Antibacterial Resistance Leadership Group (ARLG)

Performance of Nucleic Acid Amplification Tests for the Detection of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* in Extragenital Sites

GC Statistical Analysis Plan
Version 2.0

Scott Evans, Lauren Komarow, Thuy Tran

Harvard T. H. Chan School of Public Health

Updated: April 12, 2018

(Version 1.0 finalized April 10, 2017)

Pro00073558

NCT02870101

INTRODUCTION

This document describes the content proposed for the statistical monitoring and primary statistical analysis of the study titled "Performance of Nucleic Acid Amplification Tests for the Detection of Neisseria gonorrhoeae (NG) and Chlamydia trachomatis (CT) in Extragenital Sites." The focus of this analysis will include the primary and secondary objectives to evaluate diagnostic accuracy of the following three nucleic acid amplification test (NAAT) platforms for the detection of NG and CT from the extragenital sites of the pharynx and the rectum:

- Company 1 Assay [Company 1]
- Company 2 Assay [Company 2]
- Company 3 Assay [Company 3]

A subset of these analyses (as described herein) will form the basis of reports provided to the independent statistician as part of independent interim study monitoring while the study is ongoing. Therefore, this analysis plan includes a description of the key analyses which might lead to the modification of the study sample size, and hence also forms the core of any presentation or publication used to disseminate the primary conclusions of the study.

THE CURRENT STATISTICAL ANALYSIS PLAN (VERSION 2.0) IS BASED ON PROTOCOL VERSION 4.0 (DATED DECEMBER 4, 2017). PROTOCOL HISTORY

- Protocol Version 1.0 (April 25, 2016, Protocol ARLG_pNAAT-Yr3)
- Protocol Version 2.0 (August 1, 2016, Protocol ARLG_pNAAT-Yr3)
- **Protocol Version 3.0 (September 5, 2017, Protocol ARLG_pNAAT-Yr3)**
- **Protocol Version 4.0 (December 4, 2017, Protocol ARLG_pNAAT-Yr3)**

REVISIONS TO STATISTICAL ANALYSIS PLAN, VERSION 1.0 (FINALIZED APRIL 12, 2017):

- **Master-GC: Statistical Analysis Plan, Version 1.0 (April 10, 2017)**
  - **SAP was based on Protocol Version 2.0 (August 1, 2016, Protocol ARLG_pNAAT-Yr3)**

- **Master-GC: Statistical Analysis Plan, Version 2.0 (April 12, 2018)**
  - **SAP is based on Protocol Version 4.0 (December 4, 2017, Protocol ARLG_pNAAT-Yr3)**
  - **Updates include:**
    - **Final Analysis Considerations, pg 9**
      - **Clarification regarding final analysis population.**
      - **Includes exclusion of 167 participants from the final analysis population due to protocol deviations.**
    - **Swab Collection Completeness and Complications, pg 11**
      - **Clarification regarding swabs not collected per assigned order.**
    - **Swab Testing Completeness and Monitoring, pg 12**
      - **Added table to examine completeness of tiebreaker testing.**
    - **NAAT Test Results, pg 12**
      - **The "Possible Test Results" and "Notes" columns updated to reflect changes from Protocol Version 3.**
      - **Footnotes added to Tiebreaker assay section.**
    - **Anatomic Site Infection Status (ASIS) Determination, pg. 13**
      - **Row 42: Anatomic Site Infection Status was changed from "Indeterminate" to "Invalid, remove from analysis" to reflect changes from Protocol Version 3.**
    - **Table 9: Frequency of comparator NAAT result and tiebreaker result combinations by anatomical site and organism for test under consideration XX, pg 17**
      - **Added "2" to second column (omitted in the original SAP)**
    - **Final Analysis section, pg. 19**
      - **Section title updated from "Endpoint Definitions" to "Outcome Measures"**

- **Additional text clarifying data for final analysis was added to the Final Analysis Considerations, pg19.**
- **The following analysis tables were added to the Primary Analyses section, pg 19.**
    - **Frequency of re-tests by clinic and platform**
    - **Listing and frequency of re-test results by platform**
    - **Observed infection rate by clinic.**
- **Table 12: Calculation of the positive and negative percent agreement: Company 2 Assay (Company 2) and Company 3 Assay (Company 3), pg 21**
    - **Removed incorrectly labeled "(95% CI)" from PPA and NPA column headers**
    - **Added cell F to PPA denominator (was omitted in SAP, v1.0)**
    - **Added cell D to NPA denominator (was omitted in SAP, v1.0)**
- **Table 13: Result of Test under Consideration versus ASIS: Company 1 Assay (Company 1) and Company 3 Assay (Company 3), pg. 21**
    - **For clarity, added "Detected" and "Not detected" to "Results of Test under Consideration" rows.**
    - **Removed "Invalid" from "No result" row to reflect changes from Protocol Version 3.**
- **Table 14: Calculation of the positive and negative percent agreement: Company 2 Assay (Company 1) and for the Company 3 Assay (Company 3), pg 21**
    - **Added cell D to PPA denominator (was omitted in SAP, v1.0)**
    - **Added cell F to NPA denominator (was omitted in SAP, v1.0)**

LIST OF ABBREVIATIONS

ARLG	Antibacterial Resistance Leadership Group
ASIS	Anatomic Site Infection Status
CDC	Center for Disease Control and Prevention
CFR	Code of Federal Regulations
CT	*Chlamydia trachomatis*
DCRI	Duke Clinical Research Institute
DNA	Deoxyribonucleic acid
FDA	Food and Drug Administration
HIV	Human immunodeficiency virus
ICMJE	International Committee of Medical Journal Editors
ID	Identification
IFU	Instructions for Use
IRB	Institutional Review Board
ISRC	Independent Study Review Committee
LGBT	Lesbian, gay, bisexual and transgender
NAAT	Nucleic acid amplification test
NR	This means no test was run and there is no result
NPA	Negative percent agreement
NPV	Negative predictive value
NG	*Neisseria gonorrhoeae*
NIH	National Institutes of Health
OHRP	Office of Human Research Protections
PPA	Positive percent agreement
PPV	Positive predictive value
PI	Principal Investigator
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SDMC	Statistical and Data Monitoring Center
STD	Sexually transmitted diseases
WHO	World Health Organization

STUDY SCHEMA AND OBJECTIVES

DESIGN	A cross-sectional, single visit study to evaluate the diagnostic accuracy of three nucleic acid amplification tests (NAATs) for detection of *Neisseria gonorrhoeae* and *Chlamydia trachomatis* from a set of four swabs each collected from both the pharyngeal and rectal sites, respectively.
DURATION	This is a single visit study. It is estimated that the study will take between 6 and12 months after enrollment of the first participant to fully enroll. Note: Study duration was updated to 12 to 24 months during the study.
SAMPLE SIZE	Up to 2,500 participants Note: Sample size was increased to up to 3000 participants per recommended changes by the independent statistician during the study.
POPULATION	Symptomatic or asymptomatic male, female, or transgender participants: o Who are patients attending a participating clinic for evaluation of sexually transmitted disease (STD) , and o ≥18 years of age at date of screening, and o Able and willing to provide informed consent, and o Willing to comply with study procedures, including collection of 4 swabs each from the pharynx and rectum for NG and CT testing.
NUMBER OF SITES	Up to 10
STRATIFICATION	Randomization of the swab order will not be stratified. Swabs will be collected from all participants at each extragenital site.
DIAGNOSTICS	While the diagnostic accuracy of three NAATs will be evaluated, a total of four swab kits will be collected from both the pharyngeal and rectal sites, respectively (8 swabs in total). The swab kits associated with the NAATs under investigation will be tested using the corresponding laboratory test assay and test system as defined in **Table 1**. The remaining NAAT will only be performed in cases of discordant results and serve as the tiebreaker assay. There will be no evaluation for diagnostic accuracy for the tie breaker NAAT.

1.1. PRIMARY OBJECTIVES

1.1.1. For each NAAT under evaluation, estimate the positive percent agreement (PPA) and negative percent agreement (NPA) for detection of the organism and extragenital site combinations listed below.

- *Neisseria gonorrhoeae* in rectal swabs
- *Neisseria gonorrhoeae* in pharyngeal swabs
- *Chlamydia trachomatis* in rectal swabs
- *Chlamydia trachomatis* in pharyngeal swabs

1.2. SECONDARY OBJECTIVES

1.2.1. Global analyses

For each NAAT under evaluation, positive predictive values (PPVs), negative predictive values (NPVs), positive likelihood ratios, and negative likelihood ratios will be calculated for detection of the organism and extragenital site combinations listed below.

- *Neisseria gonorrhoeae* in rectal swabs
- *Neisseria gonorrhoeae* in pharyngeal swabs
- *Chlamydia trachomatis* in rectal swabs
- *Chlamydia trachomatis* in pharyngeal swabs

1.2.2. Subgroup analyses

For each NAAT under evaluation, to estimate the PPAs, NPAs, PPVs, and NPVs for detection of NG and CT from rectal and pharyngeal swab specimens by sex and by anatomic site-specific symptom status.

1.3. EXPLORATORY OBJECTIVES

1.3.1. Application of developed diagnostic benefit:risk analyses, including BED-FRAME methodologies[1].

Table 1: List of Corresponding Swab Collection Kits, Laboratory Assay and Laboratory Machine

Swab Collection Kit Name	Corresponding Laboratory Test Assay	Corresponding Laboratory Test System (Machine)
NAATs Under Investigation (as defined in Section 1.2 of Protocol)		
Company 1 Specimen Collection Kit	Company 1 assay	Company 1 System
Company 2 Specimen Collection Kit	Company 2 assay	Company 2 system
Company 3 Specimen Collection Kit	Company 3 assay	Company 3 System
Tiebreaker Assay (as defined in Section 1.2 of Protocol)		
Tiebreaker Specimen Collection Kit	Tiebreaker assay	Tiebreaker system

[1] Evans, S.R., Pennello, G., Pantoja-Galicia, N., Jiang, H., Hujer, A.M., Hujer, K.M., Manca, C., Hill, C., Jacobs, M.R., Chen, L. and Patel, R., 2016. Benefit-risk evaluation for diagnostics: a framework (BED-FRAME). Clinical Infectious Diseases, p.ciw329.

ANALYSIS PLAN OVERVIEW

A. General Analysis Considerations

Data summaries and analyses will be presented overall and by sex or, where appropriate, by anatomic site and organism.

Because the Master-GC study is a single visit study (i.e, no follow-up), the only date captured in this study is the date of a participant's clinic visit (recorded on the DEMOG case report form) and represents both the "study entry" and "off study" dates.

To ensure participant confidentiality, any listing of individual patient level data will be minimized as much as possible during study monitoring and for interim and final analyses. If such data are provided, they will be indexed with a unique blinded identifier or with identifiers removed. Study dates will not be presented.

Of note, while participants may need to seek additional care beyond this single visit, this is beyond the purview of the study. Similarly, this study will not be used to inform or determine treatment practices.

B. Statistical Monitoring Considerations

Routine statistical monitoring of accrual, study conduct, and completeness of swab and data collection will be conducted by the Harvard statistical team. A monitoring report summarizing these components will be distributed by the Harvard team on a bi-monthly basis until completion of study accrual; see sections listed below for details. The report will be distributed to the core protocol team (see Section E for details). It is planned at this time that distribution of the first monitoring report will occur two weeks after transfer of data from DCRI to Harvard has been tested and successfully confirmed by both groups.

1. Study Population (Section G, Part 1, pg 10)
2. Study Status (Section G, Part 2, pg 11)
3. Swab Collection Completeness and Monitoring (Section G, Part 3, pg 11)
4. Swab Testing Completeness (Section G, Part 4, pg 12)

Additional monitoring and querying of all data will be conducted concurrently with the statistical monitoring report to ensure cross-form consistency and data quality.

C. Interim Analysis Considerations

The Master-GC study will undergo interim review every 500 participants or every 3 months after the enrollment of the first participant (whichever occurs first) by an independent statistician who will not have an association with the protocol or device companies.

The study cannot be stopped at the interim analysis for reaching regulatory goals in order to preserve error rates / coverage probability and to ensure enough data for subgroup analyses. Since the trial cannot be stopped for attainment of the regulatory goal, no adjustment to confidence levels is necessary.

Infection rates will be evaluated by the independent statistician to determine whether sample size adjustments are warranted to ensure sufficient number of infected participants to estimate PPAs with desired precision. The sample size will not be adjusted based on the observed PPAs and NPAs (which will not be reviewed while the study is ongoing). If infection is more prevalent than expected, a smaller sample size may be accepted. If infection is rarer than anticipated, then increases to sample size will be considered. Detailed analysis considerations are provided in Section 5.3.

The independent statistician will also monitor study accrual and review endpoint evaluability with particular focus on the frequency of tests with equivocal results, invalid results, or no results for each platform. Sample size adjustments may also be considered if there is evidence that results categorized as equivocal, invalid, or no results will impact study accrual time and/or data analyses.

Two separate analysis reports will be prepared and distributed for each interim review. The summary below provides the sections of this analysis plan that are relevant to each report; see sections for details.

Open Administrative Report: This report will be distributed to the independent statistician, core study team and DMID representative (see Section E for details). Presentation of data by swab or platform will be minimized as much as possible; any inclusion will use generic identifiers (i.e., platform or swab 1, 2, etc.) as appropriate. The report will include summaries of:

1. Study Population (Section G, Part 1, pg 10)
2. Study Status (Section G, Part 2, pg 11)
3. Swab Collection Completeness and Monitoring (Section G, Part 3, pg 11)
4. Swab Testing Completeness (Section G, Part 4, pg 12)

Closed Administrative Report: This report will be distributed only to the independent statistician (see Section E). The report will include analysis results from all assay platforms (swab and platform names will be presented).

1. Study Population (Section G, Part 1, pg 10)
2. Study Status (Section G, Part 2, pg 11)
3. Swab Collection Completeness and Monitoring (Section G, Part 3, pg 11)
4. Swab Testing Completeness (Section G, Part 4, pg 12)
5. Infection Status by Anatomic Site and Organism (Section G, Part 5, pg 12)

Of note, some analyses, tables or figures may be omitted at interim analyses if there are insufficient data to warrant analysis. Additional analyses may be provided if requested by the independent statistician.

Upon completing review of the interim analysis reports, it is anticipated that the independent statistician will provide recommendations to the protocol chair, protocol clinician and project lead. The choice of sharing these recommendations with members of protocol team will be left to the discretion of this latter group.

D. Final Analysis Considerations

The primary analysis will be conducted once data from the last participant enrolled has been received. The final analysis report will be distributed to the protocol team after Harvard receives the final, locked data transfer from DCRI and validation of primary analyses and internal review at Harvard are completed. The protocol team is defined in Section E. **The Final Analysis Report will include the following components.**

1. **Study Population (pg 10)**
2. **Study Status (pg 11)**
3. **Swab Collection Completeness and Monitoring (pg 11)**
4. **Swab Testing Completeness (pg 12)**
5. **Infection Status by Anatomic Site and Organism (pg 12, excludes Exploratory Analyses section)**

The final analysis population will include participants who meet all eligibility criteria **and are assigned a randomized swab order, and** provide four swabs from at least one anatomic site. **However, during the course of the study, the team discovered that swabs collected from 167 participants were not stored at the correct temperature per protocol before lab testing. As a result, the final analysis population will exclude all results from these 167 participants. The exclusion of these participants assumes data are missing at random (MAR) as the temperature deviation was random and not related to the lab test results. We do not expect the exclusion of these 167 participants to bias estimates of PPA and NPA, though some loss in precision is possible. Sensitivity analyses will not be conducted as part of final analysis unless requested otherwise by the team.**

For all final analyses, data collected from eCRFs or data entered by the laboratories in the laboratory information management system (LIMS) will be used. The primary analysis will evaluate the result for each diagnostic test (i.e., test under consideration) and compare with the ASIS for each anatomic site and organism combination. For each diagnostic test, positive percent agreement (PPA) and negative percent agreement (NPA) will be estimated using 95% confidence intervals. Confidence intervals will be estimated using the Score method. The primary analysis will also follow FDA guidance for incorporating indeterminate ASIS or test results for the test under consideration that are equivocal. As a result, all combinations of Infected/Indeterminate/Not Infected with all outcomes from the test under consideration (Positive/Equivocal/Negative/No result or Invalid) will be presented. If the test under consideration has “no result” because the test was not run and no attempt was made to test the sample, it will be excluded from the primary analysis.

Sensitivity analyses associated with the primary analyses will be conducted to examine the impact of different classifications of indeterminate results (i.e., all infected, all not infected, account for symptom status). Additional subgroup and secondary analyses are planned.

E. Report Distribution List

Unless otherwise mentioned, distribution lists for monitoring and analysis reports are comprised of:

Protocol Team: arlg.gc@mc.duke.edu members

Core Team: Protocol Chair, Protocol Clinician, Project Lead, Statisticians, Data Management, Clinical Trials Manager, Clinical Research Associate, Regulatory Associate.

DMID: DMID representative

DIR: Designated Independent Reviewer

Table 2: Report Distribution Summary

Report	Frequency	Distribution List
Harvard Monitoring Report	Bi-monthly (or as determined by core team)	Core Team
Open Administrative Interim Report	Every 3 months or every 500 participants (whichever comes first)	DIR, Core Team, DMID
Closed Administrative Interim Report	Every 3 months or every 500 participants (whichever comes first)	DIR, Statisticians
Final Analysis Report	End of study per study timeline	Protocol Team

F. Application Validation

All study-specific programs for creation of derived datasets for derivation of the primary outcomes defined in this document will require application validation per standing operating procedures (SOP) defined in CBAR PROG.10033 as appropriate; when applicable, requirements for independent results verifications of these datasets and application validation requirements for analysis programs are provided as annotations throughout the analysis plan. A copy of CBAR PROG.10033 is provided in Section H.

G. Analysis Plan

Throughout, annotations in square brackets ([xxx]) provide the data source.

1. Study Population

1.1. Accrual and Eligibility Violations

1. Table: Number (%) enrolled overall and by month and site.

   Note: Dates of first and last enrollments will be provided in a footnote to the table.

2. Table: Number (%) enrolled by month, site and reported sex at birth [DEMOG].
3. List: Description of violations of eligibility criteria if applicable. [INCEXC]

   Note: Participants enrolled and later found ineligible will be excluded from all analyses and be included in this listing. Information on whether swabs were collected for these participants will be noted.

4. Figure: Observed, cumulative and targeted accrual by month.

   Note: Targeted accrual is assumed to be approximately 209 participants per month for 12 months or 417 participants per month for 6 months to achieve full accrual of 2500 participants.

1.2. Study Population Characteristics

Table summaries will present the following study population characteristics overall and by sex. All variables, as noted below, will be analyzed on the continuous scale or as categories or both as appropriate.

For continuous variables, summary statistics will include # of participants, # of missing data points, mean and standard deviation, median (Q1-Q3), P10 and P90, and minimum and maximum.

For categorical variables, summary statistics will include number (%) for each category. In calculation of percentages, participants with missing data will not be included in the denominator.

1. Demographics
    a. Sex at birth: By category (male/female) [DEMOG]
    b. Gender: By category (man, woman, transman, transwoman, genderqueer, additional category, decline to answer) [DEMOG]
    c. Self-reported race, ethnicity and race/ethnicity as defined by NIH reporting standards: By category [DEMOG]
    d. Current age on day of study entry (years): Continuous and by age group (18-29, 30-39, 40-49, 50-59, 60+) [DEMOG]
2. Health Status [*For interim and final analyses only*]
    a. Abnormalities or symptoms in the pharynx in the past 7 days: By category (Yes/No) [SIGNS AND SYMPTOMS]
    Note: If yes, sub-categorization of the reported symptom will also be provided. This includes sore throat, painful swallowing, swollen/tender lymph nodes in the neck, and other symptom (with listed reasons).
    b. Abnormalities or symptoms in the rectum in the past 7 days: By category (Yes/No) [SIGNS AND SYMPTOMS]
    Note: If yes, sub-categorization of the reported symptom will also be provided. This includes rectal discharge, rectal bleeding, rectal itching, painful bowel movements, and other symptom (with listed reasons).

2. Study Status

As noted, the date of "study entry" is the same as the "off study" date. Study status will be determined by cross checking the study completion form with the sample collection form.

1. Table: Number (%) by category of study status.

    Categories: Completed study per sample collection form; incomplete- subject withdrew consent; incomplete – investigator decision; other reasons (with listing of reasons). [STUDY COMPLETION AND SAMPLE COLLECTION]

    Note: Study completion is defined as collection of at least four swabs from one anatomic site.

3. Swab Collection Completeness and Monitoring

All tables will be presented by anatomic site.

Of note, swab collection data will also be examined by study site and presented if low data completeness or high numbers of complications are observed.

3.1. Swab Collection Completeness and Complications

1. Table: Number (%) of participants reporting collected swabs.

    Categories: All 4 swabs, 3 swabs, 2 swabs, 1 swab, no swabs. [SAMPLE COLLECTION]

    Note: [CLOSED administrative report at interim analysis and final analyses only]: List swabs by name for each numeric category.

2. Listing/Table: Number (%) of participants with swabs not collected per assigned swab order. [SAMPLE COLLECTION]

    Note: If few deviations are reported, list the reported order the swabs were collected. ***Swabs not collected per assigned order will still contribute to interim and final analyses.***

3. Table: Number (%) of participants reporting sample collection complications.

Categories: Patient declined – due to excessive discomfort; patient declined – other reason; problem with testing materials, other (with listed reasons). [SAMPLE COMPLICATIONS]

Note: If few complications are reported, include the number of swabs collected (4, 3, 2, 1 or no swabs) for each complication. For CLOSED administrative report at interim analysis and final analyses only, list by swab name.

3.2. Laboratory Device Monitoring

The testing laboratories will maintain a log of all unanticipated device-related complications leading to no test, such as absence of transport media, quantity not sufficient, interference issues during testing, specimen transport collection system damage or incorrect transport system. No additional monitoring will be conducted.

4. Swab Testing Completeness and Monitoring

1. Table: Number (%) of participants with swabs tested (presented by anatomic site).

   Categories: All 4 swabs tested; 1-3 swabs tested; no swabs tested. [LAB DATA]

   For participants with fewer than 4 swabs tested, reasons why the laboratory did not test the swabs will be listed. [**Where available, summarize reasons from the lab issue tracking log;** LAB DATA]

2. **Table: Number (%) of participants with tiebreaker swabs tested (presented by anatomic site).**

   Categories: All 4 swabs tested; 1-3 swabs tested; no swabs tested. [LAB DATA]

5. Infection Status by Anatomic Site and Organism

[NOTE: Interim (Closed Report) and Final Analyses Only]

5.1. NAAT Test Results

Possible test results for each NAAT platform (as listed in **Table 3**) will be used for each anatomic site (pharynx or rectum) and organism (NG or CT) combination. These four combinations include NG of pharynx, CT of pharynx, NG of rectum, and CT of rectum.

Of note, if an expected repeat test result is missing, then the final test result will be derived as "NO RESULT" for the ASIS determination. This does not apply, however, to an initial test result of "EQUIVOCAL". In this case, if the expected repeat test result is missing, then the final result for ASIS determination will be "EQUIVOCAL".

Table 3: Summary of NAAT Test Results

NAAT	Possible Test Results	Notes
Company 1 Assay (Company 1)[2]	1. Not detected 2. Detected 3. Invalid (sample processing control or sample adequacy control failed) 4. Error (probe check control failed) 5. No result (insufficient data was collected, e.g. test aborted).	Initial invalid, error, or no result tests will be repeated. If the repeat test returns invalid, error, or no result, the final result will be considered an invalid and will be categorized as no result (NR) for the ASIS determination. If the repeat test returns not detected (negative) or detected (positive), this will be the result used for the ASIS determination.
Company 2 Assay (Company 2)[3]	1. Negative 2. Positive 3. Equivocal (result between positive and negative) **4. Invalid (run status is FAIL or other technical failure)** **5. Error (sample was not tested due to an**	Initial equivocal, invalid, **and error** test results will be repeated. **If the repeat test result returns equivocal, the final test result will be considered an <u>equivocal test result</u> for the ASIS determination. If the repeat test result returns invalid or error, the final test result will be categorized as <u>equivocal</u> if the initial test was equivocal and as <u>no result (NR)</u> if**

[2] Company 1 Assay package insert. Vol. XXX, Rev (Company 1). Company 2
[3] Assay package insert. Vol. XXX Rev .

	error detected by the instrument).	**the initial test was invalid or error for the ASIS determination.** If the repeat test returns negative or positive, this will be the result considered for the ASIS determination.
Company 3 assay (Company 3)[4]	For NG: 1. Positive (detected, with cycle number less than or equal to the assay cut-off) 2. Negative (no evidence of amplification or cycle number greater than the assay cut-off). **3. Error** ***Note: An equivocal interpretation does not apply.*** For CT: 1. Positive (detected, with cycle number less than or equal to the assay cut-off) 2. Negative (no evidence of amplification) 3. Equivocal (cycle number beyond the assay cut-off). **4. Error**	**A sample with initial interpretation of error (both CT and NG) or equivocal (CT only) will be retested. If the repeat test returns negative or positive, this will be the result considered for the ASIS determination and statistical analyses. If the repeat test result is equivocal (CT only), the final test result will be considered equivocal for the ASIS determination and statistical analyses below. If the repeat test result is error, the final test result will be categorized as no result (NR) for the ASIS determination and statistical analyses below if the initial test result was error and as equivocal if the initial test result was equivocal (CT only).**
Tiebreaker assay[5,]	1. Negative 2. Positive 3. Equivocal (result between negative and positive ranges) 4. **Invalid (run status is FAIL or other technical failure)** 5. **Error (sample was not tested due to an error detected by the instrument)**	Initial equivocal, invalid **and error test results** test results will be repeated. **If the repeated test result is equivocal, it will be considered an <u>equivocal test result</u> for the ASIS determination. If the repeat test result returns invalid or error, the final test result will be categorized as <u>equivocal</u> if the initial test was equivocal and as <u>no result (NR)</u> if the initial test was invalid or error for the ASIS determination.** If the repeat test returns negative or positive, this will be the result considered for the ASIS determination.

5.2. Anatomic Site Infection Status (ASIS) Determination

Per protocol, determination of the ASIS will be NAAT-specific and evaluated for each anatomic site and organism combination.

Possible ASIS outcomes include:

- Infected
- Not infected
- Indeterminate
- Invalid, exclude from analysis

The anatomic site is considered to be **infected** when both reference test results are positive/**detected**.

The anatomic site is considered to be **not infected** when both reference test results are negative/**not detected**.

If there is discordance between the reference tests, an additional NAAT test will be performed as a tiebreaker. In this case, agreement of 2/3 of the reference NAATs will determine the ASIS. If two tests are equivocal or one equivocal and one not run, the third test result will stand as the ASIS if positive or negative. If two tests are not run, the ASIS will be considered invalid and will be excluded from the analysis.

[4] Company 3 Package Insert. Vol. XXX.
[5] **Tiebreaker Assay Package Insert. Vol. XXX.**
[6]

All possible test result combinations are shown in **Table 4**. The tiebreaker test will be run by the lab if any NAAT is not concordant with the others and interpreted only in the case of discordant results between the two planned reference tests for each assay. [redacted].

To determine the ASIS, the test result for each respective site (pharynx or rectum) and each organism (NG or CT) for each NAAT platform will be used.

Table 4: Determination of the Anatomic Site Infection Status (ASIS)
Note:
*E **=** equivocal result;
****NR = no result. This can occur either because the test result was invalid or because the test could not be run (e.g. too little sample, improperly shipped, no sample received).

Comparator NAAT 1 Result	Comparator NAAT 2 Result	Tiebreaker NAAT Result	Anatomic Site Infection Status (ASIS)
+	+	Not indicated	Infected
+	-	+	Infected
+	E*	+	Infected
+	NR**	+	Infected
+	-	-	Not infected
+	-	E	Indeterminate
+	-	NR	Indeterminate
+	E	-	Indeterminate
+	E	E	Infected
+	E	NR	Infected
+	NR	-	Indeterminate
+	NR	E	Infected
+	NR	NR	Invalid, remove from analysis
-	-	Not indicated	Not infected
-	+	-	Not infected
-	E	-	Not infected
-	NR	-	Not infected
-	+	+	Infected
-	+	E	Indeterminate
-	+	NR	Indeterminate
-	E	+	Indeterminate
-	E	E	Not infected
-	E	NR	Not infected
-	NR	+	Indeterminate
-	NR	E	Not infected
-	NR	NR	Invalid, remove from analysis
E	+	+	Infected
E	-	-	Not infected
E	+	-	Indeterminate
E	+	E	Infected
E	+	NR	Infected
E	-	+	Indeterminate
E	-	E	Not infected
E	-	NR	Not infected
E	NR	+	Infected
E	NR	-	Not infected
E	NR	E	Indeterminate
E	NR	NR	Invalid, remove from analysis
NR	+	+	Infected
NR	-	-	Not infected
NR	NR	Not indicated	Invalid, remove from analysis

NR	+	-	Indeterminate
NR	+	E	Infected
NR	+	NR	**Invalid, remove from analysis**
NR	-	+	Indeterminate
NR	-	E	Not infected
NR	-	NR	Invalid, remove from analysis
NR	E	+	Infected
NR	E	-	Not infected
NR	E	E	Indeterminate
NR	E	NR	Invalid, remove from analysis

5.3. Interim Analysis

5.3.1. Interim Analysis Considerations

The study cannot be stopped at the interim for reaching regulatory goals in order to preserve error rates / coverage probability and to ensure enough data for subgroup analyses. Since the trial cannot be stopped for attainment of the regulatory goal, no adjustment to confidence levels are necessary.

Although it is anticipated that swab collection and data completeness will be high for both anatomical sites, both the Intention-to-Diagnose (ITD) and modified Intention-to-Diagnose (mITD) infection rates[7], as described in **Table 5**, will be estimated. It is expected that disease prevalence (infection rate) of NG in the rectum, NG in the pharynx, and CT in the rectum will each be greater than 7.5% in the population under evaluation. Disease prevalence of CT in the pharynx is expected to be rare.

Table 5: Infection rate calculation for each test under consideration

Intent-to-Diagnose (ITD) Infection Rate	Number of infected ASIS results for test under consideration / Total number of ASIS results[1]
Modified Intent-to-Diagnose (mITD) Infection Rate	Number of infected ASIS results for test under consideration / Total number of ASIS results with exclusion of invalid results[2]
[1] *The denominator will include the sum of all infected, not infected, indeterminate and invalid ASIS results.* [2] *The denominator will include the sum of all infected, not infected, and indeterminate ASIS results.*	

Operationally, a range of scenarios for the unobserved data will be generated at each interim review to assist with the decision making regarding sample size adjustments. To ensure enough infected participants to estimate PPAs with desired precision, the lowest ITD/mITD infection rate across the three assays and the three organism/site combinations (except CT in the pharynx) will be used as the observed rate for interim analysis. Of note, it is assumed that there will be greater precision to evaluate NPA as it is expected that there will be more not-infected results than infected results for each anatomic site.

Table 6 and **Table 7** illustrate hypothetical scenarios when the observed infection rate after the first 1000 participants is rarer than anticipated (equal to 5%, **Table 6)** or more prevalent than anticipated (equal to 10%, **Table 7**), respectively. For both tables, a range of infection rates for the unobserved data (scenarios A-C) are also presented to demonstrate the probability and corresponding total sample sizes to obtain 150, 175 or 200 disease positive participants, respectively, at the end of study.

If infection is rarer than anticipated, then increases to sample size may be considered as demonstrated in scenario A from Table 6. In this case, should the observed prevalence of 5% remain unchanged for the unobserved data, then enrollment of approximately 3200 total participants would be needed to ensure at least 80% probability of obtaining 150 disease positive participants. However, should the infection rate increase to 9% (Scenario C, **Table** 6), then sample size adjustments may not be warranted.

Alternatively, if infection is more prevalent than expected, a smaller sample size may be considered as demonstrated in scenarios A and B from **Table 7**.

[7] Fundamental Concepts for New Clinical Trialists. A Evans, S. and A Ting, N. 9781420090871. https://books.google.com/books?id=G1IUPQAACAAJ. 2015. Taylor & Francis.

Sample size adjustments may also be considered if there is evidence that the number of ASIS results categorized as invalid will impact study accrual time and/or data analyses.

Table 6: Hypothetical illustration of simulated infection rates after first 1000 participants and observed prevalence of 5%

Observed Response Rate and Count			
ITD infection rate at interim	**Prevalence of Disease Positive (%)**	**Disease Positive (N)**	
	5%	**50 participants**	

Example Scenarios for Unobserved Data								
Scenario	**Assumed prevalence rate[1]**	**Average prevalence rate[2]**	**Estimated probability (P) and sample size (N) to obtain:**					
			150 Disease+ total participants		**175 Disease+ total participants**		**200 Disease+ total participants**	
			P	**N**	**P**	**N**	**P**	**N**
A	**5%**	**5%**	<50%	2500	<50%	2500	<50%	2500
			80%	3220	80%	3740	80%	4260
			90%	3340	90%	3860	90%	4380
B	**7%**	**6.2%-6.4%**	64%	2500	<50%	2500	<50%	2500
			80%	2589	80%	2956	80%	3323
			90%	2669	90%	3046	90%	3413
C	**9%**	**7.2%-7.6%**	80%	2232	76%	2500	<50%	2500
			90%	2302	80%	2519	80%	2807
			99%	2500	90%	2589	90%	2877

[1] The assumed prevalence rate for the remaining unobserved data. [2] Average prevalence rate is the average of the observed and assumed prevalence rates for the entire duration of study for a given sample size.

Table 7: Hypothetical illustration of simulated infection rates after first 1000 participants and observed prevalence of 10%

Observed Response Rate and Count			
ITD infection rate at interim	**Prevalence of Disease Positive (%)**	**Disease Positive (N)**	
	10%	**100 participants**	

Example Scenarios for Unobserved Data								
Scenario	**Assumed prevalence rate[1]**	**Average prevalence rate[2]**	**Estimated probability (P) and sample size (N) to obtain:**					
			150 Disease+ total participants		**175 Disease+ total participants**		**200 Disease+ total participants**	
			P	**N**	**P**	**N**	**P**	**N**
A	**10%**	**10%**	80%	1610	80%	1870	80%	2130
			90%	1670	90%	1930	90%	2190
			>95%	2500	>95%	2500	>95%	2500
B	**7%**	**8.1%-8.6%**	80%	1875	80%	2242	63%	2500
			90%	1955	90%	2322	80%	2609
			>95%	2500	>95%	2500	90%	2699
C	**5%**	**6.5%-7.3%**	80%	2220	<50%	2500	<50%	2500
			90%	2330	80%	2740	80%	3250
			>95%	2500	90%	2860	90%	3380

[1] The assumed prevalence rate for the remaining unobserved data. [2] Average prevalence rate is the average of the observed and assumed prevalence rates for the entire duration of study for a given sample size.

5.3.2. Interim Analyses

Note: Analysis programs will require independent results verification per CBAR PROG.10033.

a. Table: For each diagnostic assay, frequency of observed test outcomes by anatomical site and organism combination as shown in **Table 8**. [Data source: LAB DATA]

Note: Test results for each diagnostic assay were defined previously in Section 5.1.

Table 8: Frequency of observed test outcomes by anatomical site and organism combination

NAAT: Company 1 system Test Results					
	Not detected	Detected	Invalid	Error	No result
***NG*, rectum**					
***NG*, throat**					
***CT*, rectum**					
***CT*, throat**					
NAAT: Company 2 system Test Results					
	Negative	Positive	Equivocal	Invalid	
***NG*, rectum**					
***NG*, throat**					
***CT*, rectum**					
***CT*, throat**					
NAAT: Company 3 system Test Results					
	Negative	Positive			
***NG*, rectum**					
***NG*, throat**					
	Negative	Positive	Equivocal		
***CT*, rectum**					
***CT*, throat**					
NAAT: Tiebreaker system Test Results					
	Negative	Positive	Equivocal	Invalid	
***NG*, rectum**					
***NG*, throat**					
***CT*, rectum**					
***CT*, throat**					

b. Table/Figure: Frequency of comparator NAAT result and tiebreaker result combinations to define ASIS as shown in **Table 9**. This will be conducted for each test under consideration and be presented by anatomical site and organism combination. There will be a total of three tables to reflect the three tests under consideration. [Data source: LAB DATA]

Note: Only combinations with a frequency of one or greater will be shown. Any combination not shown will indicate that this combination was not observed (i.e., frequency equal to 0).

Table 9: Frequency of comparator NAAT result and tiebreaker result combinations by anatomical site and organism for test under consideration XX

Test Under Consideration: XXX				
Anatomical Site and Organism: *NG*, rectum				
Comparator NAAT Result 1	Comparator NAAT Result **2**	Tiebreaker NAAT result	ASIS	Frequency (n)
+	+	Not indicated	Infected	XX
+	-	+	Infected	XX
(Remaining observed combinations as described in **Table 4**.)				
Anatomical Site and Organism: *NG*, throat				
Comparator NAAT Result 1	Comparator NAAT Result **2**	Tiebreaker NAAT result	ASIS	Frequency (n)
+	+	Not indicated	Infected	XX
+	-	+	Infected	XX
(Remaining observed combinations as described in **Table 4**.)				
Anatomical Site and Organism: *CT*, rectum				

Comparator NAAT Result 1	Comparator NAAT Result **2**	Tiebreaker NAAT result	ASIS	Frequency (n)
+	+	Not indicated	Infected	XX
+	-	+	Infected	XX
(Remaining observed combinations as described in **Table 4**.)				
Anatomical Site and Organism: *CT*, throat				
Comparator NAAT Result 1	Comparator NAAT Result **2**	Tiebreaker NAAT result	ASIS	Frequency (n)
+	+	Not indicated	Infected	XX
+	-	+	Infected	XX
(Remaining observed combinations as described in **Table 4**.)				

c. Table/Figure: Number (%) of observed total test results, ASIS results, and prevalence (i.e., infection) rate. This will be conducted for each test under consideration and be presented by anatomical site and organism combination as shown in **Table 10** below. [Data source: LAB DATA]

Table 10: Number (%) of observed total test results, ASIS results, and prevalence (i.e., infection) rate for each test under consideration by anatomical site and organism

Test under consideration: Company 1 System **Comparator NAATs:** Company 2 and Company 3 Systems **Tiebreaker NAAT:** Tiebreaker System								
	Sample Size (N)		ASIS				Prevalence (P, %)	
	N (ITD)	N (mITD)	Infected	Not infected	Indeterminate	Invalid	P (ITD)	P (mITD)
***NG*, rectum**								
***NG*, throat**								
***CT*, rectum**								
***CT*, throat**								
Test under consideration: Company 2 System **Comparator NAATs:** Company 1 and Company 3 Systems **Tiebreaker NAAT:** Tiebreaker System								
	Sample Size (N)		ASIS				Prevalence (P, %)	
	N (ITD)	N (mITD)	Infected	Not infected	Indeterminate	Invalid	P (ITD)	P (mITD)
***NG*, rectum**								
***NG*, throat**								
***CT*, rectum**								
***CT*, throat**								
Test under consideration: Company 3 **Comparator NAATs:** Company 1 and Company 2 Systems **Tiebreaker NAAT:** Tiebreaker System								
	Sample Size (N)		ASIS				Prevalence (P, %)	
	N (ITD)	N (mITD)	Infected	Not infected	Indeterminate	Invalid	P (ITD)	P (mITD)
***NG*, rectum**								
***NG*, throat**								
***CT*, rectum**								
***CT*, throat**								

a. Table/Figure: Predicted infection rate summary for range of scenarios (see **Table 6** and **Table 7** described in Section 5.3.1).
b. Table: Observed frequency of tiebreaker run versus expected number of runs (number, %).

c. Table: Using Fischer's exact test, the association between of test results and randomized swab order will be examined. Associations will be conducted by anatomical site and organism combination.

5.4. Final Analysis

5.4.1. Outcome Measures

For each participant, primary and secondary endpoints will be defined for each anatomical site and organism. [Primary source data: SAMPLE COLLECTION, STUDY COMPLETION and LAB DATA]

a. Primary Endpoints

Note: Derived datasets relating to the derivation of these endpoints will undergo independent results verification in accordance with CBAR PROG.10033.

Anatomic site infection status is determined by the reference standard (described in Section 4.2).

- Infection status for Neisseria gonorrhoeae in the rectum as determined by each NAAT
- Infection status for Neisseria gonorrhoeae in the pharynx as determined by each NAAT
- Infection status for Chlamydia trachomatis in the rectum as determined by each NAAT
- Infection status for Chlamydia trachomatis in the pharynx as determined by each NAAT

5.4.2. Final Analysis Considerations

Note: All dataset derivation programs for this endpoint will undergo independent results verification in accordance with CBAR PROG.10033; validation requirements for analysis programs are stated below.

As noted previously, for all final analyses, data collected from eCRFs or data entered by the laboratories in the laboratory information management system (LIMS) will be used. Data from additional sources such as machine data or issue tracking logs may be reviewed for clarification, but will not be used for analysis.

The result for each diagnostic test will be compared with the ASIS for that anatomic site and organism. PPA and NPA will also be estimated for each diagnostic test with 95% confidence intervals. Confidence intervals will be estimated using the Score method.[8] If the test under consideration has "no result" because the test was not run and no attempt was made to test the sample, the test result will be excluded from primary analysis. If the ASIS result is invalid, this result will be excluded from the primary analysis.

FDA guidance documents will be followed as part of the primary analysis to incorporate indeterminate ASIS or test results for the test under consideration that are equivocal.[9] It is recognized that there are pros and cons to the manner in which indeterminates are handled and how these impact the resulting estimates of PPA and NPA. The primary analysis approach is the most conservative and is biased downwards. If PPA is >90% under this scenario, then the conclusion of PPA >90% is clear.

Sensitivity analyses will be conducted as appropriate to evaluate the impact of diagnostic accuracy for a range of scenarios addressing indeterminate ASIS results. This will include counting indeterminate ASIS results against the results for the test under consideration with all combinations of Infected/Indeterminate/Not Infected with all outcomes from the test under consideration (Positive/Equivocal/Negative/No result or Invalid) evaluated. These analyses may not be conservative for calculations of PPA and NPA.

5.4.3. Primary Analyses

The following will be conducted for the three diagnostic tests.

Note: Analysis programs will require independent results verification per CBAR PROG.10033.

a. Table: For each diagnostic assay **(including the tiebreaker assay),** frequency of observed test outcomes by anatomical site and organism combination. [See **Table 8** described in Section 5.3.2; data source: LAB DATA]

[8] FDA. Establishing the performance characteristics of in vivo diagnostics devices for Chlamydia trachomatis and/or Neisseria gonorrhoeae: screening and diagnostic testing. (2011).
[9] FDA. Establishing the performance characteristics of in vivo diagnostics devices for Chlamydia trachomatis and/or Neisseria gonorrhoeae: screening and diagnostic testing. (2011).

b. **Table: Frequency (%) of re-tests by clinic and platform. This will be conducted for each test under consideration and be presented by anatomical site and organism combination. [Data source: LAB DATA; DEMOG]**
c. **Table: Listing and frequency (%) of re-tests results by platform. This will be conducted for each test under consideration and be presented by anatomical site and organism combination. [Data source: LAB DATA]**
d. Table/Figure: Frequency of comparator NAAT result and tiebreaker result combinations to define ASIS. This will be conducted for each test under consideration and be presented by anatomical site and organism combination. There will be a total of three tables to reflect the three tests under consideration. [See **Table 9** previously described in in Section 5.3.2; data source: LAB DATA]
e. Table/Figure: Number (%) of observed total test results, ASIS results, and prevalence (i.e., infection) rate. This will be conducted for each test under consideration and be presented by anatomical site and organism combination. [See **Table 10** previously described in Section 5.3.2; data source: LAB DATA]
f. **Table: Observed infection rate by clinic. This will be conducted for each test under consideration and be presented by anatomical site and organism combination**. **[Data source: LAB DATA; DEMOG]**
g. Tables:
    - Cross comparison of number (%) of results of test under consideration versus ASIS;
    - Estimates of PPA and NPA with 95% Score confidence intervals;
    - Due to differences in PPA and NPA calculations for NAAT platforms that do not have an [redacted] result when it is the test under consideration, analysis results will be estimated and presented separately based on this distinction as follows:
        - **Table 11** and **Table 12**: Company 2 Assay (Company 2) and Company 3 Assay (Company 3).
        - **Table 13** and **Table 14**: Company 1 Assay (Company 1) and Company 3 Assay (Company 3).
h. Figure: Plot of PPA and 95% confidence interval band versus proportion of indeterminate ASIS results assumed to be positive (range of 0 to 1); vice-versa for NPA plot.
i. Table: Using Fischer's exact test, the association between test results and randomized swab order will be examined. Associations will be conducted by anatomical site and organism combination.

Table 11: Result of Test under Consideration versus ASIS: Company 2 Assay (Company 2) and Company 3 Assay (Company 3)

		ASIS		
		Infected	**Indeterminate**	**Not infected**
Result of Test under Consideration	**Positive**	A	D	G
	Equivocal	B	E	H
	Negative	C	F	I
	No result	Exclude from analysis[1]		

[1] *Note: If the test under consideration has "no result" because the test was not run and no attempt was made to test the sample, the test result will be excluded from primary analysis. If the ASIS result is invalid, this result will be excluded from the primary analysis.*

Table 12: Calculation of the positive and negative percent agreement: Company 2 Assay (Company 2) and Company 3 Assay (Company 3)

Analysis Type	**PPA**	**NPA**
Primary Analysis	A / (A+B+C**+F**)	I / (G+H+I **+ D**)
Sensitivity Analysis Scenarios		
Classify indeterminates[1] using symptom status reported from [SIGNS AND SYMPTOMS]	A / (A+B+C+F)	I / (G+H+I+D)
Include all indeterminate tests as infected.	(A+D) / (A+B+C+D+E+F)	I / (G+H+I)
Include all indeterminate tests as not infected.	A / (A+B+C)	(I+F) / (D+E+F+G+H+I)
Consider indeterminate and equivocal test results as "missing", with the assumption of missing at random, and model the missing results. [2]	A / (A+C)	I / (G+I)

[1] *Classify indeterminate tests on the basis of symptom status. Include indeterminate tests as Infected if the participant is symptomatic in that compartment; include indeterminate tests as not infected in the participant is asymptomatic in that compartment.*
[2] *Cells B, D, E, F and H will be assigned or weighted to cells A, C, G and I based on modeling of the missing results.*

Table 13: Result of Test under Consideration versus ASIS: Company 1 Assay (Company 1) and Company 3 Assay (Company 3)

		ASIS		
		Infected	Indeterminate	Not infected
Result of Test under Consideration	**Positive/Detected**	A	C	E
	Negative/Not detected	B	D	F
	No result	Exclude from analysis[1]		

[1] *Note: If the test under consideration has "no result" because the test was not run and no attempt was made to test the sample, the test result will be excluded from primary analysis. If the ASIS result is invalid, this result will be excluded from the primary analysis.*

Table 14: Calculation of the positive and negative percent agreement: Company 1 Assay (Company 1) and for the Company 3 Assay (Company 3)

Analysis Type	**PPA**	**NPA**
Primary Analysis	A / (A+B**+D**)	F / (C+E**+F**)
Sensitivity Analysis Scenarios		
Classify indeterminates[1] using symptom status reported from [SIGNS AND SYMPTOMS]	A / (A+B+D)	F / (C+E+F)
Include all indeterminate tests as infected.	(A+C) / (A+B+C+D)	F / (E+F)
Include all indeterminate tests as not infected.	A / (A+B)	(D+F) / (C+D+E+F)
Consider indeterminate and equivocal test results as "missing", with the assumption of missing at random, and model the missing results. [2]	A / (A+B)	F / (E+F)

[1] *Classify indeterminate tests on the basis of symptom status. Include indeterminate tests as Infected if the participant is symptomatic in that compartment; include indeterminate tests as not infected in the participant is asymptomatic in that compartment.*
[2] *Cells C and D will be assigned or weighted to cells A, B, E and F based on modeling of the missing results.*

5.4.4. Secondary (Global) Analyses

Note: Analysis programs may require independent results verification per CBAR PROG.10033.

The following will be conducted for the three diagnostic tests.

a. Table: Summary of estimated positive and negative predictive values (PPV and NPV).
b. Figure: Plot of predictive PPV and NPV estimates as a function of prevalence for each test (point estimates and 95% pointwise confidence bands)
c. Table/Figure: The 95% Score confidence interval estimates of positive and negative likelihood ratios with forest plot display (one plot per test).
d. Table/Figure: Sensitivity analyses estimating PPV, NPV, and positive and negative likelihood ratios where indeterminates are: a) counted as all infected, b) all not infected, and c) based on symptom status reported in [SIGNS and SYMPTOMS].

5.4.5. Subgroup Analyses

Note: Analysis programs may require independent results verification per CBAR PROG.10033.

For each of the three diagnostic tests, subgroup analyses will be conducted for males, females, symptomatic participants, and asymptomatic participants by pathogen and anatomic site. The analyses described in Sections 5.4.3 and 5.4.4 will be conducted for each group.

5.5. Exploratory Analyses

Application of methods for diagnostic benefit:risk analyses will be conducted, including BED-FRAME methodologies[10]. Analyses related to these objectives will be initiated upon completion of primary analyses.

[10] Evans, S.R., Pennello, G., Pantoja-Galicia, N., Jiang, H., Hujer, A.M., Hujer, K.M., Manca, C., Hill, C., Jacobs, M.R., Chen, L. and Patel, R., 2016. Benefit-risk evaluation for diagnostics: a framework (BED-FRAME). Clinical Infectious Diseases, p.ciw329.

H. APPENDIX 1: CBAR SOP PROG.10033

Center for Biostatistics in AIDS Research (CBAR)	
STANDARD OPERATING PROCEDURE	
Title: Application Validation	
Document ID: PROG.10033	Document Version: 4
	Page 23 of 27

1. Purpose

This document provides procedures for validation of programs and applications developed at CBAR as part of work on clinical studies.

2. Scope and Applicability

This document is applicable to programs and applications developed and used by CBAR workforce members in conjunction with work with data from clinical studies.

3. Introduction

Validation of programs and applications used at CBAR to create or analyze datasets in conjunction with clinical studies ensures that analysis results accurately reflect the original source data and conform to analysis specifications. Using a risk-based approach to validation, CBAR has developed standard operating procedures (SOPs) that define minimum programming standards for coding, testing and validation for the range of programming applications used at CBAR. This document provides a reference to the specific programming standards that apply to specific types of programs or applications and outlines procedures for documenting their validation.

4. Definitions

- **CBAR program:** Any of the various types of programs or applications that are created and maintained by the CBAR Programming Core for use across CBAR for the purpose of creating reports, files, or SAS datasets with minimal input from users. This includes SAS table and format programs and reporting macros as well as UNIX and R packages (e.g., MAKETOX2 and the PIPS library of functions).
- **CBAR derived dataset program:** A program, created and maintained by the CBAR Programming Core, that creates one or more standardized derived datasets for use across CBAR, which does not represent a single CRF or codebook at the DMC.
- **CBAR macro:** A generic SAS macro, created and maintained by the CBAR Programming Core, that consists of flexible SAS code that can be easily tailored to individual studies for reporting purposes (i.e., CBAR reporting macro) or perform simple operations similar to SAS functions (i.e., CBAR autocall macro).
- **CBAR R program:** An R program used by a CBAR R package created or maintained by members of the CBAR Programming Core.
- **CBAR SAS format program:** A program, created and maintained by the CBAR Programming Core, that creates permanent SAS (in)format(s) for the CBAR format catalog.
- **CBAR SAS template program:** A program, created and maintained by the CBAR Programming Core, that defines one or more SAS ODS templates (e.g., a style template, ExcelXP tagset, or graphic template).
- **CBAR UNIX package:** A collection of UNIX programs and other files, created and maintained by members of the CBAR Programming Core, for use across CBAR for the purpose of creating reports, files, or SAS datasets with minimal input from users. This includes all modules included by the main executable program file. For example, the MAKE_SMR UNIX package contains the MAKE_SMR file, smr.sas, and all of the macro program files included by smr.sas, and the MAKEDATA downloading scripts are part of a CBAR UNIX Package.
- **CBAR UNIX program:** A UNIX shell script, created and maintained by members of the CBAR Programming Core, comprised of all modules utilized by the executable program file for use across CBAR for the purpose of creating reports, files, or SAS datasets with minimal input from users. For example: Both MAKEDATA and MAKEDATA_NOSTUDY are both CBAR UNIX programs within the CBAR UNIX package MAKEDATA.

- **Clinical Study:** A clinical trial or observational study involving human subjects.
- **Miscellaneous CBAR Program:** A program or application, created and maintained by the CBAR Programming Core, that does not fall under the definition of any of the CBAR programs in Section 5.0 of PROG.10033 Application Validation.
- **SAS table program:** A SAS program, created and maintained by the CBAR Programming Core, which creates one or more SAS datasets from ASCII data to represent a single CRF or codebook from the DMC.
- **Workforce members:** Employees (both academic appointees and staff), and other persons whose conduct, in the performance of CBAR work is under the direct control of CBAR whether or not they are paid by Harvard University.

5. Procedure

The following table outlines the different types of programs or applications based on their intended purposes and programming language. The table provides the CBAR SOP that defines the relevant programming standards for the development, testing, and validation of the program or application as well as the appropriate validation form that documents the validation process.

Formal validation of programs and applications not covered by the scope of relevant programming standards SOP is not required, but adherence to the practices described therein (as appropriate) is recommended.

Development, testing, and validation standards of CBAR programs and applications of a type not covered in the table below (miscellaneous CBAR programs) are at the discretion of the Head of the CBAR Programming Core. Validation review of these programs consists of 5 areas of focus: Source Code Control, Supporting Documentation, Program Code Review, Program Logic Review, and Testing Program Review, including Input and Output. Specific criteria within each of the 5 areas are at the discretion of the validation reviewer and are approved by the Head of the CBAR Programming Core.

Type of Program/ Application	**Programming Standards SOP**	**Validation Form**	**Personnel to be Notified upon Completion of Validation**
Study-Specific Derived Dataset Creation Programs (including format and macro programs)	PROG.10066 Study-Specific Derived Dataset Creation Programming Standards	PROG.10066.f1 Study-Specific Derived Dataset Creation Programming Review	Programmer
Study-Specific Analysis programs (including macro, template, and R programs)	PROG.10067 Study-Specific Analysis Programming Standards	PROG.10067.f1 Study-Specific Analysis Programming Review	Programmer
Study-Specific Fix Files	PROG.10030 Study-Specific Fix File Programming Standards	PROG.10030.f1 Study-Specific Fix File Programming Standards	Programmer
Study-Specific User Options Files	PROG.10071 Study-Specific User Options File Standards	PROG.10071.f1 Study-Specific User Options File Review	Programmer
SAS Table Programs	PROG.10029 SAS Table Programming Standards	PROG.10029.f1 SAS Table Programming Review	None
CBAR Derived Dataset Programs	PROG.10035 SAS Derived Dataset Programming Standards	PROG.10035.f1 SAS Derived Dataset Programming Review	Programmer, cbar.cda
CBAR SAS Format Programs	PROG.10037 SAS Format Programming Standards	PROG.10037.f1 SAS Format Programming Review	None
CBAR Macros	PROG.10034 SAS Macro Programming Standards	PROG.10034.f1 SAS Macro Programming Review	Programmer, head of CBAR programming core

UNIX Programs	PROG.10031 UNIX Package Programming Standards	PROG.10031.f1 UNIX Package Programming Review	Programmer, head of CBAR programming core
CBAR R Programs	PROG.10036 R Programming Standards	PROG.10036.f1 R Programming Review	Programmer, head of CBAR programming core
CBAR SAS Template Programs	PROG.10038 SAS Template Programming Standards	PROG.10038.f1 SAS Template Programming Review	Programmer, head of CBAR programming core

5.1 Validation Form Submission

The completed and signed validation form is submitted to Document Management within 5 days of validation completion in one of the following ways:

- Place the form in the Document Management mailbox; or
- Mail the form to Document Management at CBAR; or
- Contact Document Management to arrange another method of submitting the signed form.

In the event that completion and signing of the validation form conflicts with the analysis reporting timeline, notification of validation completion may be documented via email to the required personnel. In this case, submission of the completed validation form to Document Management at CBAR occurs within 5 business days of distribution of the analysis report.

For the validation of miscellaneous CBAR programs, the validation criteria are submitted with the validation form.

5.2 Content of PROG.10033.f1 Miscellaneous CBAR Programming Review

- Program name and location
- Last date in change history
- Coding programmer
- Validation reviewer
- Head of the CBAR Programming Core
- Date of validation
- Version of PROG.10033 Application Validation used
- Miscellaneous CBAR Program Review
  - Review Criteria
  - Finding(s) during review
  - Resolution(s) prior to validation
- Signature of validation reviewer
- Date of signature
- Signature of Head of the CBAR Programming Core
- Date of signature

Referenced Documents

Document Title	Location
PROG.10029 SAS Table Programming Standards	Secure location on the CBAR network
PROG.10029.f1 SAS Table Programming Review	Secure location on the CBAR network
PROG.10030 Fix File Programming Standards	Secure location on the CBAR network
PROG.10030.f1 Fix File Programming	Secure location on the CBAR network

PROG.10031 UNIX Programming Standards	Secure location on the CBAR network
PROG.10031.f1 UNIX Programming Review	Secure location on the CBAR network
PROG.10033.f1 Miscellaneous CBAR Programming Review	Secure location on the CBAR network
PROG.10034 SAS Macro Programming Standards	Secure location on the CBAR network
PROG.10034.f1 SAS Macro Programming Review	Secure location on the CBAR network
PROG.10035 SAS Derived Dataset Programming Standards	Secure location on the CBAR network
PROG.10035.f1 SAS Derived Dataset Programming Review	Secure location on the CBAR network
PROG.10036 SAS Template Programming Standards	Secure location on the CBAR network
PROG.10036.f1 SAS Template Programming Review	Secure location on the CBAR network
PROG.10037 SAS Format Programming Standards	Secure location on the CBAR network
PROG.10037.f1 SAS Format Programming Review	Secure location on the CBAR network
PROG.10038 R Programming Standards	Secure location on the CBAR network
PROG.10038.f1 R Programming Review	Secure location on the CBAR network
PROG.10066 Study-Specific Derived Dataset Creation Programming Standards	Secure location on the CBAR network
PROG.10066.f1 Study-Specific Derived Dataset Creation Programming Review	Secure location on the CBAR network
PROG.10067 Study-Specific Analysis Programming Standards	Secure location on the CBAR network
PROG.10067.f1 Study-Specific Analysis Programming Review	Secure location on the CBAR network
PROG.10071 Study-Specific User Options File Standards	Secure location on the CBAR network
PROG.10071.f1 Study-Specific User Options File Review	Secure location on the CBAR network

Version History

Version	Changes Made	Effective Date
1	Original Version	12/1/2013
2	Rationale: This version redefines Section 7. Version History to provide more information about the reason for the new version and the major changes included. This also more clearly describes the scope, procedure, and other parts of the SOP. Purpose, Scope, Introduction, Definitions, Section 5: Clarification of language to better describe the scope and procedure; updating of outdated information and removal of unused definitions Section 7: Format of section was changed	12/1/2013

3	Purpose, Scope, and Introduction: Updated to clarify the rationale for validation and better define how the programming standards and validation procedure work together. Definitions: - The definition of CBAR program has been updated - A new definition of a clinical study has been added. Table: Minor changes to table headings 5.0 Procedure: - Minor modifications to improve clarity - Addition of standards to be defined by Head of Programming Core for CBAR programs not otherwise covered by current standards - Notification of programmer and cbar.cda for SAS Table programs and SAS format programs have been removed 5.1 Validation Form Submission - Order of the sequence of events changed to clarify that submission of the validation form occurs for the validation to be considered complete - Time-frame for validation document submission removed since, per SOP, it is required for validation completion - Flexibility is provided for validation form completion in the case of time-constraints	7/1/2014
4	Rationale: The updates made to this SOP for this version detail validation procedures for miscellaneous CBAR programs. Definitions: - Added definition of Miscellaneous CBAR program 5.0 Procedure: - Added description of the validation review of miscellaneous CBAR programs 5.1 Validation Form Submission - Described validation form submission for miscellaneous CBAR programs - Further updates provided for validation form submission 5.2 Content of PROG.10033.f1 Miscellaneous CBAR Programming Review - Added this section to describe the necessary fields for the creation of PROG.10033.f1.	12/1/2015