NICHD/FHI 360*

A multi-center, single-blind, randomized clinical trial to compare two copper IUDs: Mona Lisa NT Cu380 Mini and ParaGard

Statistical Analysis Plan

Protocol No: CCN016

Sponsor approvals:

DocuSigned by:
Gregory Kopf
Signer Name: Gregory Kopf
Signing Reason: I approve this document
Signing Time: 16-Oct-2019 | 09:51 EDT
B519F96FF52D48259DAB846010E5BE14

Approver: Laneta Dorflinger (or designee) Date
Director, Product Development & Introduction, FHI 360

DocuSigned by:
David Hubacher
Signer Name: David Hubacher
Signing Reason: I approve this document
Signing Time: 16-Oct-2019 | 09:29 EDT
841182CC26954C8EBC567B87740C22EE

Approver: David Hubacher Date
Project Leader, Product Development & Introduction, FHI 360

DocuSigned by:
Douglas Taylor
Signer Name: Douglas Taylor
Signing Reason: I have read and understand the contents of this document
Signing Time: 18-Oct-2019 | 10:33 EDT
0B7E49A4E1A3472C8249494AB8976DB9

Approver: Mario Chen (or designee) Date
Director, Senior Scientist and Director, Biostatistics Division, FHI 360

* This plan was prepared by the lead biostatistician, Pai Lien Chen, of FHI 360. The plan received additional technical input from NICHD and HD. FHI 360 will be responsible for conducting data analyses for the final clinical study report for submission to the FDA and for primary manuscripts for publication.

Contents

1. INTRODUCTION

This is a multi-site, participant-blinded, randomized clinical trial. The study will randomize 1000 eligible participants in a 4:1 ratio to two different copper IUDs: 800 to Mona Lisa NT Cu380 Mini and 200 to ParaGard. Recruitment will be approximately 80% nulliparous and 20% parous women, and approximately 10% of women will be ≥ 36 years old.

The total duration of the study is expected to be five years. The total duration for each participant is expected to be approximately 37 months. Each woman will attend a baseline visit, followed by visits at 6 weeks, 3 months, 6 months, 12 months, 24 months and 37 months, with telephones calls at 9, 18 and 30 months. Each subject will participate for up to 37 months (or less if the product is removed or expelled). Participants will use a home pregnancy test 17 days after Exit Visit procedures or any time she suspects that she may be pregnant, and report the results to their respective sites via telephone.

The analysis plan provides additional details regarding the planned statistical analysis of the study protocol (version 2.0, 09May17). Tables, figures and listings shells will be developed and included as an appendix to a future version of the plan.

2. STUDY OBJECTIVES

The study objectives described in the protocol are as follows:

2.1. Primary Objective

To measure contraceptive efficacy of the IUD Mona Lisa NT Cu380 Mini.

2.2. Secondary Objectives

To compare between study products:

- continuation rates
- the incidence of failed IUD insertion, uterine perforation, and IUD expulsion (complete and partial)
- vaginal bleeding patterns
- pelvic pain and dysmenorrhea
- the incidence of other adverse effects
- pain with and shortly after insertion
- ease of IUD insertion
- overall product satisfaction

3. STUDY ENDPOINTS

3.1. Primary Endpoint

The primary endpoint for this study is pregnancy (contraceptive method failure).

3.2. Secondary Endpoints

- IUD continuation throughout 37 months of the study duration
- Failed IUD insertion, uterine perforation, or IUD expulsion (complete and partial)
- Vaginal bleeding patterns
- Pelvic pain and dysmenorrhea
- Other adverse effects
- Pain with and shortly after insertion
- Ease of IUD insertion
- Overall product satisfaction

4. DATA SAFETY AND EFFICACY MONITORING

An independent, autonomous Data and Safety Monitoring Board (DSMB) has been established by HD and the NIH (NICHD). The DSMB will conduct periodic reviews of study progress indicators, subject safety and pregnancy. It is expected that three (3) such reviews will occur. The Project Directors at NICHD and FHI 360 may also request additional DSMB reviews (e.g. should any other findings/issues pertaining to safety or efficacy emerge requiring DSMB review outside of the planned periodic meeting dates).

The DSMB Charter will be prepared separately by HD prior to the first DSMB meeting, and then modified as needed before subsequent meetings. The Charter will address the timing of each DSMB meeting, stopping rules relative to items including SAEs, pregnancy rates, expulsion rates, as well as detailed blinding and unblinding procedures. Additional information about the DSMB composition and review procedures will be contained in the Charter. As the primary objective of the study is non-comparative, no adjustment to confidence interval coverage error rates will be made to account for interim analyses when reporting pregnancy rate data. Similarly, no multiplicity adjustment will be made for secondary objectives.

5. ANALYSIS POPULATIONS

This analysis plan describes seven different analysis populations that are used for different purposes. The final statistical report will include an accounting of all persons screened, including the number enrolled, the number followed during the study, and other key study status indicators. The reason for excluding any data from any analysis population will be documented. The analysis population flow chart can be found in Appendix A.

The general criteria for inclusion/exclusion of participants in these seven populations are described below:

5.1. Screened Population

This population consists of all participants screened for the study, including screening failures, eligible participants not enrolled, and participants eventually enrolled.

5.2. Intent-to-treat (ITT) Population

This population includes all participants in the screened population who were enrolled and randomized. This is the most inclusive study population for analysis and includes participants who did not have a successful uterine sounding and other situations where an IUD was not inserted. Any participant who received the incorrect IUD (due to randomization or allocation error) will be included in the ITT population and analyzed as randomized. This analysis population will mainly be used for analyzing participants' baseline characteristics.

5.3. Treated Population

The Treated Population is a subset of ITT Population and includes only participants who advanced to uterine sounding, had their uterus successfully measured, and then underwent an attempted IUD insertion procedure. Inadequate uterine depth (at least 5.5cm) and/or medical contraindication (such as mucopurulent discharge discovered before/during/after sounding) are examples of situations that exclude participants from the Treated Population. Participants who received the incorrect IUD are included in this population and analyzed according to the treatment they actually received. This analysis population will mainly be used for evaluating insertion results and safety and other secondary endpoints observed during the insertion process. Participants' baseline characteristics will be analyzed only if more than 5% of the ITT Population are excluded from the Treated Population.

5.4. User Population

The User Population is a subset of Treated Population who had a successful IUD insertion. It includes participants who leave the clinic with the IUD in situ as well as those who have the product removed before leaving the clinic and exit from the study. This analysis population will mainly be used for analyzing discontinuation and safety and other secondary endpoints after a successful IUD insertion.

5.5. Per Protocol (PP) Population

This is a subset of the User population, excluding all subsequent data collected from participants with a documented protocol violation (either at baseline or during follow-up) that would prevent an accurate analysis of objectives. Participants with random allocation errors will also be excluded. This analysis population will mainly be used for sensitivity analyses of discontinuation, safety, and other secondary endpoints to evaluate the impact of protocol compliance, and only if more than 5% of the User Population are excluded from the PP Population.

5.6. Efficacy Analysis Population

The Efficacy Analysis Population is a subset of the User Population that includes women who provide at least one valid pregnancy analysis cycle as defined in Session 8.3.1. This analysis population will be used for analyzing the pregnancy endpoint. The subpopulation under 36 years old at enrollment who are in the Mona Lisa NT Cu380 Mini group and provide at least one cycle

with documented intercourse and no use of other methods will be the primary population for estimating efficacy (per FDA directives).

5.7. Lost to Follow-up Population

Lost to Follow-up Population is a subset of User Population who have the following experiences:

- left the clinic with the IUD in situ and
- had no documented IUD removal and
- whose last known use of the IUD was earlier than 37 months after IUD placement.

A subset of the Lost to Follow-up Population that has no follow-up data whatsoever is referred to as "Completely" Lost to Follow-up. Participants' baseline characteristics will be analyzed if more than 5% of the User Population are excluded from the Lost to Follow-up Population.

The following table summarizes various analyses conducted on each analysis population.

Table 1: Analyses performed on each analysis population

Population	Analysis			
	Baseline	Efficacy	IUD Discontinuation	Safety and other secondary endpoint
Screened				
ITT	✓			✓
Treated	✓*			✓ ǂ
User	✓		✓	✓
Per protocol	✓		✓***	✓***
Efficacy Analysis	✓	✓		
Lost to Follow-up	✓**			

* Analysis will be performed if > 5% of ITT Population is excluded.
** Analysis will be performed if > 5% of User Population is lost to follow-up.
*** Analysis will be performed if > 5% of User Population is excluded.
ǂ Analysis will be performed for safety and other secondary endpoints observed during the insertion process.

The following table summarizes various study participant status included in each analysis population.

Table 2: Study participant status included in each analysis population

Study participant status	Population						
	Screened	ITT	Treated	User	Per Protocol	Efficacy	LTF
Early situations							
Screen failures	✓						
Eligible but not enrolled	✓						
Enrolled	✓						
Protocol violation on enrollment	✓						
Post-randomization situations							
Medical contraindication*		✓					
No successful sounding of the uterus		✓					
Uterine depth < 5.5cm		✓					
No IUD placement attempt**		✓					
Randomization errors (as assigned)		✓					
Randomization errors (as treated)			✓	✓		✓	✓
IUD placement attempts							
Attempted but failed IUD placement		✓	✓				
Successful IUD placement but immediate removal per request by participant		✓	✓	✓			
Successful IUD placement and leaves clinic with IUD in situ		✓	✓	✓	✓	✓	✓
* e.g., discovering mucopurulent cervical discharge before/during/after sounding the uterus ** IUD placement attempt is defined as handling the product and attempting to place it							

6. GENERAL STATISTICAL ISSUES

All confidence intervals will be two-sided with 95% coverage. Likewise, all significant tests will be two-sided and at 0.05 significance level.

Potential outliers will be identified by looking at summary statistics and scatter plots as needed. All outliers will be verified as-such from the source documentation. If excluding one (or several) outlier(s) change(s) the interpretation of the data, two summaries and analyses will be done, one with and one without the outlier(s).

Missing data will be treated as missing at random unless available evidence indicated that missing data is informative. If a participant has missing data at a specific time point, data at all non-missing time points will be used and included in relevant analyses. For the primary analysis of pregnancy outcomes, participants who discontinue early (including participants who were lost to follow-up) will be included but will be censored on the date of their last known pregnancy status when still using the received study product.

All comparisons between treatment groups will be made without adjustment for multiple comparisons. All statistical analyses will be done using SAS® (SAS Institute Inc., Cary) Version 9.4 (or higher).

7. STATISTICAL ANALYSIS

7.1. Study Participant Disposition and Follow-up

Information about study participant disposition and follow-up during the study will be provided and summarized for each treatment group as follows:

- A flowchart will document the flow of participants through the study and relate this to the different analysis populations;
- Tables will summarize:
    - the number and percentage of participants screened, randomized, attempt insertion, first attempt insertion, the second attempt insertion, successful insertion, insertion failure, early method discontinuation and its reasons, and protocol deviation/violation by study site and overall, incidence of uterine perforation, and IUD expulsion (complete and partial);
    - the reasons for participant non-eligibility for each of the populations;
    - the number and percentage of participants included in each of the analysis populations by study treatment;
    - the number and percentage of total study visits completed, person-years completed, and participant end of study status as indicated in the End of Trial form.
    - the person-years with IUD in place by study group.

The similarity of participant end of study status between treatment groups will be evaluated using Mantel-Haenszel tests, stratified by site.

7.2. Analysis of Baseline Data

Baseline variables will be summarized for the ITT, User, Per Protocol and Efficacy Analysis populations, and for the Treated and Lost to Follow-Up Populations under the condition stated in Section 5. The following baseline measures collected at the screening and enrolment visits will be summarized by treatment group:

- Demographic data: age, ethnicity, race, parity, marital status, level of education
- Vital signs
- Substance use: tobacco use, alcohol use, marijuana use and illicit drug use
- Medical history including prior and concomitant medications
- Gynecological history

- Infection history (Pre-treatment signs and symptoms of: (urinary tract infection, yeast infection, bacterial vaginosis, pelvic infection, Endometriosis, and STIs)
- Contraceptive history
- Pregnancy history

Data will be presented in summary tables by treatment group, site and overall. Categorical variables, and continuous variables that have been categorized at discrete levels, will be summarized by frequencies and percentages. Continuous variables will be summarized by means, standard deviations, medians, interquartile ranges, minima and maxima.

7.3. Analysis of Primary Endpoint - Efficacy Data

Efficacy data will be analyzed on the Efficacy Analysis Population. The primary efficacy analysis will include those women under 36 years old at time of enrollment who are in the Mona Lisa NT Cu380 Mini group.

Additional efficacy analyses will include the following groups

- nulliparous women under 36 years old at time of enrollment who are in the Mona Lisa NT Cu380 Mini group
- parous women under 36 years old at time of enrollment who are in the Mona Lisa NT Cu380 Mini group
- all women who are in the Mona Lisa NT Cu380 Mini group, regardless of age and parity status
- nulliparous women under 36 years old at time of enrollment who are in the ParaGard group
- all women who are in the ParaGard group, regardless of age or parity status.

No formal statistical comparisons will be conducted between above groups.

7.3.1 Definition of variables

Pregnancy: The endpoint of the primary objective is incident pregnancy. The following data will be examined to determine whether a pregnancy has occurred:

- Positive pregnancy test result recorded on the Pregnancy Test, Pregnancy Notification, or Pregnancy Outcomes forms;
- Pregnancy otherwise reported on Adverse Event, IUD Removal or End of Trial forms.

Estimated Date of Conception (EDC): The Investigator will estimate the date of conception based upon results of transvaginal ultrasound examination, quantitative β-hCG determination, last menstrual period (per daily diary information), pelvic and/or abdominal examination, and pregnancy outcome.

Without looking at the treatment group assignment in the LiveTrial eCRF, the Medical Monitors will review the assigned date of conception and other relevant data to confirm the date of conception to be used for analysis. If the Medical Monitors assign a different date during this adjudication process, their reason for assigning a different date will be documented.

On-treatment Pregnancy: An on-treatment pregnancy is a pregnancy in which the EDC is 1) on or after the date of IUD insertion and 2) on or before 7 days of post IUD removal. In addition, a "worse case" scenario will be used that considers pregnancies as "on-treatment" if the EDC or date of expulsion is not known.

Study Cycle (28-day equivalent): Every 28 study days with information about whether the participant used other contraceptive methods within those 28 study days will be considered as a study cycle (e.g., study day 28 will be considered the end of cycle 1, and day 56 for the end of cycle 2, etc., so long as information on use of other methods is documented for those periods). Note: cycles where conception occurred or where a participant had the IUD removed will be considered 'study cycles' (even if the EDC or removal did not occur on the 28th day), so long as there is information on other contraceptive methods use up until the time of conception defined as an on-treatment pregnancy or IUD removal.

Study Cycles (described previously) with no diary data will be excluded from the pregnancy analysis. Two types of pregnancy analysis cycle will be used for the efficacy analysis:

Primary Pregnancy Analysis Cycle: A primary pregnancy analysis cycle is a study cycle that

1. contains an on-treatment pregnancy, or
2. has both these characteristics:
   a. reported sexual intercourse in the cycle and
   b. no reported use of other contraceptives in the cycle

The primary pregnancy analysis cycles will be used for the primary efficacy analysis.

Secondary Pregnancy Analysis Cycle: A secondary pregnancy analysis cycle is a study cycle that

1. contains an on-treatment pregnancy, or
2. has no reported use of other contraceptives in the cycle

The secondary pregnancy analysis cycles will be used for the sensitivity efficacy analysis.

7.3.2 Statistical methods

The on-treatment pregnancy rate will be estimated using the primary pregnancy analysis cycles based on the Pearl Index (pregnancies per 100 woman-years of follow-up)

$$Pearl\ Index = \frac{Number\ of\ Pregnancies}{Total\ Number\ of\ Pregnancy\ Analysis\ Cycles} \times 1300$$

and its exact 95% confidence interval (CI) derived using a Poisson distribution assumption as

$$Lower\ 95\%\ CI\ (Pearl\ Index) = \frac{\chi^2_{0.025,\ df=v1}}{2 \times Total\ Number\ of\ Pregnancy\ Analysis\ Cycles} \times 1300$$

and

$$Upper\ 95\%\ CI\ (Pearl\ Index) = \frac{\chi^2_{0.975,\ df=v2}}{2\ \times\ Total\ Number\ of\ Pregnancy\ Analysis\ Cycles} \times 1300$$

where $v1$=2 × Number of Pregnancies and $v2$= 2 × (Number of Pregnancies +1).
Analysis will be done for each year of use and cumulatively through year three.

In addition, the typical use Pearl Index estimated by including all Study Cycles in the denominator of Pearl Index estimator will be reported. Similarly, by using all Study Cycles, the Kaplan-Meier product limit method with cycle as the "time" unit will also be used to summarize the typical use cumulative probability of pregnancy through the study follow-up period. Ninety five percent confidence intervals about the estimates of the cumulative probability of pregnancy will be provided, using the method of Peto et al. to calculate the standard error for the probabilities. Moreover, two additional sensitivity analyses will be considered to handle those cycles using other contraceptive methods in the Kaplan-Meier product limit method. The first approach is to exclude those cycles from the analysis and the remaining cycles will be renumbered to provide contiguous cycles (e.g., if Cycle 3 is excluded, Cycle 4 will be treated as Cycle 3, and so forth). The second approach is to allow participants to leave and re-enter the risk set of pregnancy of the Kaplan-Meier product limit method if they had a cycle using other contraceptive methods between two primary pregnancy analysis cycles.

In a separate analysis, these same approaches will be applied to analyze pregnancy using secondary pregnancy analysis cycles.

7.4. Analysis of Secondary Endpoint - IUD Discontinuations

The User Population will be used for IUD discontinuation analysis. Descriptive statistics will be calculated by pooling data across sites. In addition, a supporting analysis will be performed using the Per Protocol Population if more than 5% of participants in the User Population are excluded.

7.4.1 Definition of variables

Discontinuation Event: The endpoint of discontinuation is time to IUD removal for any reason or exit from the trial at the end of the study.

Reasons for discontinuation[1] are recorded in the IUD Removal and End of Trial sections of the data capture system. The following table shows the origin of the data and the endpoint classification that will be used in the analysis:

[1] A discontinuation event might have more than one reason.

IUD Removal Section	End of Trial Section	Endpoint classification*
Completed study participation	Completed Study	Completed study
Not sexually active anymore		Not sexually active anymore
Wants to get pregnant		Desire for pregnancy
Increased menstrual blood loss	Adverse event**	Increased menstrual blood loss
Intermenstrual bleeding/spotting	Adverse event**	Intermenstrual bleeding/spotting
Intermenstrual pain side effects	Adverse event**	Intermenstrual pain side effects
Menstrual pain	Adverse event**	Menstrual pain
IUD was completely expelled	Spontaneous complete expulsion	Spontaneous complete expulsion
IUD was malpositioned and had to be removed by the clinician	Malpositioned IUD but not expelled	Malpositioned IUD removed by clinician (not considered partial expulsion)
IUD perforation		IUD perforation
Method failure (pregnancy with IUD in situ)	Pregnancy	Pregnancy with IUD in situ
Other	Other	Other***
	Adverse event (without pre-coded event from IUD removal CRF)	Other adverse event
	Withdrew Consent	Withdrew Consent
	Protocol Violation	Protocol Violation
	Investigator/Sponsor Decision	Investigator/Sponsor Decision
	Use of Prohibited Medication	Use of Prohibited Medication
	Uterus not sounded or <5.5	Uterus not sounded or <5.5
	Failed IUD placement attempt	Failed IUD placement attempt
	Spontaneous partial expulsion (IUD in cervix or at os)	Spontaneous partial expulsion (IUD in cervix or at os)
	Accidental self-removal/dislodgement	Accidental self-removal/dislodgement
	Lost to Follow-up	Lost to Follow-up

* For presentation in tables, additional groupings may be created
** If corresponding, pre-coded IUD removal reason was also reported
*** Other reasons will be further subdivided if enough cases exist to combine them with a new label

7.4.2 Statistical methods

The frequency and the percentage of women who complete the study or who leave the study for specific reasons, including loss to follow-up, will be tabulated. The percentage of discontinuation for each of the categories outlined above will be provided by study groups for the entire population. The duration of IUD exposure will be presented in summary tables using means, standard deviations, medians, interquartile ranges, minima and maxima by treatment group, site and overall.

For early discontinuation probabilities and comparative analyses, participants who complete full participation will be classified as censored. The product limit method (Kaplan and Meier, 1958) will be used to estimate 12-month, 24-month and 36-month cumulative probabilities (and associated 95% confidence intervals) for IUD discontinuation. A stratified logrank test by study site will be used to compare the discontinuation probabilities between the two groups of women, separately over the 12 and 36 months study periods. A non-stratified logrank test will be used if too few events (i.e., < 5) occur in one or more study sites.

As a supporting analysis, Cox's proportional hazards regression (Cox, 1972) will be used to assess differences in the risk of discontinuation between study groups, controlling for covariates. Suspected prognostic factors such as sociodemographic characteristics (i.e., age, education, relationship status), reproductive history (i.e., parity) and previous contraceptive experience (e.g., hormonal method use) will also be considered if they are significantly related to the outcome at the alpha=0.10 level in analyses of groups combined.

The assumption of proportional hazards will be assessed graphically and by testing for the significance of interaction terms between time (or functions of time) and other variables in the model if there is no time-dependent covariate in the model. If the proportional hazard assumption is violated, further data transformations or other methods will be considered. All interactions will be tested at the alpha=0.05 level.

To measure time to discontinuation due to a specific reason, such as menstrual pain, the same analysis strategy will be used. Competing reasons for removal will be considered independent and censored such as IUD expulsion, failed IUD insertion, uterine perforation, pregnancy, personal/non-medical reasons, and other medically-related reasons.

Specific reasons for discontinuation will also be grouped into broader categories in further exploratory analyses. “Discontinuation due to side effects” will include bleeding and pain reasons. Personal (non-medical) reasons will include not sexually active, wants to get pregnant, and withdrew consent for study fatigue (burden) issues. The same analysis strategy will be used for these exploratory analyses.

7.5. Analysis of Secondary Endpoints - Diary Data

The User Population will be used for analysis of diary data. Descriptive statistics will be calculated by pooling data across sites. In addition, a supporting analysis will be performed using the Per Protocol Populations if more than 5% of participants in the User Population are excluded.

7.5.1 Definition of variables

Evaluable Cycle: A non-pregnant cycle (28-day equivalent) with complete bleeding diary data within the cycle.

Partial Evaluable Cycle: A non-pregnant cycle (28-day equivalent) with incomplete bleeding diary data within the cycle (e.g., only 10 of 28 days bleeding diary data are available within a cycle) either due to missing data or discontinuation.

7.5.2 Statistical methods

The descriptive statistics of per-subject mean and median number of evaluable cycles and the proportion of evaluable cycles that include any heavy bleeding, a prolonged bleeding episode (e.g., greater than 14 days), amenorrhea (absence of all bleeding and spotting), or bleeding in more than 1/3 of days in the evaluable cycle will be tabulated. In addition, descriptive statistics of bleeding diary data as described by Mishell, et al (2007) such as the mean, median, and range of the total number of heavy, moderate, light bleeding, or spotting days per evaluable cycle will be reported without considering the feature of repeated cycles from same study participants.

To incorporate the nature of repeated measures and partial evaluable cycles, a multinomial regression with generalized estimating equation (GEE) approach will be used to estimate and compare the mean, and total number of heavy, moderate, light bleeding, or spotting days per 28-day cycle equivalent between treatment groups (Liang and Zeger, 1986, and Lipsitz et al., 1994).

Note that if 1) there are sizeable participants discontinued early (e.g., 20% of IUD removal) due to increased menstrual blood loss or due to intermenstrual bleeding/spotting, and 2) the proportions of these two reasons of discontinuation are notably different between treatment groups (P-value < 0.1), the analyses of observed bleeding diary data are likely to be biased due to informative censoring and attrition (Hubacher et al., 2009). In this event, subgroup supporting analyses based on participants' discontinuation reasons related and non-related to blood loss/bleeding/spotting will be conducted. A stratified bleeding diary supporting analysis based on the same blood loss/bleeding/spotting related reasons will also be performed to compare the results from the above analyses. The findings of these supportive analyses will be interpreted cautiously.

For the analysis of pelvic pain data, the proportions of participants experiencing any pelvic pain during 28-day cycle, pelvic pain during period, outside of period, and both during and outside of period will be calculated using binomial distribution for each group, and compared between groups using chi-square tests. In addition, generalized linear models with GEE for repeated pelvic pain outcomes will be used to further explore the differences in frequencies of pelvic pain between study groups.

7.6. Analysis of Other Secondary Endpoints and Safety Data

The ITT population will be used for analyses involving pre-IUD insertion events, and the Treated Population will be used for the safety and other secondary endpoints analyses during the insertion process, and the User Population will be used for the safety and other secondary endpoints analyses after a successful IUD insertion. Descriptive statistics will be calculated by pooling data across

sites. In addition, a supporting analysis will be performed using the Per Protocol Populations if more than 5% of participants in the User Population are excluded.

7.6.1 Exiting the trial before attempting IUD insertion

The ITT Population will be used for estimating the incidence of exiting the trial before attempting an IUD insertion. Reasons for exiting the trial before attempting IUD insertion include:

- Could not sound the uterus
- Shallow uterine depth (less than 5.5cm)
- Medical contraindication discovered post-randomization but before IUD insertion attempt

The frequency and the percentage of women who exited the trial for different reasons before IUD insertion will be tabulated by treatment group and by study site. The proportions of participants experiencing these events will be calculated with corresponding exact confidence intervals for each group, and the difference between the groups will be compared using the Cochran-Mantel-Haenszel tests.

7.6.2 IUD successful insertion, failed IUD insertion, and incident adverse events during the insertion process

The Treated Population will be used for evaluating the incidence of successful IUD insertion and incident adverse events (e.g., perforation) occurred during the insertion process. The frequency and the percentage of women who experienced IUD placed successfully at the first attempt, at the second attempt, and incident adverse events occurred during the insertion process will be tabulated by treatment group and by study site. The proportions of participants experiencing these events will be calculated with corresponding exact confidence intervals for each group, and the difference between the groups will be compared using the Cochran-Mantel-Haenszel tests.

7.6.3 IUD expulsions, incident adverse events, and other events after a successful IUD insertion

The User Population will be used for analysis of IUD expulsions (complete and partial), incident adverse events, and other events. For each adverse effect (e.g., PID or uterine perforation), two measure units will be used in this analysis: 1) number of participants experiencing at least one adverse effect and 2) number of the AE reporting during the study. First, the proportions of participants experiencing the AE will be calculated using exact confidence intervals for each group and compared between groups using Fisher's exact tests or the Cochran-Mantel-Haenszel tests. Secondly, generalized linear models with GEE for repeated AE outcomes will be used to further explore the differences in the incidence of AEs between study groups.

Since participants with longer study duration will have more chance to experience adverse effects, the analysis will control for duration since insertion. Note that if the duration in study depends on adverse effects, then the drop out will be non-ignorable. Supporting stratified analyses by study duration or joint modeling of adverse effects and study duration will be considered as sensitivity analyses (Tsiatis and Davidian, 2004).

All expulsions (partial or complete) and accidental self-removals/dislodgements of the IUD will be examined in relationship to type of menstrual hygiene product used by the participant. Participants may use a variety of products over the course of the study. The mutually exclusive categories for analysis will include the following: Any cup use (exclusive cup use or cup use plus any other hygiene product), tampon use (no cup use but tampons used exclusively or tampon use plus other hygiene products), pad (no cup use, no tampon use, but pads used perhaps with some products other than tampons), other products (no cup use, no tampon use, no pad use). A sub-analysis will examine if the self-removal/dislodgement was known to occur during manipulation of the hygiene product. Chi-square tests and/or Fisher's exact tests will be used to examine these associations by IUD group. The proportions of participants with expulsions will be calculated using exact confidence intervals for each type of menstrual hygiene product in each group. Cochran-Mantel-Haenszel tests will be used as appropriate to compare IUD groups.

7.6.4 Pain on insertion day and pain in first 7 days post insertion

The level of pain for participants will be measured by visual analog scale at 9 time points:

- At baseline (after speculum placement but before IUD insertion)
- Immediately post IUD insertion
- 10 minutes after insertion
- Daily from Day 1 to Day 6 post IUD insertion (these last six measures are completed by the participant at home).

Thus, the Treated Population will be used for the baseline (it is anticipated that unsuccessful insertions will have a pain measurement that describes the level of pain from the attempted procedure). The User Population will include all 9 measures. The level of pain will be summarized by means, standard deviations, medians, interquartile ranges, minimum and maximum by each time point, site and study group. Changes in pain from baseline will be described by the summary statistics (mean, median, inter-quartile, etc.). T-tests or the Mann-Whitney tests will be used to compare the changes in pain between groups at each time point of post insertion. A generalized linear model will be used to evaluate the time trend of changes in pain between groups and to compare changes in pain between groups across all post insertion time points.

7.6.5 Ease of product insertion and removal

Participants in the Treated population will be used to assess ease of product insertion and Participants in the User population will be used to assess ease of product removal. Clinicians classify ease in one of four ways: easy, somewhat easy, somewhat hard, or hard. The results will be summarized in frequency and percentage by site, and compared between groups using two-sided Cochran-Mantel-Haenszel chi-square tests.

7.6.7 Characteristics of last menstrual period and product satisfaction

Acceptability data from the User Population on the last menstrual cycle and product satisfaction overall are derived from subject bleeding and satisfaction questionnaires and participant interviews conducted during the six clinic visits and during the three phone calls over the study period.

Data on bleeding and pelvic pain include number of days of bleeding, volume of blood loss (light, medium, heavy), acceptability of blood loss, level of cramping, and intermenstrual bleeding/spotting/cramping. Summaries will be described in contingency tables and displayed graphically in bar charts, by study visit and overall. Chi-squared tests for binary and ordinal outcome data (e.g., light, moderate, heavy blood loss) will be used to compare treatment groups at each visit/phone call. Generalized linear models with GEE will be used to compare the two treatment groups on these measures over time.

As noted in the previous section, if evidence of informative censoring exists due to participants' discontinuation reasons related to bleeding or pain, similar subgroup and stratified supporting analyses will be used to evaluate the impact of censoring mechanism on the results of the bleeding and pain safety data analysis.

Product satisfaction includes three measures: level of satisfaction with the study IUD, whether the participant would recommend the study IUD to others, and how the product compared to expectations. Each measure will be summarized in frequency and percentage by visit and site. A generalized linear model will be used to evaluate the time trend of product satisfaction between groups and to compare the means of changes of satisfaction between groups across all post insertion time points.

7.6.7 STI/RTIs assessment results

STI/RTIs related data collected from Treated population on Study Procedure (SP) questionnaires will be summarized for each visit in tabular form by IUD and across study sites. The summary statistics include frequencies and proportions of participants who had pelvic exam assessment, STI/RTI testing, and pap smear test and the results of all of those tests. Exploratory comparison between groups will be conducted by using Fisher exact tests or chi-squared tests.

8. REFERENCES

Cox DR "Regression Models and Life Tables (with discussion)", *Journal of the Royal Statistical Society,* B34. 1972;187-220.

Hubacher, D., Chen, P-L, Park, S. (2009) Side effects from the copper IUD: do they decrease over time? Contraception, 79: 356–362

Kaplan EL and Meier P. "Nonparametrics Estimation from Incomplete Observations", *Journal of American Statistical Association*. 1958; 53:457-481.

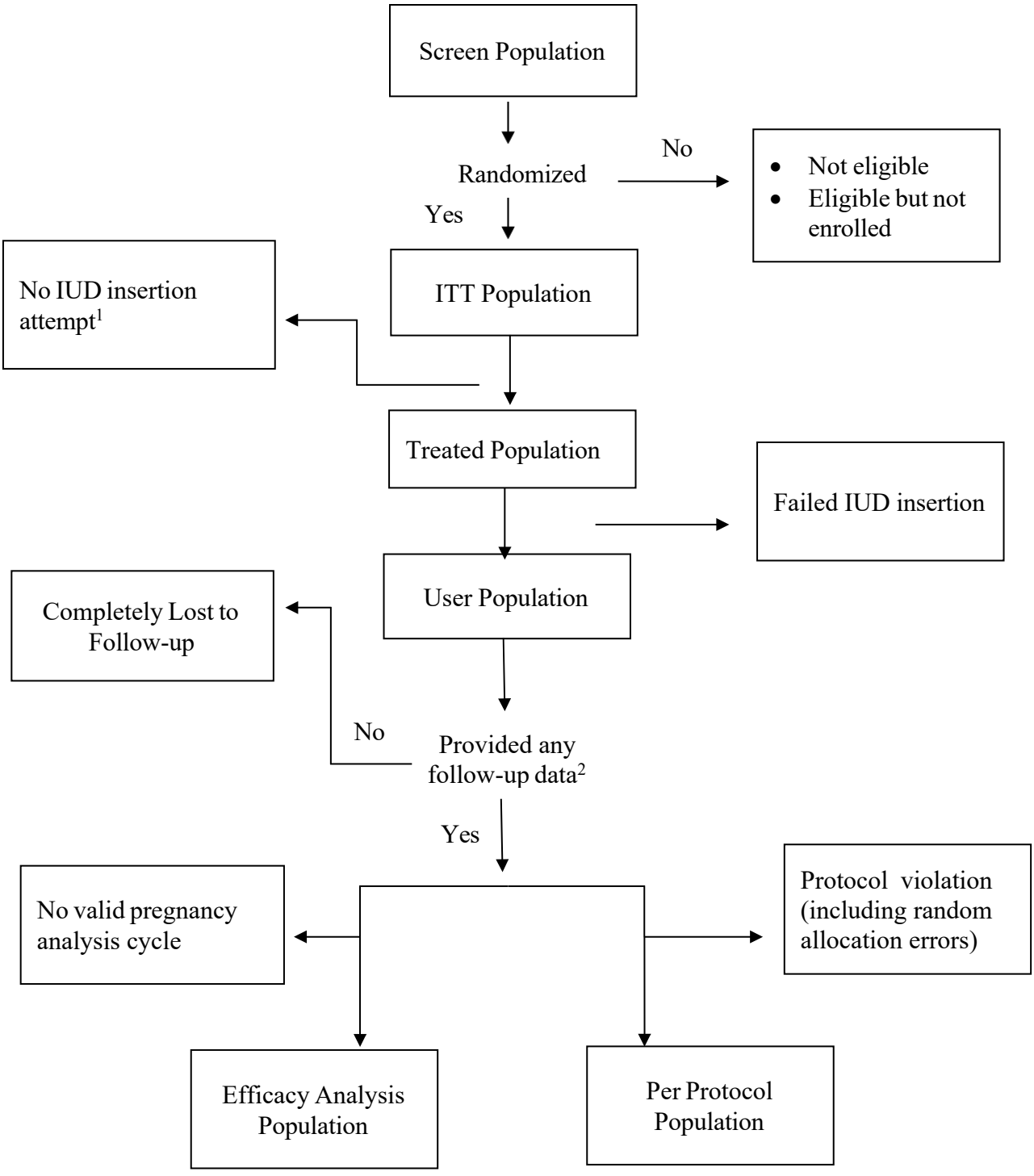
Liang, KY and Zeger, SL. Longitudinal data analysis using generalized linear models, Biometrika 1986; 73: 13-22.

Lipsitz, SR, Kim, K and Zhao, L. Analysis of repeated categorical data using generalized estimation equations, Stat Med. 1994; 13: 1149-1163.

Mishell DR Jr, Guillebaud J, Westhoff C, Nelson AL, Kaunitz AM, Trussell J, Davis AJ. Recommendations for standardization of data collection and analysis of bleeding in combined hormone contraceptive trials. Contraception 2007; 75(1): 11-1.

Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: An overview. Stat Sinica. 2004;14(3):809–34.

Appendix A: Figure 1. Analysis Population Flow Diagram



[1]No IUD insertion attempt due to unable to sound the uterus prior to attempting the IUD insertion, uterus did not sound to a minimum of 5.5mm, medical conditions, or other reasons.
[2] Loss to Follow-up Population includes participants who provided no follow-up data (completely LTF) or who provided some follow-up data but without documented IUD removal and whose last known use of the IUD was earlier than 37 months after IUD placement.

REVISION HISTORY

Document Version	Changes	Document Date
Version 1.0	First approved version.	18APR2018
Version 2.0	Incorporated several comments from external reviewers from NICHD and Health Decisions. 1. Included menstrual hygiene product used by the participant into IUD expulsion analysis. 2. Added analysis section for STI/RTIs assessment results during the study period.	14JUN2019
Version 3.0	The self-imposed blinding is no longer necessary and therefore, has been removed.	08OOCT2019