of subjects experiencing TEAEs by severity of event
- The number and percentage of subjects experiencing TEAEs by relationship to IP
- The number and percentage of subjects experiencing treatment-related adverse events (i.e. Possibly Related, Probably Related, Definitely Related)
- The number and percentage of subjects experiencing at least 1 treatment emergent SAE by MedDRA preferred term and SOC
- The number and percentage of subjects experiencing SAEs by severity of the event
- The number and percentage of subjects experiencing SAEs by relatedness to IP

Adverse events whilst receiving Lofexidine therapy during the open-label phase

All AEs reported whilst receiving Lofexidine therapy will be listed and tabulated by treatment, severity, relationship to study medication, primary system organ class and preferred term according to Version 16.0 of MedDRA. In counting the number of events reported, a continuous event, i.e., an event reported more than once and which did not cease, will be counted only once with the worst recorded severity; non-continuous AEs reported several times

by the same subject will be counted as multiple events. Events present immediately prior to the first dose of Lofexidine during the open-label phase that do not worsen in severity, will not be regarded as treatment emergent. Events with start dates more than 30 days after the administration of the last dose of Lofexidine will not be considered treatment emergent and will be listed separately. In deriving the tabulation relating to preferred term reporting, the severity of a recurrent AE will be taken to be the most severe and the relationship to study medication as the most probable.

For data analysis purposes, all TEAEs for which data are missing regarding assessments of relatedness and/or severity will be defaulted to be assessed as being related and at the highest severity grade, respectively. Missing or incomplete TEAE start dates will be imputed to correspond with the date of dosing.

Narratives of deaths, serious and other significant adverse events will be provided in the relevant section of the CSR.

A complete subject listing of all adverse events will be provided in Appendix 16.2 to the study report. This listing will include treatment, AE verbatim, MedDRA primary system organ class and preferred term, the time of onset and cessation of event relative to first dosing of study medication, duration of AE (for ongoing AEs use the date of investigator signature as the cessation date for calculation purposes), whether serious, severity, relationship to study medication, action taken and outcome.

Treatment emergent and non-treatment emergent events will be listed separately.

Clinical Laboratory Evaluations

Descriptive statistics will be calculated for the last double-blind and last open-phase post-baseline assessment as well as changes from baseline for each hematology, coagulation and chemistry variable. Each measurement will be classed as below, within, or above normal range, based on ranges supplied by the laboratory used. Shift tables in relation to the normal range from screening to each follow-up visit will be presented.

Descriptive statistics for urinalysis will be calculated for the last double-blind and last open-phase post-baseline assessment as well as changes from baseline.

Details of microscopic urinalysis will be provided in Appendix 16.2 of the report.

Vital Signs During the Double-Blind Phase

Vital signs (blood pressure and pulse [sitting or recumbent and standing], respiration, and temperature) will be measured at screening, baseline and before each dose and 3.5 hours after each dose (with the exception of the dose taken at 11 PM, where the measurement at 3.5 hours post-dose will not be done) during the double-blind phase. Summary statistics for observed and changes from baseline will be tabulated at each follow-up for each vital sign parameter and at post-baseline double-blind endpoint for each variable.

In addition the number and proportion of subjects having:

- Systolic blood pressure ≥180mmHg and an increase of ≥ 20mmHg from baseline
- Systolic blood pressure ≤ 90mmHg and a decrease ≥ 20mmHg from baseline
- Diastolic blood pressure ≥ 105mmHg and an increase ≥ 15mmHg from baseline
- Diastolic blood pressure ≤ 50mmHg and a decrease ≥ 15mmHg from baseline

will be summarized at each follow-up visits.

Vital Signs During the Open-Label Phase

Vital signs (blood pressure and pulse [sitting or recumbent and standing]) will be assessed for one dose every day at pre-dose and 3.5 hours post-dose on Days 8-13, with vital signs measured once pre-dose on Day 14. Summary statistics for observed and changes from baseline will be tabulated at each follow-up and at post-baseline open-phase endpoint for each variable.

Mean profiles by treatment will be presented for sitting and standing blood pressure and heart rate.

ECG

ECG data collected during the double-blind and open-label phase will be analyzed separately in a stand-alone report.

In addition, PI overall interpretation of the ECG (Normal, Abnormal NCS, and Abnormal CS) will be listed in Appendix 16.2.

C-SSRS

The following will be summarized and analyzed separately for the double blind and open label study phases**:**

1. To assess safety:
    a. **Suicidality:** The number and percentage of subjects reporting any suicidal ideation or behavior throughout the assessment period.
    b. **Suicidal behavior only:** The number and percentage of subjects reporting any type of suicidal behavior throughout assessment period.
    c. **Suicidal ideation only:** The number and percentage of subjects reporting any type of suicidal ideation throughout assessment period.

    In addition, taking into account baseline data regarding suicidal ideation to determine if suicidal ideation or behavior has worsened:

    d. **Emergence of suicidal ideation:** The number and percentage of subjects reporting no suicidal ideation at baseline and any type of ideation during the assessment period.

e. **Emergence of serious suicidal ideation:** The number and percentage of subjects reporting no suicidal ideation at baseline and had serious suicidal ideation (as defined above; score of 4 or 5 on suicidal ideation severity rating) during the assessment period.
f. **Worsening of suicidal ideation:** The number and percentage of subjects whose most severe suicidal ideation rating is more severe than it was at baseline.
g. **Emergence of suicidal behavior:** The number and percentage of subjects who had no suicidal behavior at baseline and any type of behavior during the assessment period.

For each of the above endpoints a-g, Fisher's Exact test will be used to test for a difference between the 2 Lofexidine treatment groups and placebo during the double blind phase only.

Physical Examination

The body systems within the physical examination data at screening, baseline (day 1), day 7 and end of the study will be summarized by treatment (Normal; Abnormal NCS, Abnormal CS). The changes from baseline to end of the study in bodyweight will also be tabulated.

Concomitant medications

Concomitant medication verbatim terms (as recorded on the CRFs) after the initial dose of study drug will be mapped to Anatomical Therapeutic Chemical (ATC) Level 2 and Drug Reference Names using the World Health Organization (WHO) dictionary (Version 3.13 Enhanced) and tabulated by treatment group and overall.

Psychosocial Therapy Sessions

- Whether subject received or participated in Psychosocial Therapy Sessions during the double blind phase (Yes, No). If yes, whether received the following
    - Group Counselling Session (Yes, No)
    - Art or Music therapy (Yes, No)
    - 12-Step Program Participation (Yes, No)
    - Other (Yes, No)

Details of the Day 30 Follow-Up

- Lost to Follow-up
- Whether subject successfully entered one of the following programs:
    - None
    - Methadone
    - Buprenorphine

- Naltrexone
- Other

- Whether subject relapsed to drug use after exiting the study (Yes, No)

6.14 Change to Planned Protocol Analysis

The following changes have been made to the planned protocol analyses sections:

Section 17.2.7 Control of the False Positive Rate and Statistical Testing Strategy

Change: All statistical tests will be two-sided at the 0.05 significance level. This is a change from the protocol, where it was specified that statistical tests would be one-sided at the 0.025 significance level.

Reason: Two-sided statistical tests at the 0.05 significance level is statistical convention for hypothesis testing.

Section 17.2.8.1 SOWS-Gossop AUC (1-7)

Change: A number of sensitivity analyses have been added and are detailed in section 16.9.2 of the statistical analysis plan.

Reason: To provide additional sensitivity analyses of $\log_e$ (SOWS-Gossop score +1) from day 1 to day 7 under differing mechanisms for missing data, including Missing at Random (MAR) and Missing Not at Random (MNAR)

Section 17.2.8.2 Completion Status

Change: A logistic regression model including fixed effects for treatment group, gender and site will be used instead of a Cochran-Mantel Haenszel test to test for a difference in completion status between the 2 Lofexidine groups and placebo.

Reason: To additionally adjust for site in the analysis and to provide an interpretation of the effect of treatment using odds ratios.

Sections: 17.2.8.3 SOWS-Gossop on Days 1, 2, 3, 5, 6 and 7 and 17.2.8.4 OOWS-Handelsman, MCGI (Subject and Rater), VAS-E, and COWS

Change: A MMRM model will be used to test for a difference in the $\log_e$ (SOWS-Gossop score +1) at each of days 1 through 7 between the 2 Lofexidine groups and placebo. This replaces planned pattern-mixture approach used in the principal analysis.

Reason: MMRM model is a standard approach for the analysis of treatment differences at different time points when there are repeated measures.

17.2.8.5 Retention Analysis (Time to Removal from Study Treatment)

Change: Gender will be included in the Cox proportional hazards regression model as covariate rather than stratifying Cox analysis by gender. In addition site will also be included in the

model. The confidence intervals for the hazard rations will also be two-tailed rather than one-tailed as specified within the protocol.

Reason: To ensure the modelling and presentation of confidence intervals is consistent with other multivariate statistical models specified within the statistical analysis plan.

7. REFERENCES

(1) Bazett HC. An analysis of time relations of the electrocardiogram. Heart 1920; 7: 353-370.3.

(2) Fridericia LS. Die systolendauer in elektrokardio-gram bei normalen menschen und bei herzkranken. Acta Med Scand 1920; 53: 469.

(3) SAS Institute Inc. The SAS System, Version 9.3. Cary, NC, SAS Institute Inc. 2012.

(4) ICH E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs.

(5) Little RJA. Pattern-mixture models for multivariate incomplete data. Journal of the American Statistical Association 1993;88:125-134.

(6) Ratitch B and O'Kelly M. Implementation of Pattern Mixture Models Using Standard SAS/STAT Procedures. PharmaSug 2011 paper SP04.

(7) Ratitch B and O'Kelly M. Missing Data in Clinical Trials: from Clinical Assumptions to Statistical Analysis using Pattern Mixture Models. Pharmaceutical Statistics 2013: DOI:10.1002/pst.1549.

Approval for implementation of

Statistical Analysis Plan

Title:	**A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety, and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)**
Reference:	**USWM-LX1-3003-1/SAP**
Version:	**1.0**
Date effective:	**18-JUN-14**
Author:	
reviewer:	

Author's signature: _____ **Date:** _____

Reviewer's signature: _____ **Date:** _____

The above Statistical Analysis Plan has been reviewed and approved by the Sponsor:

Name of Approver:

Position: **Medical Director, US WorldMeds, LLC**

Signature for _____ Date: _____

Name of Approver:

Position: **Sr. Manager, Biometrics, US WorldMeds, LLC**

Signature for _____ Date: _____

Protocol Number USWM-LX1-3003-1
Statistical Analysis Plan Addendum 1
Version: Final
Issue Date: 24-Dec-2014

STATISTICAL ANALYSIS PLAN

ADDENDUM 1

A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)

Protocol Ref: USWM-LX1-3003-1

Date effective: 24-Dec-2014

SUMMARY OF SECTIONS

1. INTRODUCTION

This document is an addendum to the statistical analysis plan dated 18th June 2014 for the analysis of the US WorldMeds, LLC study USWM-LX1-3003-1.

All decisions were taken prior to database lock.

2. CHANGES TO EXISTING ANALYSIS PLAN

Change 1

Section 6.1 General

Add following sentence

For all statistical models analyzed via log-transformed data the back-transformed LS means will be presented in the output.

Change 2

Section 6.2 Derived Data

Site Pooling for Statistical Analysis

The following text was removed:

For the purposes of statistical analyses models where site is included within the model as a fixed effect, sites randomizing < 6 subjects will be pooled from smallest to largest until the pooled site subsequently created has at least 6 subjects. In the event that the pooling of sites with less than 6 subjects does not result in a site with > 6 subjects, then that pooled site with < 6 subjects will also be pooled with the next largest site with > 6 subjects. To break ties if more than one site has the same number of subjects, the site identifier will be used as a secondary sort key.

Justification

The primary analysis no longer includes site in the model; therefore site has been removed as a variable from all subsequent statistical models.

Change 3

Section 6.3 Analysis Populations

An additional population has been added

The **per-protocol population.** This analysis set will be a subset of the mITT population and will consist of all subjects who satisfy all of the inclusion/exclusion criteria and who correctly receive the treatment to which they are randomized and do not have a major protocol deviation within the double-blind phase. The selection of subjects for the per-protocol population will be made prior to database lock. Major protocol deviation will be broadly classified into five main categories based on ICH Harmonised Tripartite Guideline Structure and Content of Clinical Study Reports E3:

1. those patients who enter the study despite not satisfying entry criteria
2. those patients who developed non-safety related withdrawal criteria during the study and were not withdrawn
3. those patients who received the wrong treatment
4. those patients who received an excluded concomitant treatment, defined as any drug that is not included in the approved list of medications (as outlined in Appendix 16.2.2) or any antibiotic
5. lack of major compliance with the protocol.

The only variable to be assessed using the per-protocol analysis set will be the primary efficacy endpoint.

Justification

Due to the identification of some subjects taking prohibited medication during the double-blind period; a per-protocol population has been added to conduct a sensitivity analysis on the primary endpoint.

Change 4

Section 6.7.2 Variables Considered

Medical History

Add the text in bold

Separate tabulations will be produced for previous and ongoing conditions with all conditions coded using MedDRA Version 16.0 (primary system organ class and preferred term). **An ongoing condition is defined as any history with any of the following**

- ≠ **Marked as ongoing**
- ≠ **With a missing stop date and ongoing not checked**
- ≠ **Or a stop date after the first dose of study medication.**

Justification

To make the definition clearer of what constitutes an ongoing condition.

Change 5

Section 6.7.2 Variables Considered

Infectious Disease Panel and Syphilis Tests at Screening

- ≠ HBsAg (Reactive, Non-reactive)

Amend to

- ≠ HBsAg (**Confirmed Positive**, Non-reactive)

Justification

Changed due to how the data is actually recorded on the laboratory vendor file.

Change 6

Section 6.9 Primary Endpoint

Amend to

The primary efficacy endpoint is the difference between the overall LS means from a pattern mixture model based on log-transformed SOWS-Gossop scores from Days 1 through 7.

The study null hypothesis H_0 and alternative hypothesis H_1 are as follows:

H_0: There is no difference in the overall means in log-transformed SOWS-Gossop scores from Days 1 through 7 between either of the 2.4 mg lofexidine and 3.2 mg lofexidine doses and placebo.

H_1: There is a difference in the overall means in log-transformed SOWS-Gossop scores from Days 1 through 7 between either of the 2.4 mg lofexidine and 3.2 mg lofexidine doses and placebo.

The null hypothesis H_0 will be rejected in favor of H_1 if there is evidence at the $\alpha = 5\%$ significance level using a 2-sided test. Furthermore if H_0 is rejected in favor of H_1 and the overall LS mean is less for the lofexidine dose in comparison to placebo, then any lofexidine dose for which there is evidence of a difference will be declared superior to placebo in the treatment of opioid withdrawal in adult patients with opioid dependence. To account for multiplicity of statistical hypothesis testing of two lofexidine doses, statistical significance of the comparison of high dose lofexidine to placebo must be established, before significance of the comparison of low dose lofexidine to placebo can be claimed.

Justification

The primary efficacy endpoint has changed from the area under the curve (AUC) from Days 1 to 7 of the SOWS-Gossop scores to the overall (i.e., averaged over study Days 1 through 7) least squares (LS) means from a pattern mixture model based on log-transformed SOWS-Gossop scores from Days 1 through 7. The new endpoint stems from a revision of the pattern mixture model that incorporates comments on the modeling approach from the US Food and Drug Administration (FDA) in relation to patients who discontinue treatment.

Using the proposed pattern mixture model, the combined process of imputation followed by MMRM does not assume MAR and provides an unbiased estimate of the treatment effect that follows under the assumption that the mechanism by which data are missing not at random (MNAR) is that upon withdrawal from the study, the SOWS-Gossop trajectory of subjects randomized to lofexidine evolves as the trajectory of placebo subjects. Thus, by assigning an

unfavorable placebo trajectory to patients who discontinue lofexidine treatment, the proposed model has addressed this concern. The nature of these changes is described in Change 7, following.

Change 7

Section 6.9.1 Principal Analysis

Amend to

A pattern-mixture approach[5] will be used in assessing SOWS-Gossop from Days 1 to 7 inclusive using the mITT population.

It will be assumed that subjects randomized to lofexidine who withdraw from the study will have a time trajectory comparable to placebo post withdrawal. The pattern-mixture model will be implemented using multiple imputations (Ratitch[6], 2011). The imputation of missing values and the analysis will be performed multiple times (20 imputed datasets) with the initial seed value set at '123', and the inference of this sensitivity analysis will be based on the combined estimates using the standard multiple imputation technique.This method is detailed in the section entitled "Pattern-Mixture Model with Control-Based Pattern Imputation" (Ratitch[7], 2013). Briefly, the initial step is to impute the relatively rare, non-monotone missing data via the MCMC Option of SAS® PROC MI. This step uses iterative methods, with which non-convergence is a possibility; should an MCMC imputation step not converge, additional imputations will be performed, and the first 20 MCMC imputations that converge will be used in the analysis. This imputation model will be implemented for each combination of gender and treatment. (It should be noted that by performing the imputation for each such combination, the imputation modelling will be richer than the model used to inferentially analyse the data. It is acceptable, even preferable, for the imputation model to be richer than for the analysis model (see Little and Rubin, 2002).

The next step applies to the monotone missing data. The values for each pattern will be imputed using the chained equation method by SAS® PROC MI option MONOTONE REG. This MNAR imputation employs only the appropriate data for each pattern. That is, the missing data for lofexidine treatment arm at time t will be imputed using only the data on placebo subjects up to time t and the data for lofexidine subjects who have data up to time t-1 but are missing data at time t. The model for imputing missing data at time t will include fixed effect terms for study day from the baseline visit through to the t-1 visit and sex. Note that a term for treatment group is not included in this imputation model. This process continues until the missing data are imputed for all time points. The imputed data sets will be analyzed using the MMRM model described below and then summarized using PROC MIANALYZE. The model will include fixed effects for treatment group, baseline $\log_e$ (SOWS-Gossop score +1), sex, study day (days 1-7), and

treatment group-by-day interaction. In this model the treatment group term in the model represents the effect of treatment over days 1 to 7. An unstructured covariance model will be used. However, in the event that this model does not converge then the following covariance models will be used in order until the model converges: Toeplitz, first-order autoregressive and compound symmetry. Analysis will be performed using PROC MIXED in SAS and the resultant F-tests will be based on Kenward-Roger's adjusted degrees of freedom.

Justification

Fitting an MMRM analysis without the imputation described above would have entailed an assumption that the missing data are missing at random (MAR). The combined process of imputation followed by MMRM does not assume MAR and provides an unbiased estimate of the treatment effect that follows under the assumption that the mechanism by which data are missing not at random (MNAR) is that upon withdrawal from the study, the SOWS-Gossop trajectory of subjects randomized to lofexidine evolves as the trajectory of placebo subjects.

Change 8

Section 6.9.1 Sensitivity Analysis

The text in bold has been edited:

The following sensitivity analyses will be conducted:

- ≠ **The above analysis will be performed using the per-protocol population.**
- ≠ The $\log_e$ (SOWS-Gossop score plus 1) from day 1 to day 7 will be analyzed using a **MMRM model on the mITT population**.

  The model will include fixed effects for treatment group, baseline $\log_e$ (SOWS-Gossop score plus 1), sex, study day (days 1-7) and treatment group-by-day interaction. In this model the treatment group term in the model represents the effect of treatment over days 1 to 7. An unstructured covariance model will be used. However, in the event that this model does not converge then the following covariance models will be used in order until the model converges: Toeplitz, first-order autoregressive and compound symmetry. Analysis will be performed using PROC MIXED in SAS and the resultant F-tests will be based on Kenward-Roger's adjusted degrees of freedom. This MMRM analysis assumes that the missing data are missing at random (MAR). This analysis provides an unbiased estimate of the treatment effect that would have been observed had all subjects continued

on treatment until day 7. It therefore assumes that the response for withdrawn subjects would have followed the trajectory of the respective treatment after discontinuation, had they not withdrawn.

≠ In a further sensitivity analysis on the mITT population a "tipping point analysis" (Ratitch[7], 2013) will be conducted in order to understand the "tipping point" δ at which the conclusions change from being statistically significant at the α=5% level using a 2-sided test in favor of lofexidine to not statistically significant. **This sensitivity analysis as per the sensitivity analysis described above begins with an imputation of the relatively rare non-monotone missing data using the MCMC option of PROC MI, based on standard MAR assumptions. The random seed for this set of analyses will be '456'. Then several analyses will be performed, indexed by the value δ, in each of which all monotone missing values are imputed using standard MAR assumptions, except in the imputation model in this instance will also include a covariate for treatment.** In the first such analysis δ = 0, which effectively is last observation carried forward. In general, δ is the adjustment that is added to the $\log_e$ (SOWS-Gossop score +1) at each visit after dropout during imputation. This increment is only added to the missing visits for the lofexidine groups. Delta is increased in increments of 1 (on the original SOWS-Gossop scale, before log transformation) until the point at which the comparison for the lofexidine group to placebo loses statistical significance. (An increment of 1 is selected, because the SOWS-Gossop scale ranges from its lowest to highest score by consecutive integers.) Effectively this analysis provides the amount by which the lofexidine subjects who discontinued early would need to be worse at each visit for the null hypothesis of no treatment difference to no longer be rejected. The imputed datasets are analyzed with the same MMRM model used in the primary analysis. The results are analyzed using PROC MIANALYZE.

≠ The studentized residuals will be assessed for normality using a histogram and quantile-quantile plot. Other diagnostic tests may be performed following the inspection of data and potential outliers. If departures from normality or outliers are detected then non-parametric analysis may be conducted. Full details will be provided within the final SAP prior to database lock. These diagnostics will be performed for all ANCOVA and MMRM analyses specified within this SAP.

Justification
As a result of changes made to the principal statistical analysis (Change 7 annotated, Section 6.9.1) of the primary endpoint, the planed sensitivity analyses also required modification. Further, a per-protocol analysis of the primary endpoint was added due to the number of reported major protocol deviations.

Change 9

Section 6.10 Secondary Endpoint

The following text was amended:

The secondary efficacy endpoint is the completion status (i.e., whether a subject receives at least one dose of study medication on Day 7 and completes the 3.5-hour post-dose SOWS-Gossop assessment on Day 7). **The proportion of subjects in each treatment arm who receive at least one dose of study medication on Day 7 and complete the 3.5-hour post-dose SOWS-Gossop assessment on Day 7 will be analyzed using a logistic regression model including fixed effects for treatment group and gender using the mITT population.** One test will compare the 3.2 mg lofexidine group to placebo; a second will compare the 2.4 mg Lofexidine group to placebo. A sensitivity analysis will be performed on this endpoint using the ITT population.

Justification

Site has been removed as a variable from all statistical models. All references for site will be removed from the statistical analysis plan where previously included in statistical models.

Change 10

Section 6.11 Tertiary/Exploratory Endpoints

Amend the following text related to SOWS-Gossop on Days 1, 2, 3, 4, 5, 6, and 7

An MMRM model will be used to test for a difference in the $\log_e$ (SOWS-Gossop score +1) at each of days 1, 2, 3, 4, 5, 6, and 7 between the 3.2 mg Lofexidine group and placebo and between the 2.4 mg Lofexidine group and placebo. **The model will include fixed effects for treatment group, baseline $\log_e$ (SOWS-Gossop score +1), sex, visit (day 1, day 2, day 3, day 4, day 5, day 6 and day 7) and treatment-by-visit interaction.** An unstructured covariance model will be used. However, in the event that this model does not converge then the following covariance matrix will be used in order until the model converges: Toeplitz, first-order autoregressive and compound symmetry. Analysis will be performed using PROC MIXED in SAS and the resultant F-tests will be based on Kenward-Roger's adjusted degrees of freedom.

Justification

The primary analysis no longer includes site in the model; therefore site has been removed as a variable from all subsequent statistical models.

Change 11

Section 6.11 Tertiary/Exploratory Endpoints

Amend the following text related to OOWS-Handelsman, MCGI (Subject and Rater), VAS-E, and COWS

The same methods used for SOWS-Gossop detailed above on days 1, 2, 3, 4, 5, 6, and 7 will be used for OOWS-Handelsman, MCGI (Subject and Rater), VAS-E, and COWS scores on days 1, 2, 3, 4, 5, 6, and 7. Depending on the distribution of these outcomes data transformations maybe used to normalise the data. Any transformation will be specified within the SAP prior to database lock.

The text in bold was added to the following:

The same methods used for SOWS-Gossop detailed above on days 1, 2, 3, 4, 5, 6, and 7 will be used for OOWS-Handelsman, MCGI (Subject [**Severity of opiate withdrawal and side effect index separately**] and Rater [**Severity of illness and side effect index separately**], VAS-E, and COWS scores on days 1, 2, 3, 4, 5, 6, and 7, **where the baseline covariate will be the baseline value of the endpoint under analysis. The same model will be used for MCGI [Subject and Rater] and VAS-E, except that there will be no baseline covariate as these endpoints do not have a baseline measurement.** Depending on the distribution of these outcomes data transformations maybe used to normalise the data. Any transformation will be specified within the SAP prior to database lock.

Justification

Certain efficacy variables are not measured at baseline. The MCGI Subject and Rater scales are each sub-divided into two components.

Change 12

Section 6.11 Tertiary/Exploratory Endpoints

The text in bold was amended for Retention Analysis (Time to Removal from Study Treatment):

Time to removal from study treatment is defined as the last study day on which the subject received treatment. **A completer for the purpose of this analysis is defined as a subject receiving at least 1 dose on Day 7 and the having the SOWS assessment completed. This will be based on the flag indicating the subject completed the double-blind phase as recorded on the CRF.** The time to removal from study treatment for completers will be censored on Day 7. The time to removal from study treatment for subjects who complete some treatment on Day 7 but who do not satisfy the definition of completer will be Day 7, uncensored.

This endpoint will be summarized descriptively for each combination of treatment group and gender with Kaplan-Meier curves and tabulations of the number and percentage of subjects newly removed from receiving study treatment on each of study Days 1 to 7. **Each Lofexidine dose group will be compared inferentially to placebo with a Cox proportional hazards regression model of time to removal from study, including covariates for treatment and sex. If the Cox proportional hazards model does not converge, then the covariate sex will be dropped from the model and the Cox proportional hazards regression model will be stratified by sex instead. The estimated hazards ratio will be reported as a descriptive measure along with the associated 95% confidence interval.**

Justification

To make the definition clearer of what constitutes a completer for this analysis and to redefine the censoring algorithm accordingly. The primary analysis no longer includes site in the model; therefore site has been removed as a variable from all subsequent statistical models.

Change 13

Section 6.11 Tertiary/Exploratory Endpoints

The text in bold was added.

Concomitant Medication Analysis

For each of Days 1 to 7, each subject's number of concomitant medication taken will be treated as a continuous variable. Descriptive statistics will be provided on the as-observed data on each study day and also overall study days.

Only days when the patient was taking active treatment will be considered. Days after active administration will be disregarded.

For the calculation of the daily number of medications taken, the number of distinct coded drug names via the WHO dictionary will be used rather than verbatim terms.

Justification

To make the definitions within this analysis clearer.

Change 14

Section 6.13 Safety analysis

Add the text in bold

All subjects who receive at least dose one of study medication will be included in the analysis of safety.

Safety measures will be summarized for the following subject cohorts for the open label phase:

- ≠ All treated subjects;
- ≠ Treated subjects without urinary evidence of illicit drug use **(i.e. any positive result in the urine drug screen during the open-label phase, including all categories)**; and
- ≠ Treated subjects with urinary evidence of illicit drug use (may be further subdivided by type of illicit drug used).

Safety measures will only be summarized for all treated subjects for the double blind phase.

Extent of exposure

Extent of exposure will be described by whether the subject took the trial medication and the number of days exposure to double-blind study medication (last date of double-blind medication minus first date of double-blind medication + 1). Average daily dose of double-blind study

medication will also be calculated **by number of doses/exposure (days)**. Exposure to open-label medication will also be calculated as last date of open-label medication – first date of open-label medication + 1. No allowance will be made for breaks in therapy in the exposure calculations. Average daily dose of open-label study medication will also be calculated. If the date of last dosing is completely missing for the trial medication then the date of last dosing will be taken for analysis purposes as the date the medication was last dispensed. **As there are no open-label dosing dates recorded on the eCRF, dates of visits will be used instead. In particular it will be assumed that open-label dosing starts one day after double-blind dosing ceases.**

Justification

To make the definitions clearer.

Change 15

Section 6.13 Safety analysis

Vital Signs During the Double-Blind Phase

The following text was amended:

Vital signs (blood pressure and pulse [sitting or recumbent and standing], respiration, and temperature) will be measured at screening, baseline and before each dose and 3.5 hours after each dose (with the exception of the dose taken at 11 PM, where the measurement at 3.5 hours post-dose will not be done) during the double-blind phase. Summary statistics for observed and changes from screening will be tabulated at each follow-up for each vital sign parameter and at post-baseline double-blind endpoint for each variable.

In addition the number and proportion of subjects having:

- ≠ Systolic blood pressure ≥180mmHg and an increase of ≥ 20mmHg from baseline
- ≠ Systolic blood pressure ≤ 90mmHg and a decrease ≥ 20mmHg from baseline
- ≠ Diastolic blood pressure ≥ 105mmHg and an increase ≥ 15mmHg from baseline
- ≠ Diastolic blood pressure ≤ 50mmHg and a decrease ≥ 15mmHg from baseline

will be summarized at each follow-up visits.

The text in bold was added to the following:

Vital signs (blood pressure and pulse [sitting or recumbent and standing], respiration, and temperature) will be measured at screening, baseline and before each dose and 3.5 hours after each dose (with the exception of the dose taken at 11 PM, where the measurement at 3.5 hours post-dose will not be done) during the double-blind phase. Summary statistics for observed and changes from **screening** will be tabulated at each follow-up for each vital sign parameter and at post-baseline double-blind endpoint for each variable.

In addition the number and proportion of subjects having:

- ≠ Systolic blood pressure ≥ 180 mmHg and an increase of ≥ 20 mmHg from **screening**
- ≠ Systolic blood pressure ≤ 90 mmHg and a decrease ≥ 20 mmHg from **screening**
- ≠ Diastolic blood pressure ≥ 105 mmHg and an increase ≥ 15 mmHg from **screening**
- ≠ Diastolic blood pressure ≤ 50 mmHg and a decrease ≥ 15 mmHg from **screening**

will be summarized at each follow-up visits.

REFERENCES

Little, RJA and DB Rubin. Statistical analysis with missing data. Version 2. John Wiley & Sons; New York: 2002

Approval for implementation of

Statistical Analysis Plan Addendum

Title:	**A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)**
Reference:	**USWM-LX1-3003-1/SAP Addendum/1**
Addendum No:	**1**
Date effective:	**24-Dec-2014**
Author:	**Principal Statistician**
reviewer:	**Principal Statistician**

Author's signature: _____ Date: _____

Reviewer's signature: _____ Date: _____

The above Statistical Analysis Plan Addendum has been reviewed and approved by the Sponsor:

Name of Approver:
Position: **Medical Director, US WorldMeds, LLC**

Signature: _____ Date: _____

Name of Approver:

Position: **Sr. Manager, Biometrics, US WorldMeds, LLC**

STATISTICAL ANALYSIS PLAN

ADDENDUM 2

A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)

Protocol Ref: USWM-LX1-3003-1

Date effective: 22-Jan-2015

SUMMARY OF SECTIONS

1. INTRODUCTION

This document is the second addendum to the statistical analysis plan for the analysis of the US WorldMeds, LLC study USWM-LX1-3003-1. The original statistical analysis plan was dated 18-Jun-14 and amendment number 1 was dated 24-Dec-14.

All decisions were taken prior to database lock.

2. CHANGES TO EXISTING ANALYSIS PLAN

Change 1

Section 6.2 Derived Data

Questionnaire Data

The following sentence was revised in bold:

In the event that component items used to calculate total scores for questionnaire data e.g. SOWS Gossop score are missing, then the total scores will **be computed as the average of the non-missing item scores multiplied by the total number of items (e.g. multiply by 10 for the SOWS-Gossop) comprising the total score. This same method will be applied to all other questionnaire data to calculate total scores in a similar fashion.**

Justification

To minimize the number of patients with missing total assessment scores due to missing single component items. This solution incorporates comments from the US Food and Drug Administration (FDA) in relation to the calculation of total scores based on non-missing item scores.

Change 2

Section 6.3 Analysis Populations

The following sentence was added in bold:

The Enrolled Population includes all subjects screened into the study irrespective of whether they received the study medication.

Three principal analysis populations are defined as follows:

- Intent-to-treat (ITT), consisting of all randomized subjects. Subjects will be assigned for analysis according to the group to which they are randomized.
- Modified intent-to-treat (mITT), consisting of all subjects in the ITT group who received at least one dose of study medication. Subjects will be assigned for analysis according to the group to which they are randomized.
- Safety consisting of all subjects who received at least one dose of study medication. **Subjects will be assigned for analysis according to treatment received.**

Justification

To further clarify the definition of the safety analysis population.

Change 3

Section 6.9.2 Sensitivity Analysis

The following text will be added:

Additionally, further efficacy analysis will be conducted using the area under the curve (AUC) from Days 1 through 7 of the SOWS-Gossop scores as the endpoint. The AUC will be calculated on the original scale of measurement based on the back-transformed log scale estimates from the pattern mixture analysis. The LS means of the daily SOWS-Gossop scores (Days 1 through 7) will be exponentiated to get the geometric mean SOWS-Gossop score for each day. These back-transformed scores will then be used to calculate the AUC using the trapezoidal rule. AUC summary statistics will be reported, no formal statistical analysis will be performed.

Justification

The additional analysis was added to incorporate comments from the US Food and Drug Administration (FDA) regarding an additional proposed efficacy analysis. Since statistical inference on the AUCs, such as confidence intervals, would be difficult and the results at best

approximate, due to the fact that a propagation of errors approach would be needed to accommodate the exponentiation of results from the log scale of measurement back to the original scale; we have elected to present the estimated AUC values as descriptive statistics.

Change 4

Section 6.10 Secondary Endpoint

The following text was amended:

The secondary efficacy endpoint is the completion status (i.e., whether a subject receives at least one dose of study medication on Day 7 and completes the 3.5-hour post-dose SOWS-Gossop assessment on Day 7). **The proportion of subjects in each treatment arm who receive at least one dose of study medication on Day 7 and complete the 3.5-hour post-dose SOWS-Gossop assessment on Day 7 will be analyzed using a logistic regression model including fixed effects for treatment group and gender using the mITT population.** One test will compare the 3.2 mg lofexidine group to placebo; a second will compare the 2.4 mg Lofexidine group to placebo. A sensitivity analysis will be performed on this endpoint using the ITT population.

Justification

Gender was inadvertently removed as a variable from the logistic regression model.

Change 5

Section 6.2 Assessments Recorded at Discontinuation Visit

Add the following text related to handling assessments recorded on discontinuation visit:

On days where a discontinuation visit is recorded on the same day as a scheduled visit, then the **scheduled visit** should be used be for analysis purposes and the discontinuation visit will be ignored. If a discontinuation visit is recorded on a study day where no scheduled visit is recorded, then it should be assigned to the **actual study day** the visit was recorded.

Justification

To further clarify a set of rules for how to handle assessments recorded on the date of discontinuation instead of the scheduled study visit.

Change 6

Section 6.9.1 Principal Analysis

The following text in bold has been revised:

A pattern-mixture approach[5] will be used in assessing SOWS-Gossop from Days 1 to 7 inclusive using the mITT population.

It will be assumed that subjects randomized to lofexidine who withdraw from the study will have a time trajectory comparable to placebo post withdrawal. The pattern-mixture model will be implemented using multiple imputations (Ratitch[6], 2011). The imputation of missing values and the analysis will be performed multiple times (20 imputed datasets) with the initial seed value set at '123', and the inference of this sensitivity analysis will be based on the combined estimates using the standard multiple imputation technique. This method is detailed in the section entitled "Pattern-Mixture Model with Control-Based Pattern Imputation" (Ratitch[7], 2013). Briefly, the initial step is to impute the relatively rare, non-monotone missing data via the MCMC Option of SAS® PROC MI. **The imputation model will include as predictors an indicator variable for gender (1 d.f.), two indicator variables for treatment (2 d.f), the Baseline value, and variables for Day 1 through Day 7 data. This step uses iterative methods, with which non-convergence is a possibility; should an MCMC imputation step not converge, alternative models will be developed by dropping time point variables from the model, one at a time starting at Day 7, until convergence is achieved. For example, if non-convergence occurs for the model that includes all the time point variables from baseline through Day 7, the next model to be considered will have all time point variables from baseline through Day 6, as well as the indicators for treatment and gender. The next model, if needed, would have all time point variables from baseline through Day 5, and so forth. Should this approach not succeed, ad hoc methods to impute non-monotone missing data will be developed and documented.**

These contingency models for non-convergence, as well as the modeling of data with a monotone missing pattern, will require additional random number seeds. As stated in the SAP, the seed for the first step of the monotone missing imputation is '234'. The following additional seeds are to be used in the order shown for any additional modeling of non-monotone missing data as well as for the subsequent imputation of monotone missing data.

1. **345**
2. **456**
3. **567**
4. **678**
5. **789**
6. **890**
7. **901**
8. **012**
9. **124**
10. **245**
11. **356**
12. **467**
13. **578**
14. **689**
15. **790**
16. **145**
17. **256**
18. **367**
19. **478**
20. **589**

Justification

To acknowledge the possibility of non-convergence and to clarify a process to implement systematic adjustments to the model to achieve convergence. To simplify the imputation model by modeling the entire data set once rather than for each separate combination of treatment and gender. To pre-specify a list of random number seeds for use in the imputation process.

Approval for implementation of

Statistical Analysis Plan Addendum

Title: **A Phase 3, Randomized, Multicenter, Double-Blind, Placebo-Controlled, Efficacy, Safety and Dose-Response Study of Lofexidine in the Treatment of Opioid Withdrawal (Days 1-7) Followed by Open-Label, Variable Dose Lofexidine Treatment (Days 8-14)**

Reference: **USWM-LX1-3003-1/SAP Addendum/1**

Addendum No: **2**

Date effective: **22-Jan-2015**

Author: **Principal Statistician**

reviewer: **Principal Statistician**

Author's signature: _____ Date: _____

Reviewer's signature: _____ Date: _____

The above Statistical Analysis Plan Addendum has been reviewed and approved by the Sponsor:

Name of Approver:
Position: **Medical Director, US WorldMeds, LLC**

Signature: _____ Date: _____

Name of Approver:

Position: **Sr. Manager, Biometrics, US WorldMeds, LLC**

Signature: _____ Date: _____