UNIVERSITÄT BERN
Faculty of Medicine
CTU Bern

# Statistical Analysis Plan (SAP)

## Title of Study: Diagnostic test accuracy of a mobile colposcope (Gynocular™), HR-HPV testing, and VIA for detection of high-grade squamous intraepithelial lesions of the cervix in women living with HIV

## Acronym of Study: Gynocular Studies

# Administrative Information

| | |
|---|---|
| Project number: | 851 |
| Trial registration number: | Trial registration number |
| SAP version: | SAP 1.0 (23.04.2020) |
| Protocol version: | 1.1 dated 10 January 2019 |

## Contributors

| Name | Affiliation | Role in SAP writing |
|---|---|---|
| Marie Roumet | CTU Bern | Author |
| Limacher Andreas | CTU Bern | Reviewer |
| | | |

## Revision history

| Revision | Justification | Timing |
|---|---|---|
| n.a. | First version | 23.04.2020 |
| | | |
| | | |

## Approved by

| Name | Affiliation | Study Role | Date and Signature (wet ink) |
|---|---|---|---|
| Marie Roumet, PhD | CTU Bern | Trial Statistician | 13.5.2020 |
| Andreas Limacher, PhD | CTU Bern | Senior Statistician | May 11th, 2020 |
| Albert Manasyan, MD, MPH | Centre for Infectious Disease Research in Zambia | Local-Investigator | May 8, 2020 |
| Julia Bohlius, PD MD MScPH | Institute of Social and Preventive Medicine, University of Bern | Sponsor-Investigator | May 11th, 2020 |

## Contents

# 1. Introduction

## 1.1 Background and rationale

Women living with HIV (WLHIV) in Southern Africa have a far higher risk of developing cervical cancer (CC) than HIV-uninfected women. The most feasible method of screening, visual inspection with acetic acid (VIA), is recommended by the WHO when no other options exist. However, this has inferior test accuracy, especially among HIV-infected women, when compared to other methods. The best strategy or combination of screening strategies in low- and middle-income countries (LMIC) remains unclear. The WHO has acknowledged the urgency of improved screening strategies for LMIC, especially among WLHIV. Evaluation of alternative screening strategies using standalone tests or combinations of tests may identify feasible screening strategies with higher test accuracy in WLHIV in LMIC.

## 1.2 Objectives

The aim of the study is to provide the first-ever assessment of the Gynocular™, a mobile, handheld colposcope, to identify women with either high-grade squamous intraepithelial lesions (HSIL) of the cervix or cervical cancer among women infected with HIV in Zambia. In this document we use the term (pre)cancer to refer to this condition.

The main objective is to evaluate and compare the ability of Gynocular™, and two already existing tests: VIA and Hr-HPV to identify (pre)cancer.

Secondary objectives include the assessment of the diagnostic accuracy when combining 2 or 3 tests together; as well as the investigation of the effects of patient characteristics (i.e. age, parity, ART status, WHO HIV stage, CD4 cell count, history of previous treatment and sexually transmitted infections (STIs)) on the diagnostic accuracy.

# 2. Study methods

## 2.1 Trial design

A total of 450 study participants will be enrolled in a single site, Kanyama Clinic (Lusaka, Zambia). Study participants are HIV-infected women enrolled in the ART program at Kanyama Clinic, Lusaka, Zambia. Study participants will receive all tests and be informed about test results. Blinding of VIA, HR-HPV, and Gynocular™ assessments will be ensured as different assessors will perform each assessment independently.

### 2.2 Randomization

Not applicable

### 2.3 Sample size

See section “Selection of participants, sampling methods and sample size” p. 26 of the protocol.

### 2.4 Framework

Prospective, diagnostic test accuracy study

### 2.5 Statistical interim analyses and stopping guidance

No interim analysis planned

### 2.6 Timing of final analysis

All the outcomes will be analyzed collectively after data validation and database lock.

### 2.7 Timing of outcome assessments

| | Baseline visit | Follow-up |
|---|---|---|
| Time point | 0 | 6 months (+/- 1 month) |
| Enrolment, collection of baseline characteristics | x | |
| Tests: VIA, HR-HPV, Gynocular™, biopsy | x | x |

All outcomes will be analyzed collectively after study completion. After completion of data entry, data validation and cleaning will be performed. Data analysis will start after database lock.

### 2.8 Blinding

Not applicable.

# 3. Data management

## 3.1 Data export

The CRFs in this trial are implemented electronically using a dedicated electronic data capturing (EDC) system (REDCap®). The EDC system is activated for the trial only after successfully passing a formal test procedure. All data entered in the CRFs are stored on a Linux server in a dedicated Oracle database. Responsibility for hosting the EDC system and the database lies with Inselspital Bern.

Data will be exported in a R format (via the online export tool) by the trial statistician for data preparation, validation and analysis.

## 3.2 Data validation

Data checks are described in the following document: ..\04_CDM_851\Data_validation_V2.xlsx

## 3.3 Data preparation

- **Occurrence of cervical (pre)cancer (reference standard)**

For the main analysis, occurrence of cervical (pre)cancer will be assessed using the valid biopsies collected at baseline and at 6 months. The disease will be considered as present when at least one of the biopsies is positive. A biopsy is considered as positive when either a CIN stage of 2 or 3 or HSIL or a cancer is detected, and negative when either CIN stage 1 or LSIL or the absence of cancer is detected. Biopsy test will be considered as “undetermined” when the test was “invalid”. In case of missing or undetermined biopsy at baseline or at 6 months results will be assessed from the available valid biopsy. In case of two undetermined biopsies, we will use different strategies. For the main analysis we will code such cases as missing and remove them from the analysis. In sensitivity analysis, they will be coded as (i) positive cases and (ii) negative cases.

In sensitivity analysis, we will consider two alternative definitions of the reference standard. First, we will base our assessment on the baseline biopsy only. Second, we will use both biopsies but consider a biopsy as positive when either HSIL or cancer are detected.

- **Results of the diagnostic tests**

Results of the screening tests (i.e. Gynocular, HPV and VIA) as well as STI test will be assessed separately at baseline and at 6 months FUP.

*Gynocular*

*Results of the HPV screening test*

HPV screening test will be considered as positive when the test was valid AND when either HPV16 or HPV18 or 45 is positive or when 'Other high-risk HPV' is 'Yes'; valid tests with negative HPV16, 18 and 45 and no 'Other high-risk HPV' will be considered as negative. The test will be considered as "undetermined" when the test was "invalid". For the main analysis we considered the undetermined tests as missing and remove them from the analysis. In sensitivity analysis, they will be coded as (i) positive cases and (ii) negative cases.

| | |
|---|---|
| a. HPV 16 | ◯ Positive<br>◯ Negative |
| b. HPV 18 or 45 | ◯ Positive<br>◯ Negative |
| c. Other high risk HPV | ◯ Yes<br>◯ No |

*Results of the VIA screening test*

VIA assessment will be considered as positive when VIA is determined positive (abnormal) or suspicious for cancer, negative otherwise.

- **STI assessment**

STI infection will be considered as present in patients with a valid Trichomoniasis test and a positive Trichomoniasis result, and absent in patients with a valid Trichomoniasis test and a negative Trichomoniasis result. The test will be considered as "undetermined" when the test was "invalid". For the ma. The test will be considered as "undetermined" when the test was "invalid". For the main analysis we considered the undetermined tests as positive as missing and remove them from the analysis. In sensitivity analysis, they will be coded as (i) positive cases and (ii) negative cases.

# 4. Statistical principles

## 4.1 Confidence intervals and *P* values

All confidence intervals will relate to the 95% level and the two-sided significance level will be set at α = 0.05.

## 4.2 Analysis populations

### 4.2.1 Full analysis set (FAS)

The full analysis set (FAS) will include all enrolled subjects that have baseline data collected and at least one diagnostic test performed.

### 4.2.2 Per-protocol (PP)

Not applicable

### 4.2.3 Safety population

Not applicable

## 4.3 Estimands

The primary analysis of the primary and all secondary measures will be based on the FAS.

# 5. Trial Population

## 5.1 Screening data

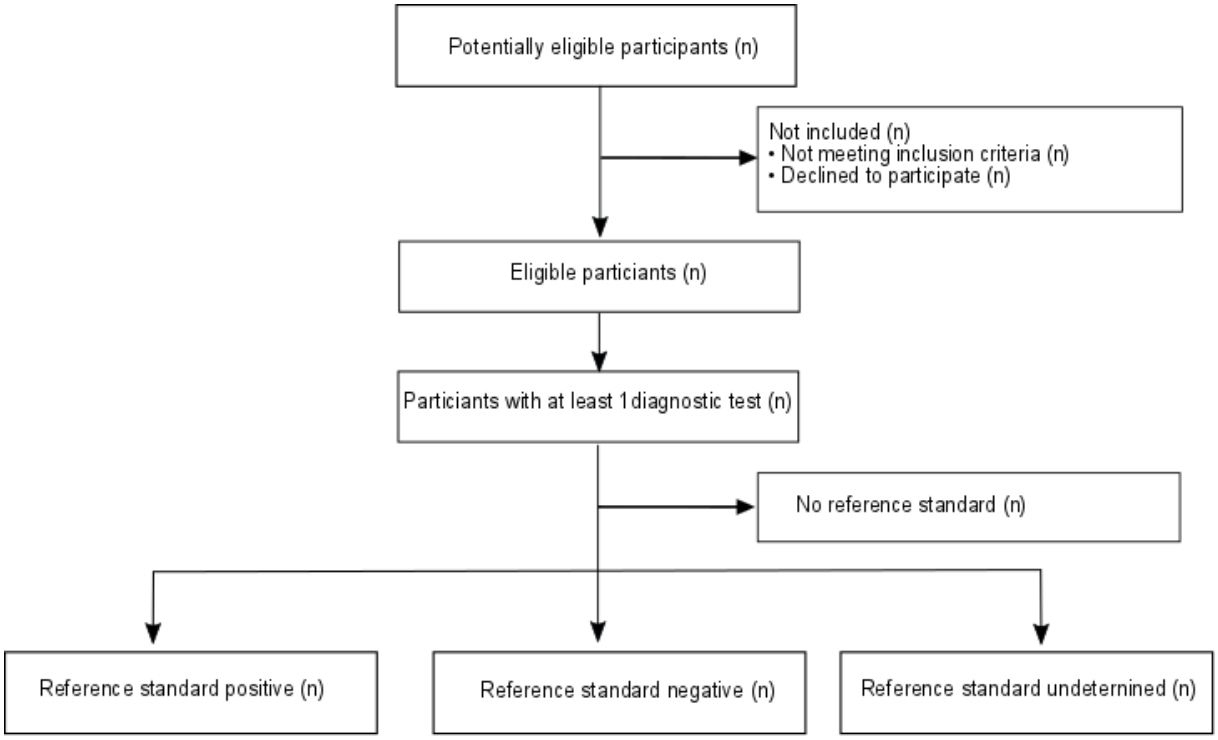
Not collected.

## 5.2 Eligibility

Women living with HIV, aged between 18 and 65 years. Women with a history of cervical cancer, who have been vaccinated against HR-HPV, have a previous hysterectomy, or are pregnant are excluded. Note: women who became pregnant after the baseline visit will be included in the analysis based on the data from the examinations at baseline.

## 5.3 Recruitment

A patient flow diagram will be drawn following the prototypical STARD diagram STARD-2015-flow-diagram.pdf




### 5.4 Baseline patient characteristics

Patient characteristics will be presented in a table stratified by disease status (i.e. biopsy negative vs. positive vs. undetermined) as number and percentage or mean and standard deviation for categorical and normally distributed continuous variables, respectively. For data severely deviating from a normal distribution, we will present median and interquartile ranges. Groups will be compared by the t-test for continuous data or by the non-parametric Wilcoxon test when normality assumption is not satisfied. For categorical data, Fisher's exact test will be used if expected frequencies are lower than 5 in any cell, else the chi-squared test will be applied.

Table x: Baseline table.

| Description | Variable | Type |
|---|---|---|
| Age | patient_age | Continuous: years |
| Age groups | patient_age | Categorical: 18-25, 26-35, 36-45, 46-55, 56-65 |
| Menopause | menopause_yn | Pre-menopausal vs. post-menopausal |

| Description | Variable | Type |
|---|---|---|
| Marital status | marital_status.factor | Categorical: Single, In a relationship/ Married, Separated/Divorced, Widowed, |
| Employment | employment_status.factor | Categorical:<br>Working (i.e. Employed part time/ Employed full time/ Self employed), vs. Not working (i.e. Able and seeking employment/ Not working, Able and not seeking employment/ Unable (caring for family)/not working, Unable (medical limitation)/ not working), Retired) |
| Education | education_level.factor | Categorical: did not finished secondary (i.e. None, Started primary, Finished primary, Started secondary), Finished secondary, more than secondary (i.e. College/University, Vocational training). |
| Income [Kwacha/month] | income_last_3m | Continuous: Kwacha/month |
| Smoking | smoking_yn.factor | Binary: Yes, No |
| Alcohol | alcohol_yn.factor | Binary: Yes, No |
| Age at sexual debut | first_intercourse_age | Continuous: years |
| Contraception use | contra_yn.factor | Binary: Yes, No |
| Contraception method | contra_pil_yn, contra_oral_yn, contra_iud_yn, contra_condoms_yn, contra_withdraw_yn, contra_implant_yn, contra_depot_provera_yn, contra_hyster_removed, contra_other_yn | Categorical:<br>Oral hormonal (i.e. The combined oral contraceptive pill/ the progesterone only pill), intrauterine device, condoms method, withdrawal method, implant, depot provera, Others |
| Gravidity | gravida_nr | Count: number of pregnancies |
| Parity | parity_nr | Count: Number of pregnancies carried to a viable gestational age |
| Parity categories | parity_nr | Categorical: 0,1-3, >3 |
| Trichomoniasis result | sti_trichom_res.factor | Binary: Positive, Negative |
| WHO stage of HIV at present? | hiv_who_stage_present.factor | Categorical: Stage 1, Stage 2, Stage 3, Stage 4, Do not know |
| On ART | art_yn.factor | Binary: Yes, No |
| Duration of ART use [years] | art_dur | Continuous: years |
| CD4 count [cells/mm3] | hiv_cd4 | Continuous: cells/mm3 |
| CD4 count [cells/mm3] | hiv_cd4 | Categorical: <200, 200-349, 350-499,>500 |

| Description | Variable | Type |
|---|---|---|
| HIV RNA load [copies/mL] | hiv_rna | Continuous: copies/mL |
| HIV RNA load [copies/mL] | hiv_rna | Categorical: Undetectable (< 50 copies/ml), 50 to 1 000 copies/ml, 1 000 to 10 000 copies/ml, > 10 000 copies/ml |

### 5.5 Procedural characteristics (if applicable)

Performance of the different screening (i.e. biopsy, VIA, Gynocular and HPV) and STI tests and availability of the results at baseline and at the 6 months visit will be presented in a table, stratified by visit, as number and percentage among all patients.

| Screening Test | | Baseline N = | 6 months N= |
|---|---|---|---|
| Biopsy | Sample taken (yes) | n [%] | n [%] |
| | Test valid (yes) | n [%] | n [%] |
| | Results available (yes) | n [%] | n [%] |
| | - Results positive | n [%] | n [%] |
| | - Results negative | n [%] | n [%] |
| | - Results undetermined | n [%] | n [%] |
| VIA | Assessment available (yes) | n [%] | n [%] |
| Gynocular | Questions of the SWEDE score are all addressed (yes) | n [%] | n [%] |
| HPV | Sample taken (yes) | n [%] | |
| | Test valid (yes) | n [%] | n [%] |
| | Results available (yes) | n [%] | n [%] |
| | - Results positive | n [%] | n [%] |
| | - Results negative | n [%] | n [%] |
| | - Results undetermined | n [%] | n [%] |
| STI | Sample taken (yes) | n [%] | |
| | Test valid (yes) | n [%] | n [%] |
| | Results available (yes) | n [%] | n [%] |

| Screening Test | | Baseline<br>N = | 6 months<br>N= |
|---|---|---|---|
| | - Results positive | n [%] | n [%] |
| | - Results negative | n [%] | n [%] |
| | - Results undetermined | n [%] | n [%] |

-

### 5.6 Adherence and protocol deviations

- The occurrence of protocol deviation (as defined in the section 4.2.2) will be checked and reported in the flow chart.

### 5.7 Withdrawal/follow-up

- Number and percentage of withdrawal and lost to follow-up will be presented in the flow chart

# 6. Analysis

## 6.1 Definition of outcome measures

### 6.1.1 Primary outcome measure

- **Diagnostic measure:**

Sensitivity and specificity of each screening test (i.e. HPV, VIA and Gynocular) performed at baseline

- **Derivation:**

*Step 1: Assessment of the occurrence of cervical (pre)cancer (reference standard)*

Main definition

- Definition: occurrence of cervical (pre)cancer will be assessed using the valid biopsies collected at baselines and at 6 months. The disease will be considered as present when at least one of the two biopsies is positive. Biopsy is considered as positive when either a CIN stage of 2 or 3 or HSIL or a cancer are detected, and negative when either the absence of cancer, CIN stage 1 or LSIL is detected. Biopsy test will be considered as "undetermined" when the test was "invalid". In case of missing or undetermined biopsy at baseline or at 6 months results will be assessed from the available valid biopsy. In case of two undetermined biopsies, we will use different strategies. For the main analysis we considered these cases as missing and remove them from the analysis. In sensitivity analysis, they will code them as (i) positive cases and (ii) negative cases.
- Derivation:
  - Positive: (biopsy_test_valid=valid & cin_stage %in% c(2,3)) OR (biopsy_test_valid=valid & biopsy_hsil_res=1) OR (biopsy_test_valid=valid & biopsy_cancer_res= positive) at baseline or at 6 months FUP
  - Negative: (biopsy_test_valid=valid & cin_stage %in% c(1)) OR (biopsy_test_valid=valid & biopsy_lsil_res=1) OR (biopsy_test_valid=valid & biopsy_cancer_res= negative) at baseline and at 6 months FUP
  - Undetermined: biopsy_test_valid=invalid
- Variable type: Binary: No, Yes

Alternative definition based on baseline biopsy (for sensitivity analysis)

- Definition: occurrence of cervical (pre)cancer will be assessed using the valid biopsies collected at baselines. The disease will be considered as present when at least one of the biopsies is positive. Biopsy is considered as positive when either a CIN stage of 2 or 3 or HSIL or a cancer are detected, and negative when either the absence of cancer, CIN stage 1 or LSIL

is detected. Result of undetermined and missing biopsies will be replaced by the results obtained at 6 months. In case of undetermined biopsy at both baseline and 6 months, the test result will be considered as missing.

- Derivation:
  - Positive: (biopsy_test_valid=valid & cin_stage %in% c(2,3)) OR (biopsy_test_valid=valid & biopsy_hsil_res=1)OR (biopsy_test_valid=valid & biopsy_cancer_res= positive) at baseline
  - Negative: (biopsy_test_valid=valid & cin_stage %in% c(1)) OR (biopsy_test_valid=valid & biopsy_lsil_res=1)OR (biopsy_test_valid=valid & biopsy_cancer_res= negative) at baseline
- Variable type: Binary: No, Yes

Definition based on both baseline at 6 months biopsies (for sensitivity analysis)

- Definition: occurrence of cervical (pre)cancer will be assessed using the valid biopsies collected at baselines and at 6 months. The disease will be considered as present when at least one of the biopsies is positive. Biopsy is considered as positive when either HSIL or cancer is detected, and negative when LSIL or absence of cancer are detected. Biopsy will test will be considered as “undetermined” when the test was “invalid”. In case of missing or undetermined biopsy at baseline or at 6 months results will be assessed from the available valid biopsy. In case of two undetermined biopsies, we will use different strategies. For the main analysis we considered these cases as missing and remove them from the analysis. In sensitivity analysis, they will code them as (i) positive cases and (ii) negative cases.
- Derivation:
  - Positive: (biopsy_test_valid=valid & biopsy_hsil_res = Positive OR (biopsy_test_valid=valid & biopsy_cancer_res= positive) at baseline or at 6 months FUP
  - Negative: (biopsy_test_valid=valid & biopsy_lsil_res = Positive OR (biopsy_test_valid=valid & biopsy_cancer_res= negative) at baseline and at 6 months FUP
- Variable type: Binary: No, Yes

*Step 2: Derive test results at baseline)*

| Test | eCRF sheet | Derivation | variable type |
|---|---|---|---|
| HPV | Test results at baseline | hpv_test_valid =”valid“<br>AND<br>(either hpv_16_res=”positive” OR hpv_18_45_res=”positive” OR hpv_other_yn=”yes”) | Binary: positive, negative |

| | | | |
|---|---|---|---|
| | | The test will be considered as "undetermined" when the test was "invalid" (hpv_test_valid ="invalid"). For the main analysis we considered the undetermined tests as missing and remove them from the analysis, in sensitivity analysis, they will code them (i) as positive cases (ii) as negative cases. | |
| VIA | VIA Examination at baseline | via_assessment_result = "VIA positive (abnormal)" Or "Suspicious for cancer" | Binary: positive, negative |
| Gynocular | General And Gynocular Exam at baseline | The SWEDE score (variable colposcopy_score) will be calculated and categorized into a "low risk,negative" and " high risk, positive" categories. The cut-off will be defined as the value optimizing both sensitivity and specificity (Youden cut-off) | Binary: positive, negative |

*Step 3: Derivation of diagnostic measure*

For each test, we will build a two-by-two table (or a three-by-three table in case of occurrence of undetermined results), where test findings are measured against the reference standard (i.e. biopsy finding). We will then estimate the sensitivity and specificity of each test.

### 6.1.2 Secondary outcome measures

i. Other statistics measuring test accuracy of VIA, HR-HPV testing and Gynocular™ at baseline.

For all the tests:

We will use the method described above to calculate: positive and negative predictive values, positive and negative likelihood ratios, false positive rate, false negative rate, number needed to screen and the diagnostic odds ratios. Diagnostic odds ratios with 95% confidence intervals will be calculated for comparison of diagnostic accuracies between tests. This is defined as the ratio of the odds of test positivity in subjects with the disease to the odds of test positivity in subjects without the disease assessed.

For the Gynocular exam:

The accuracy of the SWEDE score will be assessed by ROC curve analysis and summarized by the area under the ROC curve (AUC).

### 6.1.3 Additional secondary outcome measures (analyzed in subproject(s), see section 6.6)

- Investigations to inform telemedicine capacity

i. Description of image quality images obtained by the Gynocular™ from static assessors, in terms of the assessors' ability to adequately evaluate the images.
ii. Proportion of correctly diagnosed (pre)cancer HSIL through static images obtained by the Gynocular™. (Pre)cancer patients will be defined as (i) HSIL or cancer patients and (ii) CIN2+ or cancer patients.
iii. Area under the ROC for static image Swede score obtained by Gynocular™ in WLHIV.
iv. Cohen's kappa coefficient to assess agreement between live and static assessors.

-
- Artificial Intelligence for improving the detection of precancerous cervical lesions

i. Test accuracies of AI deep learning tool retrospectively using coded images obtained in the study through the Gynocular™ colposcope and smartphone images.
ii. Inform and improve AI deep learning tools to detect HSIL by using images or GIFs obtained in the study through the Gynocular™ colposcope and smartphone images.

### 6.1.4 Safety outcomes

Adverse events will be recorded. We will ask participants to inform the study nurse or assistant if they develop any of these, in person or by phone.

Information describing these events will include: time of onset, duration, resolution, action to be taken, assessment of intensity, relationship with study treatment. The clinical team will also complete assessments of seriousness, causality and expectedness. Participants are asked to inform the clinical team immediately if any AE occur in the 9 months following their initial enrolment. At each visit, the Study nurse/assistant will also inquire about any adverse events that have occurred, for verification.

## 6.2 Analysis methods

### 6.2.1 Primary analysis

Primary and secondary accuracy measures of the different stand-alone tests will be derived as described above and summarized in a table showing point estimates with 95% confidence intervals. The diagnostic odds ratio is defined as the ratio of the odds of test positivity in subjects with the disease to the odds of test positivity in subjects without the disease. The number needed to screen is defined as the number of patients that need to be tested to detect one correct positive case.

Sensitivity, specificity, positive and negative predictive values, as well as false positive and false negative rates will be presented with a Wilson score confidence interval. Other measures will be accompanied by a normal approximation confidence interval. For the Gynocular exam (Swede score), we will additionally report the area under the ROC curve.

| | | Primary measures | | Secondary measures | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test | Cut-off | Sens. [CI] | Spec. [CI] | PPV [CI] | NPV [CI] | PLR [CI] | NLR [CI] | FP [CI] | FN [CI] | NNS [CI] | DOR [CI] | AUC [CI] |
| HPV | NA | | | | | | | | | | | NA |
| VIA | NA | | | | | | | | | | | NA |
| Gynocular | | | | | | | | | | | | |

Sens.: sensitivity

Spec: specificity

PPV: positive predictive value

NPV: negative predictive value

PLR: positive likelihood ratio

NLR: negative likelihood ratio

FP: false positive rate

FN: false negative rate

NNS: Number needed to screen

DOR: Diagnostic odds ratios

AUC: area under the curve

The prevalence of the disease will be reported as well.

### 6.2.2 Secondary analyses

In the primary analysis, all patients will be analyzed using the FAS. In a secondary per-protocol analysis, primary and secondary accuracy measures of the different stand-alone tests will be evaluated as described above based on the PP analysis set.

### 6.2.3 Sensitivity analyses

The primary analysis will be conducted deriving the reference standard from the two alternative definitions given in section 6.1.1.

In case of undetermined results for the biopsy, HPV and STI tests, the primary analysis will be done by considering the undetermined results as (i) positive cases and as (ii) negative cases.

To explore the possibility of a “training effect”, the primary analysis will be conducted separately in the first 10% recruited patients and in the others.

### 6.2.4 Subgroup analyses

We will evaluate test accuracies in subgroups defined by age (< 25, 26-35, 36-45, >46 and pre-menopausal vs. post-menopausal), parity, ART status, co-infection with STIs, methods of contraception, CD4 cell count.

| Subgroup | eCRF sheet | Variable | Categorization |
|---|---|---|---|
| Age | crf_02_demographic_history | age | < 25, 26-35, 36-45, >46 |
| Menopause | crf_03_medical_history | menopause_yn | Binary: Pre-menopausal vs. post-menopausal |
| Parity | crf_03_medical_history | parity_nr | 0,1-3,>3 |
| ART | crf_03_medical_history | art_yn | Binary: Yes, No |
| STI | crf_08_14_test_results | STI infection will be considered as present in patient with a valid Trichomoniasis test and a positive Trichomoniasis results at baseline trichom_test_valid and sti_trichom_res | Binary: Yes, No |
| Methods of contraception | crf_03_medical_history | contra_pil_yn, contra_oral_yn, contra_iud_yn, contra_condoms_yn, contra_withdraw_yn, contra_implant_yn, contra_depot_provera_yn, contra_hyster_removed, contra_other_yn | Long acting reversible contraception (implant, depot provera, intrauterine device); oral hormonal, condoms method, withdrawal & others methods, none |
| HIV RNA | crf_08_14_test_results | hiv_rna | Binary: suppressed yes (<1000 copies/ml) versus no (≥1000 copies/ml) |
| CD4 count (cells/mm3) | crf_08_14_test_results | hiv_cd4 | <200, 200-349, 350-499,>500 |
| Education | crf_02_demographic_history | education_level | Categorical: did not finish secondary (i.e. None, Started primary, Finished primary, Started secondary), Finished |

| Subgroup | eCRF sheet | Variable | Categorization |
|---|---|---|---|
| | | | secondary, more than secondary (i.e. College/University, Vocational training). |
| History of treatment for pre-cancer | crf_03_medical_history | pre-cancer_treatm_yn | Binary: yes versus no |

Subgroup analyses will be based on the treatment policy strategy using the FAS. For binary variables (i.e. menopause, ART, STI, HIV RNA, history of treatment for pre-cancer), subgroup analyses will only be conducted if each of the subgroup contains 10% or more of the enrolled patients.

### 6.2.5 Additional analyses

All additional analyses will be based on the treatment policy strategy using the FAS.

- **Association of patient characteristics with the diagnostic tests and disease status**

We will run univariable and multivariable logistic regression analyses to investigate the possibility of an effect of patient characteristics on the result of each diagnostic test as well as on the disease status (biopsy result). We will consider the following characteristics:

- age
- Menopause
- parity
- ART status
- sexually transmitted infections (STIs) at baseline
- Methods of contraception
- HIV RNA
- CD4 cell count
- History of treatment for precancer
- Education

- **Effect of patient characteristics on the association between diagnostic test and disease status**

We will investigate possible effect modification by the characteristics listed above. We will use logistic regression models to test interaction between patient characteristic and each diagnostic test:

Disease status ~ diagn_test + characteristic + diagn_test x characteristic

Models will be run crude as well as adjusted for other patient characteristics listed above.

- **Diagnostic test accuracy of the combined tests**

We will address the diagnostic accuracy of the tests when they are done in sequence. For each combination, we will calculate sensitivity and specificity, positive and negative predictive values, and positive and negative likelihood ratios. We will consider the following scenarios and sequences.

Result of the Gynocular test will be assessed using the SWEDE score. We will use the cut-off value defined when the test was used as a "stand alone" (i.e. Youden cut-off) to separate "low risk, negative" and "high risk, positive" patients. Additionally, we investigate the use of other cut-offs maximizing either sensitivity (e.g. yielding sensitivity $\geqslant$90%), or specificity (yielding specificity $\geqslant$90%).

| Test1 | Test 2 | Considered scenario |
|---|---|---|
| VIA | Gynocular | T1 (+)/ T2 (i.e. first test positive and second test positive vs negative);<br>T1(-)/T2 (i.e. first test negative and second test positive vs negative) |
| HPV | Gynocular | T1 (+)/ T2 |
| HPV | VIA | T1 (+)/ T2 |
| VIA | HPV | T1 (+)/ T2<br>T1(-)/T2 |
| Gynocular | HPV | T1 (+)/ T2<br>T1(-)/T2 |

If it makes sense, we will also explore the combination of three tests. The studied scenario would notably include (i) HPV(+), VIA(+), Gynocular and (ii) HPV(+), VIA (-), Gynocular.

- **Characterizing co-infections of premalignant and malignant disease (STIs/HR-HPV)**

We will present in a table the number and proportion of patients with positive and negative STI and HR-HPV separate for each stage of CIN and the occurrence (or not) of persistent disease.

### 6.2.6 Additional descriptive analysis

- **The effectiveness of treatment**

We will report the proportion of women treated who remain free from disease 6 months after treatment. Women with a negative biopsy at 6 month will be considered as “free of disease”.

- **Test accuracy in treated patients**

We will also use the results of the 6 months tests to investigate accuracy of tests in women who have undergone treatment after the baseline visit. The occurrence of cervical (pre)cancer will be assessed using the valid biopsies collected at 6 months; HPV, VIA and Gynocular test results will be derived as described in section 6.1.1. For the Gynocular test we will use the cut off defined when analysing the baseline results.

## 6.3 Interim analyses

Not applicable.

## 6.4 Missing data

If there is substantial amount of missing data that could confound the interpretation of results, we will consider multiple imputation by chained equation in sensitivity analyses.

## 6.5 Safety evaluation

Adverse events occurring in the 9 months following patient enrolment will be listed. Information describing adverse events will include: time of onset, duration, resolution, action to be taken, assessment of intensity, relationship with study treatment as well seriousness, causality and expectedness.

## 6.6 Subproject (methodology will be further described in additional documents)

Secondary outcomes listed in the section 6.1.3 “Additional secondary outcomes (will be analyzed separately)”, are actually not recorded in the database and will be analyzed as two distinct subprojects. Their analysis will not be described in this SAP.

### 6.6.1 Telemedicine

Cervical images obtained from the study will be used to assess both the quality of pictures by trained colposcopists (nurses and doctors) and the accuracy of diagnoses made by assessment of static images. The agreement between live and static assessors will be assessed by Cohen's kappa coefficient.

### 6.6.2 Machine learning

The accuracy of an already established Artificial intelligence (AI) deep learning tool retrospectively using coded images obtained in this study through the Gynocular™ colposcope and smartphone images will be tested.

Moreover the images and/or GIFs obtained in this study through the Gynocular™ colposcope and smartphone images will be used to further inform and improve AI deep learning algorithms for detection of cervical cancer.

## 6.7 Statistical software

All the analysis will be done using R (version 3.3.0).

## 6.8 Quality control

The scripts used for data preparation and analysis as well as the corresponding statistical report will be reviewed by an independent statistician that is not otherwise involved in the study.

# 7. Changes from the protocol

Changes from protocol

| Header | Change | Reason |
|---|---|---|
| Change in the definition of the primary outcome. | New definition: "The disease will be considered as present when at least one of the biopsies is positive. Biopsy are considered as positive when either a CIN stage of 2 or 3 or HSIL or a cancer are detected".<br>In sensitivity analysis, we will analyse the primary outcome as initially defined in the protocol (i.e. a biopsy as positive when either HSIL or cancer are detected.) | The definition of the primary outcome is now in line with what is usually published |
| Subgroup analysis | Subroup analysis:<br>- will not be conducted for WHO HIV stage<br>- Will be conducted for Menopause, Methods of contraception, HIV RNA, Education | According to the literature, these are relevant factors to consider |
| Additional analysis, association of patient characteristics with the diagnostic tests and disease status | We will not consider the effect of WHO HIV stage. However, we will analyzed the effect of Menopause, Methods of contraception, HIV RNA, Education. | According to the literature, these are relevant factors to consider |

# 8. References