Rwanda VMB Study
Statistical Analysis Plan of Primary Endpoints and Secondary Clinical Outcomes

Author: Marijn Verwijs, Project Manager
Supervisor: Janneke van de Wijgert, Professor and Chief Investigator
Version date: 16 September 2016

Full title	Preparing for a clinical trial of interventions to maintain normal vaginal microbiota for preventing adverse reproductive health outcomes in Africa
Funding	UK Medical Research Council (Development Grant MR/M017443/1)
Sponsor	University of Liverpool (UoL)
Protocol number	UoL001129
Study site	Rinda Ubuzima (RU) KN 50th Avenue (formerly Rue Akagera 716), Kiyovu, Kigali, Rwanda P.O. Box 4560, Kigali, Rwanda Tel: +250-252-503431 Fax: +250-252-503389
Chief investigator	Professor Janneke van de Wijgert, MD PhD MPH UoL Institute of Infection and Global Health
Principal Investigator	Stephen Agaba, MD Rinda Ubuzima (see contact details above) Email: agabask@gmail.com
Data and sample collection period	5 June 2015 – 23 February 2016
Protocol approved by	Rwanda National Ethics Committee, Rwanda Ministry of Health National Health Research Committee, UoL Sponsorship Committee, and UoL Research Ethics Sub-Committee for Physical Interventions.
Trial Steering Committee	1. Prof. Jacques Ravel (see contact details above) 2. Dr. Adrian Nivoliez, representing Biose (formerly Probionov), Arpajon-sur-Cère, France 3. Dr. Jessica Younes, representing Winclove Probiotics, Amsterdam, Netherlands 4. Dr. Tania Crucitti, Institute of Tropical Medicine, Antwerp, Belgium 5. Dr. Vicky Jespers, Institute of Tropical Medicine, Antwerp, Belgium 6. Prof. Anna Maria Geretti, UoL Institute of Infection and Global Health

Table of Contents

1. Introduction

This Statistical Analysis Plan (SAP) describes the primary analyses, and one closely related secondary analysis, of the study protocol entitled: "Preparing for a clinical trial of interventions to maintain normal vaginal microbiota for preventing adverse reproductive health outcomes in Africa" (referred to hereafter as the Rwanda VMB Study). The main aim of this study is to determine the safety and preliminary efficacy of three interventions to prevent recurrence of bacterial vaginosis (BV). The study was conducted at Rinda Ubuzima (RU) in Kigali, Rwanda, and was sponsored by the University of Liverpool, Liverpool, UK.

It should be noted that, at the request of the UK Medical Research Council (MRC), the Rwanda VMB Study is a pilot study. Therefore, no formal sample size calculations were done and the sample size was mostly determined by the available budget.

1.1 Study overview

A flow diagram of study visits, and numbers of women seen at each visit, is shown in Figure 1. Women at high risk for HIV, sexually transmitted infections (STIs), and BV were targeted using community mobilisers who were high risk women themselves. Women who were potentially eligible (HIV-negative, non-pregnant, aged 18-45, and sexually active with high risk of HIV/STI/BV) and interested in participating were invited to visit the RU clinic for screening procedures. Please refer to the study protocol for a full listing of eligibility criteria. At the screening visit, a face-to-face interview, speculum examination, and testing for HIV, pregnancy, urinary tract infection (UTI), BV (by Amsel criteria and Nugent scoring), and *Trichomonas vaginalis* (TV; by wet mount and InPouch culture) were conducted. At the end of the visit, preliminary eligibility was assessed, and initial treatments and referrals were given based on available test results. Additional STI testing was conducted on stored samples after the participant had left the clinic. Results of those tests were shared with each participant at a Results Visit 1 (scheduled 2 weeks later), additional treatments were given as required, and eligibility was reassessed. At the end of the Screening and Results Visits cascade, women were either declared ineligible or diagnosed with BV by Nugent score and/or TV on wet mount or by culture and treated with 500mg oral metronidazole twice per day for 7 days.

At the Enrolment Visit, only women whose BV/TV treatment had been successful (no BV by Amsel criteria and no TV on wet mount), who were still not pregnant, and who were free of STIs and UTI, were randomised to 4 groups as follows:
Group 1) Behavioural 'vaginal practices cessation and safer sex' counselling only (also referred to as the 'control group');
Group 2) Behavioural counselling plus oral metronidazole 500mg, twice per week for two months;
Group 3) Behavioural counselling plus Ecologic Femi vaginal capsule, once per day for 5 days immediately after oral metronidazole treatment followed by thrice weekly for two month;
Group 4) Behavioural counselling plus Gynophilus vaginal tablet, once every 4 days for two months.
Participants took/inserted the first dose of their intervention (if applicable) under direct observation at the Enrolment Visit.

Participants used their study products (if applicable) for 2 months, and returned for study visits after seven days (D7), one month (M1) and two months (M2; cessation of product use), and then again at six months (M6; 4 months after cessation of product use). At these follow-up visits, participants underwent a face-to-face interview, speculum examination and a variety of laboratory tests (BV by Amsel criteria and Nugent scoring, TV by wet mount and InPouch culture, candidiasis by wet mount, and additional testing if clinically indicated). At the M6 visit, HIV, STI, pregnancy and UTI tests were repeated. Please refer to Table 1 below for a full schedule of assessments by study visit.

Figure 1: Flow diagram of study visits and numbers of women seen at each visit

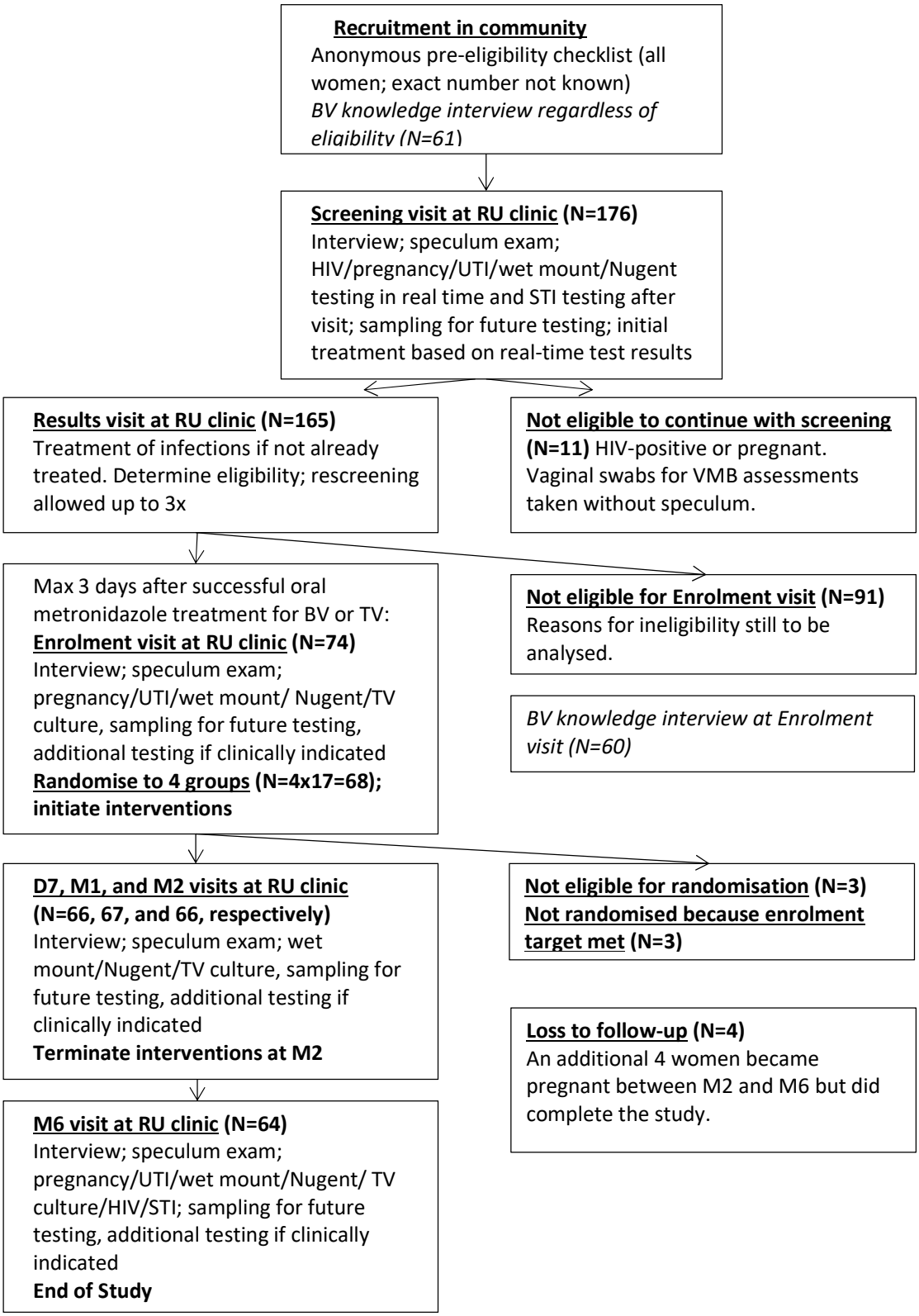

Table 1: Schedule of assessments

Procedures in clinic during visits	Screening/results	Enrolment	D7, M1, M2	M6
Informed consent	X	X		
Contact information	X	X	X	X
Assign PIN	X			
Eligibility checklist	X	X		
Randomisation: initiate interventions		X		
Cease interventions			X (M2)	
Face-to-face interview, quantitative	X	X	X	X
Discussion/interview, qualitative				
Assess AEs and social harms		X	X	X
HIV pre- and post-test counselling	X			X
HIV/STI/BV risk reduction counselling	X	X	X	X
Blood collection (10 ml)	X			X
Urine collection	X	X		X
Rapid HIV testing (whole blood)[1]	X			X
Rapid pregnancy testing (hCG urine)	X	X	[X]	X
Rapid UTI testing (urinalysis dipstick)	X	X	[X]	X
Speculum exam: visual inspection	X[2]	X	X	X
Speculum exam: vaginal pH + wet mount (TV, candida, BV by Amsel[2])	X[2,3]	X[3]	X[3]	X[3]
Speculum exam: vaginal swabs	X	X	X	X
Speculum exam: endocervical swabs	X[2]	X	[X]	X
Speculum exam: CVL	X[2]	X	X	X
Provide results and treatment if appl[1]	X	X	X	X
Laboratory procedures after visits				
Blood: syphilis and HSV-2 serology	X			X[4]
Endocervical swab: NG/CT PCR	X	[X][5]	[X][5]	X
Vaginal swabs: TV InPouch culture	X	X	X	X
Vaginal swabs: Nugent scoring	X	X	X	X
Vaginal swabs: store for VMB MiSeq	X	X	X	X
Laboratory procedures if funding permits:				
Vaginal swabs: store for additional VMB assessments (e.g. qPCR, biofilm)	X	X	X	X
CVL: store for additional VMB assessments	X	X	X	X
Vaginal smears: store for biofilm assessments	X	X	X	X

AE=adverse event; BV=bacterial vaginosis; CT=*Chlamydia trachomatis*; CVL=cervicovaginal lavage; D7=Day 7 visit; HPV=human papillomavirus; HSV-2=herpes simplex type 2; M1/2/6=Month 1/2/6 visit; NG=*Neisseria gonorrhoeae*; PIN=participant identification number; qPCR=quantitative PCR; STI=sexually transmitted infection; UTI=urinary tract infection; TV=*Trichomonas vaginalis*; VMB=vaginal microbiome.

1. By the algorithm recommended in the most recent Rwandan HIV voluntary counselling and testing guidelines
2. HIV-positive and pregnant women did not undergo a speculum exam unless clinically indicated but they were asked to donate vaginal swabs for VMB assessments.
3. BV (by Nugent) was always treated during the screening process but was only treated when symptomatic and positive by Amsel score at the Enrolment visit and during follow-up. Candida on wet mount was also only treated when symptomatic. TV on wet mount was always treated.
4. For HSV-2 only when negative at baseline.
5. Endocervical swabs were tested for NG/CT by PCR at screening and M6, but may also be tested at enrolment and/or D7/M1/M2 if deemed necessary for clinical reasons or to accurately interpret the VMB findings.

1.2 Visit windows and retention definitions

Visit windows for the study visits were as follows:

- Screening + results visits: the entire screening process should not have taken longer than 6 weeks
- Enrolment visit: maximum of 3 days after completion of oral metronidazole treatment
- Day 7: Day 5 – Day 9
- Month 1: ± 7 days
- Month 2: ± 7 days
- Month 6: ± 7 days

The following retention definitions were used:

- Early withdrawal: The participant informed a RU staff member that she no longer wants to participate in the study.
- Late visit: The participant did attend a study visit but after the visit window had closed.
- Skipped visit: The participant did not attend a study visit but did return to the study clinic for a subsequent study visit.
- Lost-to-follow-up: RU does not label a participant as lost to follow-up until study close-out. A participant is considered lost to follow-up at study close-out when multiple contact attempts by RU staff failed (as specified in the RU SOP for Retention) and the participant did also not return to the study clinic spontaneously before study close-out.
- Missed visit: All visits that should have taken place according to the study protocol but did not. This includes all study visits after an early withdrawal or after the last study visit made by a participant who is lost to follow-up, as well as all skipped visits.

2. Primary and secondary clinical endpoints

The endpoints of the study discussed in this SAP are:

Primary clinical objectives	Endpoints
1. To determine the **safety** of the three biomedical interventions compared to the control group.	– Self-reported solicited and unsolicited (serious) adverse events (AEs) and social harms – Clinician-observed speculum exam findings
2. To determine the **preliminary efficacy** of the three biomedical interventions compared to the control group **in reducing BV recurrence without increasing vaginal candidiasis incidence.**	– BV by Amsel criteria, Nugent score 4-10* (versus 0-3), TV by culture and symptomatic vaginal candidiasis at any time during the intervention (Day 7, Month 1, Month 2) – BV by Amsel criteria, Nugent score 4-10* (versus 0-3), TV by culture, and symptomatic vaginal candidiasis after cessation of the intervention (Month 6) * Abnormal Nugent scores will also be analysed as Nugent score 7-10 versus 0-6.
Secondary clinical objectives	Endpoints
2. To determine **preliminary efficacy** of the three biomedical interventions compared to the control group **in reducing the incidence of STIs and UTI**.	- Incidence of the following STIs between screening and Month 6: HIV, HSV-2, syphilis, CT and NG - Incidence of UTI between screening and Month 6

3. Participant flow

3.1 Retention (see also the flow diagram in Figure 1 above)

	N
Number of women who attended a Screening visit	176
Number of women who attended an Enrolment visit	74
Number of women who were randomised	68
Number of women who attended D7, M1, M2, and M6 visits	66, 67, 66, 64
Number of women who completed all scheduled study visits	60
Number of women who missed visits after randomisation	8
Number of women who withdrew early	4
Number of women who were lost-to-follow-up	0
Number of deaths	0
Number of completed study visits by randomised women, including the Enrolment visit	324
Number of skipped visits after randomisation	4
Number of missed visits after randomisation	12
Total person-months of data collected after randomisation	TBD
Total person-days of data collected after randomisation	TBD

3.2 Timing and reasons for early withdrawals

PID	Last scheduled visit attended	Reason for early withdrawal
058	M1	Participant came at M2 but had no symptoms anymore; she did not want to be investigated any further.
082	M2	Participant moved away from Kigali between M2 and M6.
093	Enrolment	Participant was verbally 'harassed' by partner and sister.
139	M2	Participant moved away from Kigali between M2 and M6.

3.3 Reasons for ineligibility during the screening process and at enrolment

During the screening process	N
HIV positive	
Pregnant	
Not at high risk of HIV/STIs/BV (defined as having had more than one sexual partner in the last 12 months OR having been treated for an STI and/or BV in the last 12 months)	
No BV by Nugent score or TV by wet mount or culture	
Clinician-observed genital ulcers, condylomata, or other genital abnormalities at the Screening visit	
Underwent a gynaecological surgery/invasive procedure in the 3 months prior to screening	
History of significant urogenital prolapse, undiagnosed vaginal bleeding, urine or faecal incontinence, or blood clotting disorders	
Allergic to metronidazole or any other components of the study drugs	
Participating in another health intervention study	
Other (each reason will be listed)	
TOTAL number of women	**102**

At the Enrolment visit	N
Enrolment target already met	3
Did not adhere to 7-day metronidazole treatment	
BV or TV not adequately treated	
Pregnant at the Enrolment visit	
STI or UTI at the Enrolment visit	
Other (each reason will be listed)	
TOTAL number of women	6

3.4 Numbers of and reasons for unscheduled visits

Time period	Number of women	Number of visits
Between the Screening and Enrolment visits (not including Results visits scheduled by RU staff)		
Between Enrolment and M1		
Between M1 and M2		
Between M2 and M6		
After M6		
TOTAL		

Reasons for unscheduled visits	Number of visits
To withdraw informed consent	
To obtain additional study product supplies	
To report an (S)AE or social harm	
To test for an STI or vaginal infection	
To test for a UTI	
To collect treatment for a positive lab result	
Wants to be tested for HIV or pregnancy	
For male partner HIV/STI testing and couple counselling	
To ask questions or express concerns	
Other (each reason will be listed)	
TOTAL	

4. Analyses of baseline data

4.1 Characteristics of the total screened and randomised populations (collected at Screening)

All participants who attended a Screening visit underwent a face-to-face interview, and an HV and pregnancy test. Women who tested positive for HIV or pregnancy did not always undergo a speculum examination or any further diagnostic testing.

Sociodemographics	Screened (N=174)	Randomised (N=68)
Kigali neighbourhood (n/N % per category)		
Age (median, IQR)		
Age, categorised (n/N % per category) - 18 to 25 - 26 to 35 - 36 to 45		
Marital status (n/N % per category)		

Education level (n/N % per category)		
Sexual history and behaviour	**Screened (N=174)**	**Randomised (N=68)**
Number of sex partners in lifetime (median, IQR)		
Number of sex partners last 12 months (median, IQR)		
Number of sex partners last month (median, IQR)		
Currently has a main sex partner (n/N %)		
Lives together with main sex partner (n/N %)		
Main sex partner circumcised (n/N % per category) - Yes - No - Does not know		
Main sex partner has other sex partners (n/N % per category) - Yes - No - Does not know		
Vaginal sex frequency last 2 weeks (median, IQR)		
Any condom use in past two weeks (n/N %)		
Condom use during last sex act (n/N %)		
Exchanged sex for money/goods in past month (n/N %)		
Reproductive and contraceptive history	**Screened (N=174)**	**Randomised (N=68)**
Pregnancies in lifetime (median, IQR)		
Deliveries in lifetime (median, IQR)		
Miscarriages in lifetime (median, IQR)		
Currently breastfeeding (n/N %)		
Ever severe infection with fever during pregnancy/childbirth (n/N %)		
Ever baby with severe infection with fever within days after birth (n/N %)		
Currently using a modern method of contraception other than condoms (n/N % per category): - Pills - Injections - Implant - Copper IUD - Sterilised - Other (described in footnote)		
Medical history	**Screened (N=174)**	**Randomised (N=68)**
Any significant medical and/or surgical history or current condition (n/N %; describe in footnote)		
Ever had surgery (n/N %; describe in footnote)		
Any gynaecological surgery / invasive procedure last 3 months (n/N %)		0*
Any chronic diseases (n/N %; describe in footnote)		
Past or current significant urogenital prolapse (n/N %)		0*
Past or current undiagnosed vaginal bleeding (n/N %)		0*
Past or current urine or faecal incontinence (n/N %)		0*
Past or current blood clotting disorder (n/N %)		0*
Ever had allergic reaction (n/N %)		0*

Ever treated for BV (n/N %)		
Ever treated for vaginal candidiasis (n/N %)		
Treated for STI in the past 12 months (n/N %)		
Current medications and symptoms	**Screened (N=174)**	**Randomised (N=68)**
Any current medication use (n/N %; describe in footnote)		
Any urogenital symptoms now or in last 2 weeks, patient-reported (n/N%)		
If yes, what symptoms[1] (n/N % per category): – Burning when passing urine – Frequent urination or urgent need to urinate – Blood in urine – Genital burning – Genital itching – Pain during sex – Lower abdominal pain – Unusual vaginal discharge – Sores in the genital and/or anal area (including buttocks) – Other (additional categories to be added if needed)		
If yes, already received treatment (n/N %)		
If yes, some or all of the symptoms are currently ongoing (n/N %)		
Medical procedures by study physician	**Screened (N=174)**	**Randomised (N=68)**
Any abnormal pelvic exam findings (n/N % per category)		
Any abnormal bimanual exam findings (n/N % per category)		
Clinical diagnosis (n/N % per category)		
Treatment given (n/N % per category)		
Laboratory results	**Screened (N=174)**	**Randomised (N=68)**
HIV serology positive[2] (n/N %)		0*
Urine pregnancy test positive[3] (n/N %)		0*
UTI by urinalysis positive[4] (n/N %)		0*
BV by modified Amsel criteria positive[5] (n/N %)		0*
Nugent 4-10 on Gram stain (n/N %)		
Nugent 7-10 on Gram stain (n/N %)		
Nugent 4-6 on Gram stain (n/N %)		
TV on wet mount (n/N %)		0*
TV by InPouch culture (n/N %)		
Candida on wet mount (n/N %)		0*
Syphilis serology positive[6] (n/N %)		0*
HSV-2 serology positive[7] (n/N %)		
CT PCR positive (n/N %)		
NG PCR positive (n/N %)		
Positive for either CT or NG by PCR (n/N %)		

* These were exclusion criteria and can therefore not be present in the randomised population.

1. Percentages can add to more than 100% because participants were allowed to give multiple answers.
2. Kehua HIV Rapid Test (Kehua Bioengineering Co., Shanghai, China) with Determine HIV Rapid Test (Abbott Laboratories, Tokyo, Japan) for confirmation of positive results and the Unigold HIV Rapid Test (Trinity Biotech Plc, Bray, Ireland) as tie-breaker.
3. Nova hCG urine pregnancy test (Nova, Waltham, USA).
4. Nova urinalysis dipstick test (Nova, Waltham, USA), considered UTI-positive if nitrites were positive or at least 1+ leucocytes (in

accordance with Rwandan guidelines).

5. The modified Amsel score was positive if 2 or 3 of the following criteria were positive: 1) >20% clue cells on wet mount 2) a positive KOH (amine) test; 3) vaginal pH > 4.5.
6. Spinreact *T. pallidum* Haemagglutination test with confirmation of active infection by Spinreact Rapid Plasma Reagin test (Spinreact, Girona, Spain).
7. Kalon HSV-2 test (Kalon, Guildford, United Kingdom), with optical density (OD) > 1.1 of the mean OD of the cut-off defined as positive) and < 0.9 of the OD of the cut-off defined as negative.

4.2 Baseline characteristics per randomisation group (collected at Screening)

For all participants that were enrolled, the same baseline characteristics at Screening will be shown in the following table. The differences between all four study arms will be compared by eyeballing; no formal statistical tests will be performed.

The laboratory results of the enrolled and randomized participants (during the Screening visit) will be reported. Some of the tests were only performed in a subset of participants (for instance, participants were ineligible due to being HIV-positive or pregnant, and were therefore not further tested for STIs such as CT, NG, HSV-2, etc). The laboratory results per enrolled arm will also be reported.

Sociodemographics	**Group 1 N=17**	**Group 2 N=17**	**Group 3 N=17**	**Group 4 N=17**
Kigali neighbourhood (n % per category)[1]				
Age (median, IQR)				
Age, categorised (n % per category) - 18 to 25 - 26 to 35 - 36 to 45				
Marital status (n % per category)				
Education level (n % per category)				
Sexual history and behaviour	**Group 1**	**Group 2**	**Group 3**	**Group 4**
Number of sex partners in lifetime (median, IQR)				
Number of sex partners last 12 months (median, IQR)				
Number of sex partners last month (median, IQR)				
Currently has a main sex partner (n %)				
Lives together with main sex partner (n %)				
Main sex partner circumcised (n % per category) - Yes - Does not know - No				
Main sex partner has other sex partners (n % per category) - Yes - Does not know - No				
Vaginal sex frequency last 2 weeks (median, IQR)				
Any condom use in past two weeks (n %)				
Condom use during last sex act (n %)				
Exchanged sex for money/goods in past month (n %)				
Reproductive and contraceptive history	**Group 1**	**Group 2**	**Group 3**	**Group 4**
Pregnancies in lifetime (median, IQR)				
Deliveries in lifetime (median, IQR)				
Miscarriages in lifetime (median, IQR)				

Currently breastfeeding (n %)				
Ever severe infection with fever during pregnancy/childbirth (n %)				
Ever baby with severe infection with fever within days after birth (n %)				
Currently using a modern method of contraception other than condoms (n % per category): - Pills - Injections - Implant - Copper IUD - Sterilised - Other (described in footnote)				
Medical history	**Group 1 N=17**	**Group 2 N=17**	**Group 3 N=17**	**Group 4 N=17**
Any significant medical and/or surgical history or current condition (n %; describe in footnote)				
Ever had surgery (n %; describe in footnote)				
Any chronic diseases (n %; describe in footnote)				
Ever treated for BV (n %)				
Ever treated for vaginal candidiasis (n %)				
Treated for STI in the past 12 months (n %)				
Current medications and symptoms	**Group 1**	**Group 2**	**Group 3**	**Group 4**
Any current medication use (n %; describe in footnote)				
Any urogenital symptoms now or in last 2 weeks, patient-reported (n%)				
If yes, what symptoms[2] (n % per category): – Burning when passing urine – Frequent urination or urgent need to urinate – Blood in urine – Genital burning – Genital itching – Pain during sex – Lower abdominal pain – Unusual vaginal discharge – Sores in the genital and/or anal area (including buttocks) – Other (additional categories to be added if needed)				
If yes, already received treatment (n %)				
If yes, some or all of the symptoms currently ongoing (n %)				
Medical procedures by study physician	**Group 1**	**Group 2**	**Group 3**	**Group 4**
Any abnormal pelvic exam findings (n % per category)				
Any abnormal bimanual exam findings (n % per category)				
Clinical diagnosis (n % per category)				
Treatment given (n % per category)				
Laboratory results	**Group 1**	**Group 2**	**Group 3**	**Group 4**
Nugent 4-10 on Gram stain (n %)				
Nugent 7-10 on Gram stain (n %)				
Nugent 4-6 on Gram stain (n %)				
TV by InPouch culture (n %)				

Candida on wet mount (n %)				
HSV-2 serology positive[3] (n %)				
CT PCR positive (n %)				
NG PCR positive (n %)				
Positive for either CT or NG by PCR (n %)				

1. If missing values, a footnote detailing the number of missing values for a particular variable will be added to the table.
2. Percentages can add to more than 100% because participants were allowed to give multiple answers.
3. Kalon HSV-2 test (Kalon, Guildford, United Kingdom), with optical density (OD) > 1.1 of the mean OD of the cut-off defined as positive) and < 0.9 of the OD of the cut-off defined as negative.

4.3 Other baseline characteristics by randomisation group (collected at Enrolment)

Vaginal practices	**Group 1 N=17**	**Group 2 N=17**	**Group 3 N=17**	**Group 4 N=17**	**Total N=68**
Weekly frequency of washing body (median, IQR)					
Ever washing the genitalia, and if yes, location of washing (n % per category) [1] - Yes, outside only - Yes, both inside and outside - Yes, inside only - No					
Weekly frequency of cleaning inside the vagina (median, IQR)					
Timing of cleaning inside the vagina (n % per category)[2] - Before sex - After sex - While bathing - During/after menses - Other (described in a footnote)					
Vaginal practices in the past 12 months, other than vaginal practices to manage menses (n % per category)[2] - Water only using fingers - Water and soap using fingers - Water with paper/cloth/cotton wool - Water and soap with paper/cloth/cotton wool - Paper, cloth, cotton, wool without liquid - Traditional herbs, stones, or powder - Western vaginal medicine - Other (described in a footnote) - No vaginal practices					
Practices to manage menstrual blood or spotting in the past 12 months (n % per category)[2] - Tissue, toilet paper, paper, cloth or cotton wool put inside the vagina - Tissue, toilet paper, paper, cloth or cotton wool placed in underwear - Sanitary pad - Tampon - Water, without soap, inside the vagina - Water, with soap, inside the vagina					

- Other (describe in a footnote) - Nothing was used - Did not have menses in past 12 months					
Used anything inside the vagina other than plain water in the last two weeks (n/N %)					
If yes, what was used (n/N % per category)[1] - Soap - Paper/cloth/cotton wool - Tampon - Traditional herbs, stones or powders - Western vaginal medicine - Other (described in a footnote)					
Menses	**Group 1 N=17**	**Group 2 N=17**	**Group 3 N=17**	**Group 4 N=17**	**Total N=68**
Regular menstrual cycle (n % per category) - Yes - No, irregular - Amenorrhea due to hormonal contraception - Amenorrhea due to other reason (described in a footnote)					
Duration of bleeding(median, IQR)					
Amount of bleeding (n % per category) - Light - Moderate - Heavy					
History of painful periods (n %)					
Self-reported vaginal sex during menses in the past 12 months (n %)					
Probiotics	**Group 1**	**Group 2**	**Group 3**	**Group 4**	**Total**
Ever heard of 'probiotics' before this study (n %)					
Ever bought probiotic pills for oral use in a store or on a market (n %)					
If yes, used any probiotic pills in past two weeks (n %)					
Frequency of eating yoghurt (n %) - Never - Less than once per week - More than once per week					

1. If missing values, a footnote detailing the number of missing values for a particular variable will be added to the table.
2. Percentages can add to more than 100% because participants were allowed to give multiple answers.

5. Associations between laboratory results and signs/symptoms

All women at all study visits were tested for BV by Amsel criteria, BV by Nugent scoring, TV by wet mount, TV by InPouch culture and vaginal candidiasis by wet mount. For the analyses described in this section, all laboratory results at all study visits will be used.

Modified Amsel criteria (n %)	Nugent score (n %)			Total	Fisher's exact p[1]
	0-3	4-6	7-10		
Positive					
Negative					
Total					
Vaginal pH (n %)					
≥ 4.5					
< 4.5					
Total					
>20% Clue cells on wet mount (n %)					
Positive					
Negative					
Total					
KOH/amine test (n %)					
Positive					
Negative					
Total					
Candida on wet mount (n %)					
Positive					
Negative					
Total					
TV on wet mount (n %)					
Positive					
Negative					
Total					
TV InPouch culture (n %)					
Positive					
Negative					
Total					
Unusual vaginal discharge self-reported (n%)					
Present					
Absent					
Total					
Unusual vaginal discharge clinician observed (n %)					
Present					
Absent					
Total					
Abnormal genital odour clinician observed (n %)					
Present					
Absent					
Total					

1. When p<0.05, Nugent 4-6 and Nugent 7-10 will each be compared to Nugent 0-3.

The sensitivity, specificity, positive predictive value and negative predictive value of the modified Amsel criteria compared to the Nugent score (with 7-10 defined as positive, 0-3 as negative) will also be calculated.

TV on wet mount (n %)	TV culture		Total	OR (95% CI)[1]	Fisher's exact p
	Positive (n%)	Negative (n%)			
Positive					
Negative					
Total					
Unusual vaginal discharge self-reported (n %)					
Present					
Absent					
Total					
Unusual vaginal discharge clinician observed (n %)					
Present					
Absent					
Total					
Vaginal itching self-reported (n %)					
Present					
Absent					
Total					

The sensitivity, specificity, positive predictive value and negative predictive value of TV on wet mount compared to TV culture will also be calculated.

Unusual vaginal discharge self-reported (n %)	Candida on wet mount		Total	OR (95% CI)[1]	Fisher's exact p
	Positive (n %)	Negative (n %)			
Present					
Absent					
Total					
Unusual vaginal discharge clinician observed (n %)					
Present					
Absent					
Total					
Vaginal itching self-reported (n %)					
Present					
Absent					
Total					

6. Analyses of follow-up data: primary efficacy outcomes

6.1 Intent-to-treat incidence rates of primary efficacy outcomes

All women at all study visits were tested for BV by Amsel criteria, BV by Nugent scoring, TV by wet mount, TV by InPouch culture and vaginal candidiasis by wet mount.

During follow-up, antibiotic and antifungal treatments were given as follows:
- BV was only treated when a woman had BV by Nugent score or modified Amsel criteria but also reported symptoms typical for BV (unusual vaginal discharge, fishy smell) or these were observed by the study physician during a pelvic examination. A diagnosis could almost always be made on the day of the study visit.
- TV was always treated when identified on wet mount or by TV culture. The wet mount results were always available on the day of the study visit, but the InPouch culture results only 5 days later. If a woman had tested negative by wet mount but positive by InPouch culture, she was asked to return to the RU clinic for treatment.
- Vaginal candidiasis was only treated when Candida was present on the wet mount and the participant reported symptoms that are typical for vaginal candidiasis (itching, curd-like discharge) or these were observed by the study physician during a pelvic examination. A diagnosis could almost always be made on the day of the study visit.

We will describe any antibiotic or antifungal use between randomisation and M2 and between M2 and M6 by indication, drug type and drug dosing schedule.

Each positive case during follow-up will be assessed separately to differentiate between incident infections and persistent infections. We will assess whether treatments were dispensed appropriately and whether participants adhered to the treatments.

The incidence of each primary outcome will be defined as follows:
- Incident BV by modified Amsel criteria: positive by modified Amsel criteria after having been negative by modified Amsel criteria at the previous visit or after a case of symptomatic BV (diagnosed by modified Amsel criteria or by Nugent score) was adequately treated.
- Incident BV by Nugent score 4-10: a Nugent score of 4-10 after a Nugent score of 0-3 at the previous visit or after a case of symptomatic BV (diagnosed by modified Amsel criteria or by Nugent score) was adequately treated.
- Incident BV by Nugent score 7-10: a Nugent score of 7-10 after a Nugent score of 0-6 at the previous visit or after a case of symptomatic BV (diagnosed by modified Amsel criteria or by Nugent score) was adequately treated.
- Incident TV by InPouch culture: a positive TV culture after a negative TV culture at the previous visit, or after a case of TV (diagnosed on wet mount or by InPouch culture) was adequately treated.
- Symptomatic vaginal candidiasis: Candida on wet mount plus the presence of symptoms (which the study physician deemed eligible for treatment) after a negative wet mount at the previous visit, a positive wet mount at the previous visit but in the absence of treatment (i.e.no treatment was given), or after adequate treatment for symptomatic vaginal candidiasis.
- In all definitions, one woman can have multiple incident infections, and all of these will be included in incidence rate (IR) and incidence rate ratio (IRR) calculations.

Incidence rates: IRs will be calculated as the number of incident cases per person-days (PD) of follow-up with 95% confidence intervals using the IR function in STATA. IRs will be calculated for each primary endpoint in each of the 4 randomisation groups and for the entire study population. IRs will be calculated for three time periods: between the Enrolment visit and the M2 visit (during product use), between the M2

visit and M6 visit (after cessation of product use) and for the entire study period after randomisation. They will be reported in the following table (one table for each time period):

	Group 1		Group 2		Group 3		Group 4		Total	
	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]
Incident BV by modified Amsel criteria										
Incident Nugent 4-10										
Incident Nugent 7-10										
Incident TV by culture										
Incident symptomatic vaginal candidiasis										

1. Number of women (n) who developed at least one incident infection during the specified time period as a proportion of the women who had at least one follow-up visit in that time period (N).
2. Incident rate (IR): Number of incidence cases per person-day of follow-up, with 95% confidence intervals (CI).

The median Nugent score and vaginal pH category at each study visit for each study group and the total study population will be visualised using line graphs as follows:
- Nugent score as a count between 0 and 10 (median for the group)
- Vaginal pH as the ordinal categories on the pH strip (median category for the group)

6.2 Intent-to-treat incidence rate ratios comparing product use groups to the control group

IRRs with 95% confidence intervals will be calculated comparing each product use group (groups 2-4) to the control group (group 1) using the IR function in STATA. In addition, IRRs with 95% confidence intervals will be calculated comparing each vaginal probiotic group (group 3 and 4) to the metronidazole group (group 2). As before, these IRRs will be calculated for each primary efficacy outcome and for each of the three time periods: between the Enrolment visit and the M2 visit (during product use), between the M2 visit and M6 visit (after cessation of product use) and for the entire study period after randomisation. They will be reported in the following table (one table for each time period):

	Group 2 vs 1 IRR (95% CI)	Group 3 vs 1 IRR (95% CI)	Group 4 vs 1 IRR (95% CI)	Group 3 vs 2 IRR (95% CI)	Group 4 vs 2 IRR (95% CI)
Incident BV by modified Amsel criteria					
Incident Nugent 4-10					
Incident Nugent 7-10					
Incident TV by culture					
Incident symptomatic vaginal candidiasis					

6.3 Per protocol incidence rate ratios comparing product use groups to the control group

A few significant protocol deviations occurred that could have an impact on the primary analysis results:

1. CT/NG PCR testing was not available at the RU clinic and was outsourced to the National Reference Laboratory in Kigali, Rwanda. Test results were usually available only weeks after sample collection. We realised that this might be an issue early on and increased the sample size of each randomisation group to 17 instead of 15 to allow for exclusion of women who were randomised while having an ongoing CT or NG infection. Unfortunately, the number of women with a CT or NG infection in each randomisation group turned out to be higher than expected: 8/17 in Group 1, 7/17 in Group 2, 4/17 in Group 3, and 7/17 in Group 4 (Fisher's exact p = 0.53).

2. Discrepancy between modified Amsel criteria and Nugent score at the Enrolment visit: Women could be enrolled if they had negative modified Amsel criteria after having completed the 7-day metronidazole treatment. Nugent scores became available later. In a handful of cases, women who were randomised because they no longer had BV by modified Amsel criteria turned out the have a Nugent score of 7-10 at the Enrolment visit.

3. Another special case is the use of antibiotics and/or antifungals during follow-up. If these were prescribed by the study physician due to symptomatic BV, TV, or symptomatic vaginal candidiasis, their use is causally related to the primary endpoints. These women will therefore be included in the per protocol analyses. However, if women used antibiotics and/or antifungals extensively during follow-up, including for reasons not related to the primary endpoints, they may be excluded in the per protocol analyses. This will be determined on a case-by-case basis by the Chief Investigator and each case will be described.

The analyses described in 6.1 and 6.2 will be repeated after exclusion of the participants described in 1, 2, and 3 above. If additional significant protocol deviations are discovered during data analyses (which is unlikely), additional exclusions might be made. These will also be determined on a case-by-case basis and each case will be described.

6.4 Adjustment for confounding: generalised estimating equation (GEE) models

Other behavioural and clinical characteristics might also influence the primary efficacy outcomes. Some of these variables were collected at baseline but also at the M2 and M6 visits.

We will assess any behavioural changes over time in the total study population as follows:

Sexual behaviour	**Baseline**	**M2**	**M6**	**p-value[2]**
Number of sex partners in last month at baseline or per month during follow-up period (median, IQR)				
Currently has a main sex partner (n %)[1]				
Obtained a new main partner during period (n %)	NA			
Lives together with main sex partner (n %)				
Main sex partner circumcised (n % per category) - Yes - Does not know - No				

Main sex partner has other sex partners (n % per category) - Yes - Does not know - No				
Vaginal sex frequency last 2 weeks (median, IQR)				
Any condom use in past two weeks (n %)				
Exchanged sex for money/goods in past month (n %)				
Contraceptive practices, vaginal practices, menses	**Baseline**	**M2**	**M6**	**p-value[2]**
Currently using a modern method of contraception other than condoms (n % per category): - Pills - Injections - Implant - Copper IUD - Sterilised - Other				
Weekly frequency of vaginal practices since the last visit (includes water only using fingers, water and soap using fingers, paper/cloth/cotton wool with or without liquid, and traditional herbs, stones, or powders)				
Amenorrhea due to hormonal contraception (n %) Amenorrhea due to other reason (n %) If no amenorrhea, duration of bleeding (median, IQR) If no amenorrhea, any vaginal sex during menses (n %)				

1. If missing values, a footnote detailing the number of missing values for a particular variable will be added to the table.
2. McNemar's test for binary or categorical variables and Wilcoxon test for medians.

We will use generalised estimating equation (GEE) models with Nugent 4-10 (compared to Nugent 0-3) at any study visit after randomisation as the outcome and randomisation group as the main predictor (and indicator variable with Group 1 as the reference group) to adjust for the potential effects of these characteristics. A similar model will be run with Nugent 7-10 (compared to Nugent 0-6) as the outcome.

Assessment for effect modification:
We will assess CT and NG infection at Screening (combined variable: one or both infections diagnosed by PCR at Screening and not treated until later on in the study) for effect modification. If statistically significant effect modification is present, the relevant women will be excluded from the GEE models. If no effect modification is present, the presence of CT and/or NG infection will be adjusted for (see below). We expect the number of women with a Nugent score of 7-10 at the Enrolment visit, or who used antibiotics and/or antifungals extensively, to be too small to assess for effect modification or confounding.

Adjustment for confounding:
Since the sample size of the Rwanda VMB Study is small, we do not have sufficient statistical power to adjust for many confounding variables. We will focus on two variables: CT/NG infection (see above) and study product adherence (see below). However, if we find that any of the variables in the tables in sections 4.2 and 4.3 and in the table are imbalanced between the randomisation groups, or if any behaviours changed significantly over time (see table above), one or two additional variables may be added to the models.

Adherence: Adherence to the study interventions was assessed at the D7, M1 and M2 visits by structured interviewer-administered questionnaire, review of a diary card that the participant completed in between

study visits, review of returned used packaging and unused product, and by asking the participant to complete a self-rating adherence scale. These different sources of adherence data will be triangulated to arrive at an overall level of adherence (between 0-100%) for each participant between the Enrolment and D7 visits, D7 and M1 visits, M1 and M2 visits, and the Enrolment and M2 visits. Women were allowed to cease product use during menstrual bleeding and these days will therefore not be taken into account. Adherence will be plotted in a line graph for randomisation groups 2, 3 and 4 with study visits (D7, M1, and M2) on the horizontal axis and the adherence percentage on the vertical axis. In the GEE models, only one overall adherence measure (% between the Enrolment and M2 visits) will be used, with the adherence level for women in Group 1 (who did not use any study product) set at 100%.

7. Analysis of follow-up data: primary safety outcomes

7.1 Description of serious adverse events

PID	Event / Diagnosis	Description	Outcome	Related to product?	Date of onset
027	Typhoid fever	Participant was generally unwell, had a fever, epistaxis and a headache.	Participant was hospitalized for 7 days but recovered completely.	No	19 Nov 2015
145	Malaria during pregnancy	Participant became generally unwell, had a headache, nausea, and fever. She was 8 weeks pregnant at the time.	Participant was hospitalized for 2 days but recovered completely.	No	22 Jan 2015

7.2 Description of social harms

PID	Description	Severity / Outcome	Date of onset
036	The participant was beaten by partner due to participation in self-sampling group. RU staff visited the participant at home.	Medium. The participant stopped frequent self-sampling but continued study participation.	29 Jun 2015
093	The participant was verbally “harassed” by her partner and sister because of her participation in the study.	Mild. The participant decided to withdraw from the study.	26 Jul 15

7.3 Adverse events: structurally assessed at each study visit

Some safety outcomes were structurally assessed at each study visit as part of the face-to-face interview and pelvic and bimanual examination. Laboratory test results will not be included because they are covered under the primary and secondary efficacy outcomes. The safety outcomes will be reported as the number of women who experienced the event at least once during the time period by study group. Differences between the randomisation groups will be assessed by Fisher’s exact test. If the p-value is < 0.05, pairwise comparisons (also using Fisher’s exact test) between study groups will be conducted. If an AE occurred more than once in individual women, the analysis will be repeated for the number of events (as opposed to number of women) per study group. All analyses will be done for three time periods: between the Enrolment and M2 visits (during product use), between the M2 and M6 visits (after cessation of product use), and for the overall follow-up period between the Enrolment and M6 visits.

Number of women Between Enrolment – M2	Group 1 N=17	Group 2 N=17	Group 3 N=17	Group 4 N=17	p-value[1]
Any patient-reported urogenital symptoms (n %)					
If yes, symptoms reported (n % for each)[2]: – Burning when passing urine – Frequent urination or urgent need to urinate – Blood in urine – Genital burning – Genital itching – Pain during sex – Lower abdominal pain – Unusual vaginal discharge – Sores in the genital and/or anal area (including buttocks) – Other (describe in footnote)					
Any abnormal findings during pelvic exam clinician-observed (n %)					
If yes, findings observed (n % for each)[2]: – Abnormal (genital) odour – Enlarged/tender inguinal lymph node – Condylomata (any location genitalia) – Ulcers/blisters suggestive of STI on vulva – Any other lesion on vulvar epithelium – Abnormal vaginal discharge – Vaginal mass (polyp, myoma, etc.) – Ulcers/blisters suggestive of STI in vagin – Any other lesion on vaginal epithelium – Abnormal cervical discharge/pus – Any lesion on cervical epithelium – Other (describe in footnote)					
Any abnormal findings during bimanual exam clinician-observed (n %)					
If yes, findings observed (n % for each)[2]: – Uterine mass – Adnexal mass – Uterine tenderness – Adnexal tenderness – Cervical motion tenderness – Other (describe in footnote)					
Diagnoses based on the above (but not including the efficacy endpoints) (n % for each): - Pelvic inflammatory disease (PID) - Mucopurulent cervicitis - Genital ulcer disease (GUD) - Genital warts/condylomata					

1. Fisher's exact test
2. Percentages can add to more than 100% because participants were allowed to give multiple answers or could have multiple conditions.

7.4 Adverse events: not structurally assessed

In addition to the safety outcomes that were structurally assessed, the study physician also asked the following questions at each study visit: Since your last study visit, did you experience any other new physical complaints or were you diagnosed with any other new conditions? Since your last study visit, have you taken any new medications? The answers given to those questions, and any other spontaneously reported adverse events, will be coded using MedDRA (version 19.0, March 2016) and will be reported in the same manner as described above for the structurally assessed adverse events.

8. Analyses of follow-up data: secondary efficacy outcomes

8.1 Incidence and incidence rate ratios of secondary efficacy outcomes

It has been hypothesised that a lactobacilli-dominated VMB would protect women from HIV, other STIs, and UTIs, and would facilitate pregnancy. While the sample size of this study was too small to adequately test these hypotheses, we did include incidence of HIV, STIs (other than TV), UTIs, and pregnancy as secondary outcomes. Women were tested for HIV and STIs (HSV-2, syphilis, CT, and NG) at the Screening and M6 visits, and for UTI and pregnancy at the Screening, Enrolment and M6 visits. In all cases, additional testing may have been done at the D7, M1 and M2 visits when clinically indicated. Incidence of these secondary outcomes will be determined between baseline (using endpoint data from the Enrolment visit if available and otherwise from the Screening visit) and M6.

Incident cases will be defined as follows:

- HIV: a positive result for the HIV serology algorithm at M6. Women who were HIV-positive at Screening were excluded from the study; therefore, all positive results at M6 will be treated as incident cases. Time at risk will start at the Enrolment visit.
- HSV-2: a positive result for HSV-2 serology at M6. Only those participants who tested negative for HSV-2 at Screening were retested at M6; therefore, all positive cases at M6 will be treated as incident cases. Time at risk will start at the Enrolment visit.
- Syphilis: a positive result for the syphilis serology algorithm at M6. All participants positive for syphilis at Screening were adequately treated prior to randomisation; therefore, all positive cases at M6 will be treated as incident cases. Time at risk will start at the Enrolment visit.
- NG: a positive result for NG PCR at M6 after a negative NG PCR result at Screening (time at risk will start at the Enrolment visit) or after having been successfully treated for NG during follow-up (time at risk will start on the treatment completion date).
- CT: a positive result for CT PCR at M6 after a negative CT PCR result at Screening (time at risk will start at the Enrolment visit) or after having been successfully treated for CT during follow-up (time at risk will start on the treatment completion date).
- UTI: a positive urinalysis result (according to the Rwandan guidelines) at M6 or earlier during follow-up if tested due to symptoms after a negative urinalysis dipstick result at an earlier visit or after adequate UTI treatment. Time at risk will always start at the Enrolment visit because all UTI's at the Screening and Enrolment visits were treated prior to randomisation.
- Pregnancy: a positive urine pregnancy test at M6 (or earlier during follow-up). As being pregnant at Screening or Enrolment was a reason for exclusion in this study, all positive results will be treated as incident cases.

Incidence rates: IRs will be calculated as the number of incident cases per person-days (PD) of follow-up with 95% confidence intervals using the IR function in STATA. IRs will be calculated for each primary endpoint in each of the 4 randomisation groups and for the entire study population. In all cases, IRs will be calculated for the entire follow-up period as described above.

	Group 1		Group 2		Group 3		Group 4		Total	
	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]	n/N[1]	Cases/PD (IR, 95% CI)[2]
HIV										
HSV-2										
Syphilis										
NG										
CT										
UTI										
Pregnancy										

1. Number of women (n) who developed at least one incident infection during the specified time period as a proportion of the women who completed all follow-up visits in that time period (N).
2. Incident rate (IR): Number of incidence cases per person-days of follow-up, with 95% confidence intervals (CI).

Incidence rate ratios (IRRs) with 95% confidence intervals will be calculated comparing each product use group (groups 2-4) to the control group (group 1) using the IR function in STATA. In addition, IRRs with 95% confidence intervals will be calculated comparing each vaginal probiotic group (group 3 and 4) to the metronidazole group (group 2). As before, these IRRs will be calculated for each primary efficacy outcome for the entire follow-up period.

	Group 2 vs 1 IRR (95% CI)	Group 3 vs 1 IRR (95% CI)	Group 4 vs 1 IRR (95% CI)	Group 3 vs 2 IRR (95% CI)	Group 4 vs 2 IRR (95% CI)
HIV					
HSV-2					
Syphilis					
NG					
CT					
UTI					
Pregnancy					

The number of incident infections will likely be too small to allow for per-protocol analyses and GEE models as described for the primary efficacy outcomes in 6.3 and 6.4.