

Sponsor Name: BioNTech SE.

Protocol Title: A Phase I randomized, double-blind, placebo-controlled, safety, tolerability, and pharmacokinetic trial of BNT331 administered in single ascending doses in healthy women and multiple ascending doses in women diagnosed with bacterial vaginosis

Protocol Version Date: 13 AUG 2024

Compound Name: BNT331 vaginal inserts

Study Phase: Ia/b

Regulatory Agency Identifier(s):	IND Number: 163436. EUDRACT Number: N/A
----------------------------------	--

Registry Identifier(s):	Registry Name	Registry ID
	clinicaltrials.gov	NCT06469164

SAP Version History

SAP Version	Approval Date	Based on Clinical Trial Protocol (Version, Approval Date)	Brief Description of Change(s)/Revision(s) and Rationale
1.0	18 FEB 2025	version 2.0, 13 Aug 2024	Original document creation.
1.1	24 FEB 2025	Version 2.0, 13 Aug 2024	To correct version error in header from 4.0 to 1.1
2.0	25 SEP 2025	Version 2.0, 13 Aug 2024	<p>Updates / corrections specified in respective Note to File and in Data Presentation Plan Version 1.0 prior to unblinding for interim analysis:</p> <ul style="list-style-type: none"> Section 3.1: Update of mITT set definition Section 4.1.2: Inclusion of analysis time windows used for external data file handling Section 4.2: Clarification that reasons for exclusion from analysis sets will only be listed Section 4.3.1: Clarification that Amsel's criteria will be summarized at baseline Section 4.3.3: Specification of considered time periods for summary of BV history Section 4.3.5.1: Clarification that rescue medications only apply to Part B Section 4.5.2: Correction of analysis set used for ADA response analysis – Biomarker set Section 4.8.1: Clarification that exposure data will only be listed for Part A <p>Further updates / clarifications:</p> <ul style="list-style-type: none"> Section 3.1: Clarification of Biomarker set definition Section 4.1: Clarification that display of a maximum number of significant digits applies to PK Section 4.1.1: Clarification added that procedures starting on first day of study treatment will be considered post-baseline Section 4.1.1 and 4.8.1: Clarification of sources of exposure data for Part B Section 4.1.1: Addition to specify disjoint categories for subgroup "Method of contraception" Section 4.1.3 and 4.5.3: Clarification added that binary efficacy endpoints to be set to non-response after start of rescue medication refer to Clinical cure, Nugent Score cure, as well as Responder outcome. Sections 4.6.3 and 4.6.4: Addition of handling rules for valid measurements with no bacteria of a considered taxus detected

			<ul style="list-style-type: none"> • Section 4.6.3: Clarification that change in Nugent score will be summarized for the ITT set for Part A • Minor corrections of typos/formats
--	--	--	--

TABLE OF CONTENTS

SAP VERSION HISTORY	2
TABLE OF CONTENTS	4
LIST OF TABLES	6
LIST OF FIGURES	6
LIST OF ABBREVIATIONS	7
1. INTRODUCTION	10
1.1 Study Objectives, Endpoints and Estimands	10
1.2 Study Design	13
1.2.1 Treatment Randomization and Blinding	15
1.2.2 Central Review	16
1.2.3 Data Monitoring	16
2. STATISTICAL HYPOTHESES AND MULTIPLICITY ADJUSTMENT	16
3. ANALYSIS SETS	16
3.1 Definition of Analysis Sets	16
3.2 Violations and Deviations	17
4. STATISTICAL ANALYSIS	17
4.1 General Considerations	17
4.1.1 Definitions, Derivations and Classifications	19
4.1.2 Windowing, Unscheduled Visits and Repeat Test Results	22
4.1.3 Dropouts, Missing/Incomplete Data, and Outliers	23
4.2 Subject Disposition	23
4.3 Subject Characteristics	24
4.3.1 Demographic and Baseline Characteristics	24
4.3.2 Baseline Assessments	24
4.3.3 Medical and Other History	24
4.3.4 Disease-Specific Characteristics	25
4.3.5 Medications, Therapies and Procedures	26
4.3.5.1 Prior and Concomitant Medications	26
4.3.5.2 Prior and Concomitant Procedures	26
4.4 Primary Endpoint Analyses	27
4.5 Secondary Endpoint Analyses	27
4.5.1 Secondary Endpoint 1 Analysis – Pharmacokinetic Analysis (Serum)	27
4.5.1.1 Pharmacokinetics – Analytical Approach for Serum	27
4.5.2 Secondary Endpoint 2 Analysis - ADA Response Analysis (Serum)	29

4.5.3	Secondary Endpoint 3 Analysis – Clinical and Nugent Score Cure and Responder Outcome Analyses (Part B Only)	31
4.6	Exploratory Endpoint Analyses	31
4.6.1	Exploratory Endpoint 1 Analysis – Pharmacokinetic Analysis (Vaginal Fluid)	31
4.6.2	Exploratory Endpoint 2 Analysis – ADA Response Analysis (Vaginal Fluid)	32
4.6.3	Exploratory Endpoint 3 Analysis - Change in the Overall Vaginal Microbiome	32
4.6.4	Exploratory Endpoint 4 Analysis - Change in the Levels of <i>Gardnerella</i> and <i>Lactobacillus spp.</i>	34
4.6.5	Exploratory Endpoint 5 Analysis – Long-term Treatment Outcome Analysis (Part B Only)	35
4.6.6	Exploratory Endpoint 6 Analysis – Long-term Psychosocial Outcome Analysis (Part B Only)	35
4.7	Other Analyses	36
4.8	Safety Analyses	36
4.8.1	Extent of Exposure	36
4.8.2	Adverse Events	37
4.8.3	Laboratory Data	38
4.8.4	Vital Signs	39
4.8.5	Other Safety Assessments	39
4.8.5.1	Physical Examination	39
4.8.5.2	Pelvic Examination	39
4.8.5.3	Trichomonas rapid test (Part B only)	39
4.8.5.4	Pregnancy Test	40
4.8.5.5	Local Tolerability	40
4.9	Subgroup Analyses of Interest	40
5.	INTERIM ANALYSIS AND ANALYSIS SEQUENCE	40
6.	SAMPLE SIZE DETERMINATION	41
7.	CHANGES FROM AND ADDITIONS TO METHODS PLANNED IN THE CLINICAL TRIAL PROTOCOL	41
8.	REFERENCES	41
9.	SUPPORTING DOCUMENTATION	41
	Appendix 1 Schedule of Procedures/Assessments	42
	Appendix 2 Date Imputation Rules	50
	Appendix 3 Health Reported Outcome Surveys and Questionnaires	52

LIST OF TABLES

Table 1:	Objectives and Endpoints	11
Table 2:	Dose Cohorts (Part A).....	13
Table 3:	Dose Cohorts (Part B).....	14
Table 4:	Schedule of Activities – Part A	43
Table 5:	Schedule of Activities – Part B	46

LIST OF FIGURES

Figure 1:	Study Schema.....	15
-----------	-------------------	----

LIST OF ABBREVIATIONS

Abbreviation	Description
ADA	anti-drug antibody
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine transaminase
AST	aspartate transaminase
ATC	Anatomic Therapeutic Chemical
AUC	area under the concentration-time curve
AUC _{0-inf}	area under the concentration-time curve from pre-dose to infinity
AUC _{0-t}	area under the concentration-time curve from pre-dose to time of last quantifiable concentration
%AUC _{ex}	percentage of AUC _{0-inf} by extrapolation
BLQ	below the limit of quantitation
BMI	body mass index
BUN	blood urea nitrogen
BV	bacterial vaginosis
CHG	change from baseline
CL/F	clearance
C _{max}	maximum concentration
CRO	clinical research organization
CSFQ-F	changes in sexual functioning questionnaire –female version
CSR	clinical study report
CTMS	clinical trial management system
CTP	clinical trial protocol
CV%	coefficient of variation percentage
C. trachomatis	Chlamydia trachomatis
DL	dose level
eCRF	electronic Case Report Form

Abbreviation	Description
FDA	Food and Drug Administration
FIH	first-in-human
h	hour(s)
IA	interim analysis
IC	informed consent
ICH	International Council for Harmonisation
IMP	investigational medicinal product
IPD	important protocol deviation
IRC	Internal Review Committee
ITT	intent-to-treat
KOH	potassium hydroxide
λ_z	terminal rate constant
λ_z lower	starting point of time interval to determine λ_z
λ_z upper	ending point of time interval to determine λ_z
LLOQ	lower limit of quantitation
MAD	Multiple ascending dose
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCS	mental component score
MCV	mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified Intent-to-Treat
MRT_{inf}	mean residence time
N	total number of subjects in the given cohort
NIPD	non-important protocol deviation
NOAEL	no observed adverse effect level
N. gonorrhoeae	Neisseria gonorrhoeae
PCHG	percent change from baseline

Abbreviation	Description
PCS	physical component score
PD	pharmacodynamics
PK	pharmacokinetics
PROM	patient-reported outcome measure
PT	preferred term
Q1	lower quartile
Q3	upper quartile
qPCR	quantitative polymerase chain reaction
RDW	red cell distribution width
SAE	serious adverse event
SAD	single ascending dose
SAP	statistical analysis plan
SD	standard deviation
SF-36	short form (36) health survey
S/N	signal to noise
SOC	system organ class
spp.	species
STI	sexually transmitted infection
$t_{1/2}$	elimination half-life
TEAE	treatment-emergent adverse event
TLF	tables, listings, figures
$t_{\max}$	time of $C_{\max}$
ToC	test of cure
V_z/F	apparent volume of distribution
WGS	whole genome sequencing
WHO-DD	World Health Organization-Drug Dictionary

1. INTRODUCTION

This is a first-in-human (FIH), two-part, randomized, double-blind, placebo-controlled, multi-center Phase I study evaluating the safety, tolerability, and pharmacokinetics (PK) of BNT331 administered in single ascending doses in healthy women (Part A) and multiple ascending doses in women diagnosed with bacterial vaginosis (BV) (Part B).

This statistical analysis plan (SAP) describes the detailed procedures for the planned statistical analyses of the safety, immunogenicity, pharmacokinetic, pharmacodynamic (PD) and efficacy data that are collected. The endpoints are defined, and the statistical methods used to analyze them are presented.

The Clinical Research Organization (CRO) Biostatistics may support the following exploratory analysis, where feasible. Alternatively additional information may be detailed in a separate document:

- Evaluating the changes in the vaginal microbiome by metagenomic sequencing of subjects following treatment with BNT331.
- Evaluating the change in the levels of *Gardnerella* and *Lactobacillus species (spp.)* by quantitative polymerase chain reaction (qPCR).

Table shells for the planned tables, listings, and figures (TLFs) are included in a separate document, entitled “Data Presentation Plan”.

All analysis will be described in the International Council for Harmonisation (ICH) E3 clinical study report (CSR). The ICH E3 CSR will be referred to in this document as “CSR” for conciseness.

This version of the SAP has been developed using clinical trial protocol (CTP) version 2.0, dated 13 AUG 2024. Any need to describe changes to the CTP, if applicable, will be described in Section 7.

1.1 Study Objectives, Endpoints and Estimands

[Table 1](#) displays the objectives and endpoints for Part A and Part B of the study.

Estimands are not planned in the initial SAP.

Table 1: Objectives and Endpoints

Objective	Endpoint
PRIMARY	
Part A and Part B	
To evaluate the safety and tolerability of BNT331 when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.	<p>For each cohort of BNT331 group and combined placebo groups:</p> <ul style="list-style-type: none"> • Occurrence of AEs with onset after the first treatment dose and until 7 days post-dose (Part A) and 120 days after the first dose (Part B). • Occurrence of SAEs with onset after the first treatment dose and until the 7 days post-dose (Part A) and 120 days after the first dose (Part B).
SECONDARY	
Part A and Part B	
To determine the systemic PK of BNT331 active substance, including quantification of systemic absorption when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.	<p>For each cohort:</p> <ul style="list-style-type: none"> • <u>Part A</u>: The serum concentrations of BNT331 active substance at pre-specified timepoints among subjects who received one single administration. • <u>Part B</u>: The serum concentrations of BNT331 active substance at pre-specified timepoints among subjects who received all scheduled administrations.
To determine ADA response in healthy women and in women diagnosed with BV.	<p>For each cohort:</p> <ul style="list-style-type: none"> • ADA prevalence and change of binding titers against BNT331 active substance in blood before study treatment (pre-dose) and at 7 days post-dose (Part A) and at 6 days after the first dose, 21 to 30 days after the first dose, and 60 days after the first dose (Part B).
Part B only	
To evaluate the proportion of subjects with diagnosed BV who meet the definitions of clinical cure, Nugent score cure, and responder outcome.	<p>For each cohort of BNT331 group and combined placebo groups:</p> <ul style="list-style-type: none"> • Occurrence of Clinical cure defined as normalization of the vaginal discharge, a negative KOH "Whiff" test, and clue cells <20% of the total epithelial cells/high power field on microscopic examination of the vaginal fluid at 6 days after the first dose and 21 to 30 days after the first dose. • Occurrence of Nugent score cure/Microbiological cure defined as score of <4 at 6 days after the first dose and 21 to 30 days after the first dose. • Responder outcome defined as an occurrence of clinical cure with a normal Nugent score (<4) 6 days after the first dose and 21 and 30 days after the first dose.
EXPLORATORY	
Part A and Part B	
To determine the local PK of BNT331 active substance when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.	<p>For each cohort:</p> <ul style="list-style-type: none"> • <u>Part A</u>: The vaginal fluid concentration of BNT331 active substance over time before (pre-dose), and after study treatment administration (4 h and 24 h) and at 7 days post-dose. • <u>Part B</u>: The vaginal fluid concentration of BNT331 active

Objective	Endpoint
	substance over time before (pre-dose), and 8 h to 24 h after the last study treatment administration.
To determine local ADA response when BNT331 is administered as SAD in healthy women and MAD in women diagnosed with BV.	For each cohort: <ul style="list-style-type: none"> Vaginal ADA presence will be detected as S/N ratio against BNT331 before study treatment (pre-dose) and at 7 days post-dose (Part A) and 21 to 30 days after the first dose (Part B).
To evaluate the changes in the vaginal microbiome of subjects following treatment with BNT331.	For each BNT331 group and combined placebo groups from 5 mg cohort and 20 mg cohort in Part A and Part B, respectively. <ul style="list-style-type: none"> Change of relative abundance of bacterial taxa by shotgun metagenomic sequencing (such as identification of BV associated bacteria including <i>Gardnerella vaginalis</i> and <i>Lactobacillus spp.</i>) from baseline (before treatment) until 7 days post-dose (Part A) or until 120 days after the first dose (Part B). Change of Nugent score.
Change in the levels of <i>Gardnerella</i> and <i>Lactobacillus spp.</i> from baseline (before treatment).	For each BNT331 group from 5 mg cohort and 20 mg cohort in Part A; for each BNT331 group and combined placebo groups from 5 mg cohort and 20 mg cohort in Part B: <ul style="list-style-type: none"> Quantitate the levels of <i>Gardnerella</i> and <i>Lactobacilli</i> from baseline (before treatment) to 7 days post-dose (Part A) and to 21 to 30 days after the first dose (Part B).
Part B only	
To evaluate the long-term treatment outcome of subjects with diagnosed BV who meet the definitions of clinical cure following treatment with BNT331.	For each cohort of BNT331 group and combined placebo groups: <ul style="list-style-type: none"> Occurrence of subjects reporting new BV symptoms and episodes until 120 days after the first dose.
To measure the long-term psychosocial outcome of subjects with diagnosed BV.	Patient-reported outcome before and after treatment: For each cohort of BNT331 group and combined placebo groups, including subjects after rescue treatment: <ul style="list-style-type: none"> Change over time in scores of SF-36 and CSFQ-F at three timepoints (baseline, 21 to 30 days after the first dose, and 120 days after the first dose).
ADA = anti-drug antibody; AE = adverse event; BV = bacterial vaginosis; CSFQ-F = Changes in Sexual Functioning Questionnaire; h (s)= hours; KOH = potassium hydroxide; PK = pharmacokinetics; MAD = multiple ascending doses; SAD = single ascending doses; SAE = serious adverse event; SF-36 = Thirty-six-item short form health survey; spp. = species.	

1.2 Study Design

This is a two-part, randomized, double-blind, placebo-controlled study to assess the safety, tolerability, PK, PD and efficacy (for Part B) of BNT331 in healthy women (Part A) and in women diagnosed with BV (Part B).

Approximately, 96 subjects who are at least 18 years of age will be enrolled in this two-part study.

Since BNT331 will be evaluated for the first time in human, the study uses a staggered, sequential approach for subject dosing. Pausing/stopping rules will be followed (see details in Section 7.1 of the CTP). The study will be monitored by an Internal Review Committee (IRC) that may pause or even permanently discontinue administration of study treatment, or even permanently terminate the study at any point in time.

Part A

In Part A, 24 healthy women, who are non-pregnant, non-breastfeeding and pre-menopausal will be randomized at a ratio of 3:1 to BNT331 or placebo in three single ascending dose (SAD) level cohorts. [Table 2](#) displays the dose cohorts in Part A.

Table 2: Dose Cohorts (Part A)

Cohort	Single dose (mg)	Number of subjects (N = 24)	
		BNT331	Placebo
1	2	6	2
2	5	6	2
3	20	6	2

Part A of this study will be initiated with Cohort 1, where two subjects – ensuring that at least one of the subjects is randomized to BNT331 – will be dosed at the same site and will complete at least 24 hours of follow-up post-dose. Based on tolerability and stopping/pausing rule assessments, the remaining six subjects of Cohort 1 will be dosed. Once the last subject of Cohort 1 completes Visit 3 (Follow-up visit) at seven days post-dose, all relevant data of the eight subjects will be reviewed by an IRC before enrollment for the next dose level cohort will be opened. Subsequent dose level cohorts will dose under the same rules and will follow the same dose-escalation/assessment process. Enrollment for Part B will be opened once all relevant data of Cohort 3 (Part A) subjects have been reviewed.

The planned study duration for each subject in Part A is up to 30 days of screening, one day of treatment and 12 days of follow-up, which amounts to a maximum total study duration of 43 days per subject.

Part B

In Part B, 72 non-pregnant, non-breastfeeding, pre-menopausal adult women with a diagnosis of symptomatic BV (by Amsel's criteria) will be randomized at a ratio of 2:1 to BNT331 or placebo in two multiple ascending dose (MAD) level cohorts (i.e., one dose daily for 5 days consecutively). [Table 3](#) displays the dose cohorts in Part B.

Table 3: Dose Cohorts (Part B)

Cohort	Five doses of (mg)	Number of subjects (N = 72)	
		BNT331	Placebo
4	5	24	12
5	20	24	12

Part B of this study will be initiated with Cohort 4. Once all 36 subjects have completed Visit 2 (Early Response visit), all relevant data will be reviewed by the IRC. Based on tolerability and stopping/pausing rule assessments, dosing will be opened for the next/last dose level cohort (Cohort 5, 20 mg).

The planned study duration for each subject in Part B is up to three days of screening, five days of treatment and follow-up up to Day 120 ± seven days which amounts to a maximum total study duration of 130 days per subject.

An IRC will monitor the accumulating safety data during periodic safety review meetings. A separate document, "IRC Charter" provides additional details.

An interim analysis (IA) – comprising Part A and Part B data – is planned soon after all Part B subjects have completed the test of cure (ToC) visit.

[Figure 1](#) below displays the study design.

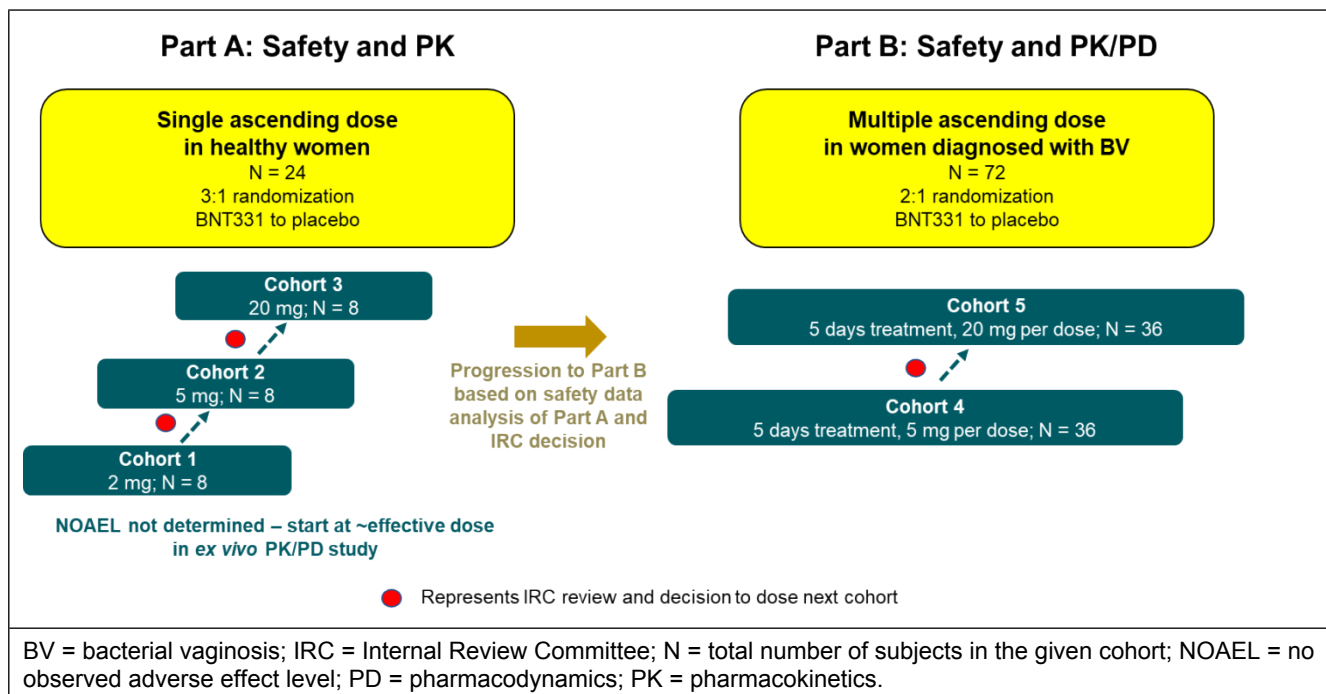


Figure 1: Study Schema

Table 4 and Table 5 in Appendix 1 display the schedule of assessments for the study.

1.2.1 Treatment Randomization and Blinding

Subjects will be randomized 3:1 to BNT331:placebo in Part A and 2:1 to BNT331:placebo in Part B via an online randomization tool.

All study site personnel, including the investigator, investigator's staff, and subjects, will remain blinded with regard to study treatment allocation until the final database lock. The IRC members will also remain blinded except in certain situations as outlined in the IRC Charter. A limited number of BioNTech and/or IQVIA personnel may be unblinded to individual subject treatment assignments for the purposes of review and analysis of PK, PD and anti-drug antibody (ADA) data.

Sponsor personnel who routinely have contact with site personnel, clinical research organization (CRO) personnel and/or are involved in the study conduct or with any other interest in this double-blind study will remain blinded with regard to the study treatment allocation until a data extract is performed after the last subject in Part B has reached ToC Visit (Visit 3) for IA.

In Part A and Part B, selected IRC members will be unblinded for the safety review.

During the study a limited number of BioNTech Bioanalytical personnel and/or IQVIA personnel may be unblinded to individual subject treatment assignments for the purposes of

review and analysis of PK, PD and ADA data. Further details are specified in the “Blind Management Plan” document.

All planned analyses identified in the SAP will be performed by IQVIA following BioNTech authorization of the full SAP, IA Database Lock, Analysis Sets and Unblinding (release of actual randomization schedule).

1.2.2 Central Review

Not applicable.

1.2.3 Data Monitoring

As outlined in the IRC Charter, an IRC will be established to provide medical and safety oversight of subject safety based on periodic in-depth review of blinded safety data by subject, cohort, and cumulatively.

The IRC will act as a decision-making body for dose-escalation, and progression from Part A to Part B and the IRC will periodically assess the overall benefit-risk of the study treatment.

2. STATISTICAL HYPOTHESES AND MULTIPLICITY ADJUSTMENT

No adjustment for multiplicity will be performed.

No formal statistical hypotheses will be tested in this study.

3. ANALYSIS SETS

3.1 Definition of Analysis Sets

Agreement and authorization of subjects included/excluded from each analysis set will be conducted prior to the planned IA (see details in Section 5).

The following analysis sets are defined for this study.

Population	Definition
Screened Set	All subjects who signed informed consent.
Safety Set	All subjects who received at least a single dose of study treatment (BNT331 or placebo).
Intent-to-treat (ITT) Set	All randomized subjects.
Modified Intent-to-Treat (mITT) Set	All randomized subjects for Part B, excluding those who subsequently demonstrate a positive test result for other concomitant vaginal or cervical infections at baseline (e.g., <i>C. trachomatis</i> , <i>N. gonorrhoeae</i>), or had any important protocol deviations which may interfere with the efficacy assessment for BV, or who have a baseline Nugent score <7 in central laboratory. Subjects who did not take any study treatment (BNT331 or placebo) are excluded from the mITT.
Pharmacokinetic (PK) Set	All subjects who receive all CTP-required doses of BNT331 and with at least one measurable PK concentration from available samples.

Population	Definition
Biomarker Set	All subjects who receive at least one dose of CTP-required doses of study treatment (BNT331 or placebo) and with at least one measurable biomarker outcome from available samples.
BV = bacterial vaginosis; CTP = Clinical trial protocol; ITT = intent-to-treat; mITT = modified intent-to-treat; PK = pharmacokinetic.	

The Safety Set will be used for Part A and Part B safety analyses, while the modified Intent-to-Treat (mITT) Set will primarily be used for Part B efficacy analyses.

Subjects who receive the wrong treatment will be analyzed according to the actual treatment they received in the efficacy analyses and safety analyses.

3.2 Violations and Deviations

Clinical trial protocol deviations are failures to adhere to the inclusion/exclusion criteria and CTP requirements and will be classified as important or not important protocol deviations. More details are provided in the “Protocol Deviation Management Plan”. Clinical trial protocol waivers or exemptions which are prospective approval of CTP deviations to recruitment and enrollment criteria, are not permitted.

Important protocol deviations (IPD) will be identified and categorized following medical review before database snapshot/database lock for analysis. Important protocol deviations are deviations that may significantly impact the completeness, accuracy, and/or reliability of key study data or that may significantly affect a subject’s rights, safety, or well-being.

4. STATISTICAL ANALYSIS

Any deviations from the planned analyses, for the primary and secondary endpoints, described in this SAP will be described and justified in the CSR.

4.1 General Considerations

If any randomized subject is found not to have valid documented informed consent (IC), that subject’s data will be excluded from the CSR, except as necessary to document the error. Unless specified otherwise, each study part will be summarized and analyzed separately. Tables, listings, and figures will be presented by part, cohort (and treatment group), visit and timepoint as applicable.

All safety analyses will be performed using the Safety Set, unless otherwise specified. Subjects will be analyzed according to the treatment they actually received. Safety data that will be summarized includes study treatment exposure, adverse events (AE), clinical laboratory parameters, physical examinations, and vital sign assessments.

All efficacy analyses will be performed using the mITT Set, unless otherwise specified. The Intent-to-Treat (ITT) Set may be used to provide a sensitivity analysis of selected endpoints only.

All medical history and AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA) and the primary system organ class (SOC) and preferred term (PT) will be used for analyses. All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) and Anatomic Therapeutic Chemical (ATC) coding will be used for analyses. The version of the dictionaries will be the dictionary current at the time of data snapshot (data extraction).

Except where specified otherwise, all continuous variables will be summarized with descriptive statistics and all categorical variables will be summarized with frequency counts and percentages in each category, and by treatment group (and cohort). Overall summaries may be provided.

All continuous variables are listed below in the preferred order. If the original value has N decimal places, the default summary statistics are to be presented with the following decimal precision (using SAS® function ROUND as the very last step prior to presentation):

- Number (n) of subjects in each category.
- Mean and/or geometric mean (data dependent): N + 1.
- Standard deviation (SD): N + 2.
- Median: N + 1.
- Lower quartile (Q1): N + 1.
- Upper quartile (Q3): N + 1.
- Minimum: N.
- Maximum: N.
- Number of missing values.

For PK, a maximum of four significant digits will be displayed.

All categorical variables will be summarized as follows:

- Number (n) of subjects with the measurement available.
- Percentage (%) of subjects in each category presented to one decimal place (using SAS® function ROUND as the very last step prior to presentation) and calculated relative to either one of the following, unless otherwise specified:
 - Number (N) of subjects in the analysis set by treatment group.
 - Number (N) of subjects in the analysis set by treatment group with data available (observed cases).

For conversion of days to years, months, or weeks, the following conversion factors will be used:

- 1 year = 365.25 days.
- 1 month = $365.25/12 = 30.4375$ days.

- 1 week = $365.25/52 = 7$ days.

For any laboratory value reported using a non-numeric qualifier (e.g., less than [$<xx.x$] a certain value, or greater than [$>yy.y$] a certain value), the inequality sign will be removed, and the resulting numeric value will be used as the observed value, or for change from baseline (CHG) calculations. When evaluating the value to determine the Food and Drug Administration (FDA) toxicity grading, the numeric values will be used; however, if the numeric value is equal to a cutoff value, the inequality signs ($<xx.x$ or $>yy.y$) will be considered.

The intensity of (serious) AEs will be graded by the investigator by following the Table for Grading the Severity of Adult and Pediatric Adverse Events, [DAIDS 2017](#). Division of AIDS (DAIDS) Table for Grading the Severity of Adult and Pediatric Adverse Events. Version 2.1, published, July 2017. Concentrations below the limit of quantitation (BLQ) will be set to either zero or missing in the calculation of summary statistics wherever appropriate. These values will not be imputed in the by-subject data listings. Details of how BLQ values will be treated are described in Section [4.5.1.1](#).

All analyses will be conducted using SAS® version 9.4 or higher.

4.1.1 Definitions, Derivations and Classifications

Definition of Baseline

The baseline value for all safety endpoints (i.e., laboratory parameters, physical examinations and vital signs) will be the last available non-missing value (scheduled or unscheduled) prior to first dose of study treatment (BNT331 or placebo). Assessments taken on the same day as study treatment will be assumed to be before treatment unless indicated to be obtained post-treatment.

However, AEs and medications/procedures starting on the first day of study treatment will be considered post-baseline, i.e., as treatment-emergent adverse events (TEAEs) or concomitant medications.

Definition of Day 1

The date of first and last dose of study treatment will be based on the earliest (minimum) and the latest (maximum) dose date of BNT331 or placebo in each of the study parts. Treatment details will be obtained as indicated on the “Study Drug Administration” eCRF (Part A), on the “End of Treatment” eCRF (Part A and B), and in the e-diary (Part B).

Study Day 1 will be defined as the first day of study treatment for each of the study parts. This will also be considered the overall reference start date and it may be presented in selected by-subject data listings where an assessment date or event date is presented.

Definition of Study Day

Study day will be computed as follows:

- Study Day = ($\text{Date of event} - \text{date of first dose}$) + 1 if the date of the event is on or after the overall reference start date.

- Study Day = (<Date of event> – <date of first dose>) if the date of the event is prior to the date of first dose.
- Treatment Day = (<Treatment date> – <date of first dose>) + 1.
- Study Day post-treatment (relative day post-treatment) = (<Date of event> – <date of most recent treatment>) + 1 where most recent treatment refers to the last treatment received before the date of interest.

Study days will only be presented for subjects with at least one dose of study treatment or part thereof. Unless otherwise specified, if an assessment date is partial or missing, study day, and any corresponding durations will be presented as partial or missing in the by-subject data listing.

Derivations Involving Dates

The number of days between two dates, the date of interest and the reference date, will be calculated as

- Number of days = (<date of event> – <date of first dose>) + 1 when the event date is on or after the date of first dose.
- Number of days = (<date of first dose> – <date of event>) when the event date is before the date of first dose.

The resulting value is then converted to years, months, or weeks, as appropriate. Partial dates will be imputed before calculating the number of days (See [Appendix 2](#) for partial or unknown date imputation rules).

Common Calculations

Change from baseline will be calculated as:

- Change from baseline = (<Actual [observed] value at post-baseline visit> – <baseline value>).

Percent change from baseline (PCHG) will be calculated as:

- Percent change from baseline (%) = (<Change from baseline at post-baseline visit> / <baseline value>) x 100.

If baseline is not available, the CHG and PCHG will not be calculated and will remain missing.

Demographics and Disposition

Age in years at time of enrollment will be derived relative to the date of IC using the following SAS® code:

- Age (years) = (year of informed consent – year of birth).

Race as recorded on the 'Demographics' eCRF:

- White.
- Asian.

- Black or African American.
- American Indian or Alaska Native.
- Native Hawaiian or Other Pacific Islanders.
- Not reported.
- Unknown.
- Other.

Body mass index (BMI) was derived as weight in kilograms (kg) divided by height in meters squared (i.e., weight/[height*height] with unit kg/m²) by the data entry system and as indicated on the “Physical Measurements” eCRF. Body mass index will be categorized into four categories:

- <18.5.
- ≥18.5 to <25.
- ≥25 to <30.
- ≥30.

Duration of subject participation will be derived relative to the overall reference start date. The date of the visit corresponding to the Follow-up visit (Visit 3 for Part A and Visit 5 for Part B) will be used as the end date of subject participation in the study.

Subgroup Definitions

The following subgroups may be derived and assessed for statistical analysis.

- Race:
Refer to ‘Demographics and Disposition’ section.
- BMI:
Refer to ‘Demographics and Disposition’ section
- Number of previous BV episodes (Part B) in the past 12 months:
 - ≥2.
 - <2.
- Method of contraception at Screening (Part B; in case more than one method was used/documented, the below hierarchy will be used to assign the subjects to disjoint subgroup categories):
 - 1) Hormonal contraception (estrogen and progestogen).
 - 2) Hormonal contraception (progestogen).
 - 3) Intra-uterine device.

- 4) Tubal occlusion.
- 5) Vasectomy.
- 6) Abstinence.

Subjects who cannot be categorized to a subgroup due to missing data will be excluded from the corresponding subgroup analysis. The number of subjects with missing data will be provided with each subgroup analysis.

4.1.2 Windowing, Unscheduled Visits and Repeat Test Results

In general, for all by-visit summaries, data recorded at nominal visits (expected visits) are to be presented. Unscheduled assessments are not to be included in by-visit summaries, but are to contribute to baseline, and abnormal results as per specified criteria or best/worst case results where required.

For any incidence summaries of laboratory grades, all visits are to be considered including unscheduled visits.

In the case of a retest (same sample collection date), the latest available result for that visit will be used for by-visit summaries.

By-subject data listings are to include scheduled, unscheduled, and retest data (if applicable).

As part of external data file handling, the following visit windowing will be used:

Part	Analysis visit (code)	Analysis visit	Target day	Time window
Part A	0	SCREENING		-30 to -1
Part A	10	VISIT1 DAY1	1	1
Part A	20	VISIT2 EOT	2	2
Part A	24	TELEPHONE A	3	3 to 6
Part A	30	VISIT3	7	7 to 12
Part B	0	SCREENING		-2 to <1
Part B	10	VISIT1 DAY1	1	1
Part B	20	VISIT2 EOT	6	2 to 6
Part B	28	TELEPHONE B	8	7 to 18
Part B	30	VISIT3	21	19 to 45
Part B	40	VISIT4	60	46 to 90
Part B	50	VISIT5	120	>90

4.1.3 Dropouts, Missing/Incomplete Data, and Outliers

Missing data occurs when any requested data are not provided, leading to blank fields on the collection instrument. These data will be indicated using a “blank” in the by-subject data listing displays. Answers such as “Not applicable” and “Not evaluable” are not considered to be missing data and should be displayed as such.

For the binary efficacy endpoints (i.e. for Clinical and Nugent Score cure, as well as for Responder outcome), if rescue medications as referenced in the CTP Section 6.9.2 are used, the subject will be specified as a non-responder from the time the rescue medication is used. If a subject withdraws from the study, this subject will be counted as a non-responder for endpoints after withdrawal.

Missing or incomplete dates will be imputed according to the rules specified in [Appendix 2](#).

Uncoded events and procedures at the time of analysis will be assigned the string “UNCODED” as the SOC, and the reported term will be used as the PT so they can be included in the summary tables. In the final dataset, all the events and procedures should be coded.

Uncoded medications at the time of analysis will be assigned the string “UNCODED” as the ATC summary level(s) and the verbatim term will be used as the preferred name, to be included in the summary tables. In the final dataset, all the medications should be coded.

4.2 Subject Disposition

Subject disposition will be analyzed using the Screened Set by treatment group, unless otherwise specified.

A subject is considered to have completed the study if they have completed the Follow-up visit (Visit 3 or Visit 5 for Part A or Part B, respectively).

Subjects who permanently discontinued study treatment will not be discontinued from the study, except when a subject withdraws consent for further continuation in the study and with study procedures.

The number and percentage of subjects who have signed informed consent (screened subjects), are randomized, received study treatment, and completed the study will be summarized. Reasons for not completing treatment and the study will also be summarized. A subject’s study participation status (study completion/study discontinuation) will be based on the “End of Study” eCRF as indicated by the investigator. Treatment status will be based on the “End of Treatment” eCRF.

For those subjects who are considered to have failed screening as indicated on the “Eligibility” eCRF, the reason(s) for failure will be summarized.

A summary of subject enrollment by site will also be presented.

Important protocol deviations will be summarized by deviation category and the by-subject data listing will also include IPDs only.

The number and percentage of subjects included and excluded from the analysis sets will be summarized; reason(s) for exclusion from each set will additionally be listed.

4.3 Subject Characteristics

4.3.1 Demographic and Baseline Characteristics

Descriptive summaries of demographic, and baseline characteristics (physical measurements) will be presented by part, cohort (and treatment group) for the Safety Set as indicated by the investigator on the “Demographics”, and “Physical Measurements” eCRFs.

Demographic data include:

- Age (years) at IC.
- Race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Not Reported, Unknown, and Other. Subjects reporting multiple races will be included in a “Multiple” category.
- Ethnicity: Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown.

Baseline characteristics (physical measurements) include:

- Weight (kg).
- Height (cm).
- BMI (kg/m^2).
- BMI group: <18.5 , ≥ 18.5 to <25 , ≥ 25 to <30 , ≥ 30 kg/m^2 .

Descriptive statistics will be used to summarize the baseline assessments on Nugent score for subjects in 5 mg and 20 mg DL cohort and number of previous BV episodes (Part B only), by treatment group. In addition, the presence of all four Amsel’s criteria (4/4) at baseline for Part B subjects will be summarized using descriptive statistics. Summaries will be produced for each part, by cohort (and treatment group) for the Safety Set.

4.3.2 Baseline Assessments

Disease-specific characteristics are included in Section [4.3.4](#).

4.3.3 Medical and Other History

Medical history and menstrual cycle history will be recorded at Screening as indicated by the investigator on the “Medical History”, and “Menstrual Cycle History” eCRFs. For Part B subjects, BV history will also be recorded at Screening as indicated by the investigator on the “Bacterial Vaginosis History” eCRF.

Section 4.1.3 contains the data handling rules for uncoded events. Medical history will be classified using MedDRA terminology and summarized using numbers and percentages by primary SOC, PT, for each part, by cohort (and treatment group) for the Safety Set.

The following characteristics of menstrual history will be summarized for Part A and Part B:

- Duration of period/menstruation (days).
- Number of days in cycle.

The following characteristics of BV history will be summarized for Part B only:

- Number of prior BV episodes (ever, and in the last 12 months).
- Number of subjects with prior BV episodes (ever, and in the last 12 months).
- Use of treatment for previous BV (including types of treatment if further specified).
- Previous BV treatment outcome.

Menstrual cycle and BV history will be listed in by-subject data listings for the Screened Set. Medical history will be listed in a by-subject listing for the Safety Set.

4.3.4 Disease-Specific Characteristics

Disease History

For menstrual cycle history and BV history, refer to Section 4.3.3.

Nugent Score

For Part A and Part B subjects descriptive summaries of baseline Nugent score will be presented by part, cohort (and treatment group), and visit for the Safety Set.

Amsel's Criteria

For Part A enrollment, all subjects need to be negative for Amsel's criteria. None of the four (0/4) Amsel's criteria may be met at Screening.

For Part B subjects, eligibility for enrollment will be confirmed by the presence of all four Amsel's criteria (4/4) at Screening. The clinical diagnosis of BV will be based on Amsel's criteria while the microbiological diagnosis will be based on Nugent scoring (Nugent et al. 1991). Per the Amsel's criteria, BV is defined as the presence of at least three of the following four criteria (3/4) at Visits 2, 3, 4, and 5:

- Homogeneous, thin, grayish-white discharge that smoothly coats the vaginal walls.
- Positive Whiff-amine test, defined as the presence of a fishy odor when 10% potassium hydroxide (KOH) is added to a sample of vaginal discharge.
- Clue cells (i.e., vaginal epithelial cells covered with bacteria) on saline wet mount, comprising at least 20% of epithelial cells per high power field.
- Vaginal pH >4.5. Only for inclusion, vaginal pH is not included as a component of the clinical cure for BV.

Clinical diagnosis will be presented as indicated by the investigator on the “Amsel’s Criteria” eCRF. The number and percentage of subjects will be summarized by part, cohort (and treatment group) as applicable for the following:

- Amsel’s criteria assessment performed: Yes, No.
- Reasons for not collecting Amsel’s criteria.
- Presence of grayish-white discharge: Yes, No.
- Whiff-amine test: Positive, Negative.
- Clue cells on saline wet mount, comprising more or less than 20% of epithelial cells per high power field: $\geq 20\%$, $< 20\%$.
- Vaginal pH: > 4.5 , ≤ 4.5 .

4.3.5 Medications, Therapies and Procedures

4.3.5.1 Prior and Concomitant Medications

Prior and concomitant medications will be presented as indicated in the “Prior and Concomitant Medications” eCRF.

Prior medications are medications and therapies with a start and end date before the first dose of study treatment.

Concomitant medications are medications with a start date after the first dose of study treatment until the end of study, except the medications for treating recurrent BV in Part B subjects after the ToC visit and any medications that are considered as the BV salvage treatment for Part A and Part B subjects. Medications with a start date before the first dose of study treatment and still being taken after the first dose of study treatment will also be classified as concomitant medications. If a medication or therapy cannot be classified due to completely missing dates, it will be classified as concomitant.

Rescue medications, defined as any medications that are considered as the BV salvage treatment or medications for treating recurrent BV in Part B subjects after the ToC visit, will be summarized for Part B, by cohort (and treatment group) for the Safety Set. Details for identifying rescue medications will be documented in the study data specifications.

Medications and therapies (prior, concomitant, rescue) will be summarized by ATC therapeutic class (ATC Level 2), ATC pharmacological class (ATC Level 3), and chemical substance (preferred name) for each part, by cohort (and treatment group) for the Safety Set. In addition, medications will also be presented in by-subject data listings.

4.3.5.2 Prior and Concomitant Procedures

Prior and concomitant procedures and/or non-drug therapy data will be presented as indicated on the “Concomitant Procedures” eCRF and included in a by-subject data listing for the Safety Set.

Prior procedures and/or non-drug therapies are procedures/non-drug therapies with a start and end date before the first dose of study treatment. Concomitant procedures and/or non-drug therapies are procedures/non-drug therapies with a start date after the first dose of study treatment up to and including the last dose of study treatment.

4.4 Primary Endpoint Analyses

The primary objective is to evaluate the safety and tolerability of BNT331 when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.

Endpoints for each dose cohort of BNT331 and combined placebo groups include:

- Occurrence of AEs with onset after the first treatment dose and until seven days post-dose (Part A) and 120 days after the first dose (Part B).
- Occurrence of SAEs with onset after the first treatment dose and until seven days post-dose (Part A) and 120 days after the first dose (Part B).

The primary endpoint is a safety endpoint. Refer to Section 4.8.2 for more details about the analysis methods to be implemented.

4.5 Secondary Endpoint Analyses

For the PK and ADA analyses described in the study objectives, the parameters will be determined using serum concentrations (secondary objective) of BNT331 active substance.

4.5.1 Secondary Endpoint 1 Analysis – Pharmacokinetic Analysis (Serum)

The objective is to determine the systemic PK of BNT331 active substance, including quantification of systemic absorption when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.

Endpoints for each treatment group of BNT331 include:

- Part A: The serum concentrations of BNT331 active substance at pre-specified timepoints among subjects who received one single administration.
- Part B: The serum concentrations of BNT331 active substance at pre-specified timepoints among subjects who received all scheduled administrations.

Pharmacokinetics – Serum BNT331 Active Substance Concentration

Blood samples will be collected by the investigator as indicated on the “Blood Draw for PK Samples (Multiple Timepoints)”, and “Blood Draw for PK Samples” eCRFs and tested in the PK/ADA specialty lab. The systemic PK analysis will be performed based on the PK Set. A by-subject data listing will be provided based on the Safety Set.

4.5.1.1 Pharmacokinetics – Analytical Approach for Serum

Serum concentrations for Part A and B will be summarized using descriptive statistics (number of subjects [N], number of observations [n], arithmetic mean, SD, coefficient of

variation percentage [CV%], geometric mean, geometric CV%, median, minimum, maximum; where feasible, geometric mean and geometric CV% may also be calculated) for each part, cohort (and treatment group), visit and timepoint as applicable based on the PK Set. In addition, a by-subject data listing will be provided based on the Safety Set.

The PK serum concentrations of BNT331 active substance will be summarized using descriptive statistics for each dose level cohort in both Part A and Part B. BNT331 active substance concentration-time profiles will be generated for each dose level cohort in both Part A and Part B.

Pharmacokinetic analyses will be performed using computer program Phoenix® WinNonlin® version 8.3, or higher (Certara, L.P., Princeton, New Jersey, USA). Pharmacokinetic parameters will be estimated for Part A, if data permits, using a non-compartmental analysis method with the actual date and time of each sample. Individual and mean ($\pm$ SD) serum concentration will be plotted over time on semi-logarithmic scale for Part A. Individual serum concentration will also be plotted over time on semi-logarithmic scale for Part B.

Clinical trial protocol deviations and study data will be reviewed to identify subjects or data points which may be affected, and thus excluded from PK analyses and PK summaries. Samples collected outside of the time window of the specified time post-dose will be included for the purpose of PK parameter calculation, but will be excluded for the purpose of summarizing concentrations of each time point. Changes to the procedures or events which may impact the quality of the PK data, will be considered IPDs or events, and will be described within the CSR. These changes or events will include any circumstances that will alter the evaluation of the PK. Examples include, but may not be limited to sample processing errors that lead to inaccurate bioanalytical results, and/or missed or inaccurate dosing on the day of PK measurements. In the case of an important protocol deviation or event, affected PK data will be excluded from any summaries or analyses. Other CTP deviations or events which are not expected to markedly impact the PK results and associated data interpretation will not be considered important protocol deviations for PK analyses.

Serum concentrations that are BLQ will be set to zero for the computation of descriptive statistics and for individual figures. For PK parameter calculations, post-dose BLQ concentrations that precede quantifiable samples in the initial portion of the PK profile will be assigned a value of zero. A BLQ value that occurs between quantifiable data points, especially prior to C_{max} , will be evaluated to determine if an assigned concentration of zero makes sense, or if exclusion of the data is warranted. Following C_{max} , BLQ values embedded between two quantifiable data points will be treated as missing when calculating PK parameters. If a BLQ value occurs at the end of the collection interval (after the last quantifiable concentration), it will be set to zero. If consecutive BLQ concentrations are followed by quantifiable concentrations in the terminal portion of the concentration curve, these quantified values will be excluded from the PK analysis by setting them to missing, unless otherwise warranted by the concentration-time profile. If an anomalous concentration value is observed at the terminal phase and is judged as physiologically unreasonable, it will be excluded from computing PK parameter estimates. Excluded data and the reason(s) for exclusion will be described in a by-subject data listing.

For PK parameter calculations, pre-dose samples that are BLQ or missing for the first dose will be assigned a numerical value of zero. Any anomalous concentration values observed at pre-dose will be identified in the CSR and used for the computation of PK parameters. If the anomalous pre-dose concentration is not greater than 5% of the C_{max} in the profile, the PK parameters for the given profile will be calculated and reported in the by-subject data listing and included in statistical summaries. If the anomalous concentration is greater than 5% of C_{max} in the profile, the impact of the anomalous pre-dose concentration will be reviewed with the Sponsor for possible exclusion of the affected subject from summaries for both concentrations and derived PK parameters.

The following serum PK parameters to characterize systemic disposition of BNT331 in Part A may be calculated using non-compartmental methods and actual elapsed time from dosing, as appropriate, where possible.

C_{max}	Maximum observed concentration ($\mu\text{g/mL}$), obtained directly from the observed concentration versus time data, with at least one post-dose quantifiable concentration.
t_{max}	First time to reach C_{max} (h), obtained directly from the observed concentration versus time data.
AUC_{0-t}	Area under the concentration-time curve from time zero (pre-dose) to time of last quantifiable concentration ($\mu\text{g/mL h}$), calculated by linear up/log down trapezoidal summation with at least two consecutive post-dose concentrations.
AUC_{0-inf}	Area under the concentration-time curve from time zero (pre-dose) extrapolated to infinity ($\mu\text{g/mL h}$), calculated by linear up/log down trapezoidal summation and extrapolated to infinity by addition of the last quantifiable observed concentration divided by the elimination rate constant: $AUC_{0-last} + C_{last}/\lambda_z$, where possible.
$t_{1/2}$	Apparent elimination half-life (h), determined as $\ln 2/\lambda_z$, where possible.
CL/F	Apparent clearance after intravenous dosing (L/h), calculated as dose divided by AUC_{0-inf} , where possible
V_z/F	Apparent volume of distribution following dosing, calculated as apparent clearance (L/h) divided by apparent terminal rate constant λ_z (1/h), where possible
Rsq adjusted	Adjusted coefficient of determination for calculation of λ_z . If Rsq_{adj} is less than 0.800, associated parameters AUC_{0-inf} , λ_z , CL/F , V_z/F , and $t_{1/2}$ will be excluded from descriptive statistics but will be listed.
λ_z	Apparent terminal rate constant (1/h), determined by linear regression of the terminal points of the log-linear concentration-time curve. The Best Fit method utilized by WinNonlin will be used to identify the terminal linear phase of the concentration-time profile, with visual assessment and adjustment of the selected data points by the PK scientist if warranted. A minimum of three data points will be used for determination.
λ_z lower	The starting time point (h) of the time interval of the log-linear regression to determine λ_z .
λ_z upper	The ending time point (h) of the time interval of the log-linear regression to determine λ_z .
$\%AUC_{ex}$	Percentage of AUC_{0-inf} obtained by extrapolation, calculated as $[(C_{last}/\lambda_z)/AUC_{0-inf}] \times 100$. If the $\%AUC_{ex}$ is greater than 20.0% of AUC_{0-inf} , then AUC_{0-inf} and parameters derived from it will be listed but not included in summary and inferential statistics.

4.5.2 Secondary Endpoint 2 Analysis - ADA Response Analysis (Serum)

The objective is to determine ADA response in healthy women and in women diagnosed with BV.

The endpoint for each treatment group of BNT331 includes:

- Part A and Part B: Anti-drug antibody prevalence and change of binding titers against BNT331 active substance in blood before study treatment (pre-dose) and at seven days post-dose (Part A) and at six days after the first dose, 21 to 30 days after the first dose, and 60 days after the first dose (Part B).

Blood samples will be collected by the investigator as indicated on the “Blood Draw for Anti-Drug Antibody” eCRF and tested in the PK/ADA specialty laboratory. The ADA response analysis will be performed based on the Biomarker Set.

The ADA samples will be batched for analysis in the PK/ADA specialty laboratory. First a qualitative screening test will be performed, followed by a qualitative confirmation test. If the qualitative confirmation test is positive, then the sample will be tested for antibody titer. Only the results from the last test performed will be reported.

Subjects will be classified as “ADA-negative” or “ADA-positive” depending on if their baseline levels were negative or positive, respectively; each analyte will be classified independently. Specifically, the following definitions will be used.

ADA-positive sample	A sample confirmed to be positive for the presence of reactive antibodies. A sample must be positive in both the screening assay and the confirmation assay. The result will be reported as “Titer” which is considered positive for the presence of antibodies.
ADA-negative sample	A sample that is reported as negative by the screening assay or by the confirmation assay.
Treatment-emergent ADA	Defined as any post-treatment positive ADA assay response when the baseline ADA result is negative or missing.
Treatment-boosted ADA	Defined as any post-treatment positive ADA assay response that is greater than or equal to 4-fold over baseline titer level when baseline is ADA-positive in the ADA assay.
Pre-existing immunoreactivity ADA	Defined as either a positive ADA assay response at baseline with all post-treatment ADA assay results negative, or a positive assay response at baseline with all post-treatment ADA assay responses less than 4-fold over baseline titer levels.

The number and percentage of subjects with ADA-positive or negative samples at baseline and post-baseline will be summarized by part, cohort (and treatment group), and visit as applicable for the Biomarker Set.

A by-subject data listing will be provided based on the Biomarker Set.

In addition, for each post-baseline visit, the number and percentage of subjects with treatment-emergent ADA, treatment-boosted ADA, and pre-existing immunoreactivity ADA will be summarized by part, cohort (and treatment group), and visit as applicable for the Biomarker Set.

The 4-fold threshold in the definitions of treatment-boosted ADA and pre-existing immunoreactivity ADA may subject to changes. The Sponsor may update the threshold based on the actual data and clinical meanings, and justifications for the changes will be provided in

the CSR. If insufficient data to generate any meaningful conclusions, only the by-subject listing will be provided.

4.5.3 Secondary Endpoint 3 Analysis – Clinical and Nugent Score Cure and Responder Outcome Analyses (Part B Only)

For Part B, the objective is to evaluate the proportion of subjects with diagnosed BV who meet the definitions of clinical cure, Nugent score cure, and responder outcome at Visit 2 (early response) and Visit 3 (ToC).

Endpoints for each treatment group of BNT331 and combined placebo groups include:

- Occurrence of Clinical cure defined as normalization of the vaginal discharge, a negative KOH Whiff test, and clue cells <20% of the total epithelial cells/high power field on microscopic examination of the vaginal fluid at six days after first dose and 21 to 30 days after first dose.
- Occurrence of Nugent score cure/Microbiological cure defined as a normal Nugent score (<4) at six days after first dose and 21 to 30 days after first dose.
- Responder outcome defined as an occurrence of clinical cure with a normal Nugent score (<4) at six days after first dose and 21 to 30 days after first dose.

The number and percentage of subjects for each of the above endpoints will be summarized by treatment group and visit as applicable for the mITT Set. If rescue medication is used, the subject will be specified as a non-responder (for Clinical and Nugent Score cure, as well as for Responder outcome) from the time the rescue medication is used. If a subject withdraws from the study, this subject will be counted as a non-responder for endpoints after withdrawal.

Sensitivity analysis may be conducted on observed data, i.e., all subjects with available data at the relevant visit with no imputation for missing data or rescue medication usage.

For the Nugent scoring system, the number of subjects and percentages will be summarized by the following categories by treatment group and visit as applicable for the mITT Set:

- Normal vaginal microbiota: 0 to 3.
- Intermediate vaginal microbiota: 4 to 6.
- Bacterial vaginosis: 7 to 10.

Vaginal swab samples will be collected by the investigator as indicated on the “Vaginal Swab for Nugent Score” eCRF and tested in the central lab.

4.6 Exploratory Endpoint Analyses

4.6.1 Exploratory Endpoint 1 Analysis – Pharmacokinetic Analysis (Vaginal Fluid)

The objective is to determine the local PK of BNT331 active substance when administered vaginally as SAD in healthy women and MAD in women diagnosed with BV.

Endpoints for each cohort include:

- Part A: The vaginal fluid concentration of BNT331 active substance over time before (pre-dose), and after study treatment administration (four and 24 hours) and at seven days post-dose.
- Part B: The vaginal fluid concentration of BNT331 active substance over time before (pre-dose), and eight to 24 hours after the last study treatment administration.

Pharmacokinetics – Vaginal Fluid BNT331 Active Substance Levels

Vaginal swab sampling for BNT331 active substance topical PK will be performed at the visits given in the SoA for the study ([Appendix 1](#)). After pre-dose sample, the first post-dose timepoint will be four hours after dose in Part A and eight to 24 hours after the last dose in Part B and was chosen to minimize interference with local distribution and systemic absorption. The vaginal swabs intended for measurement of BNT331 active substance levels will be collected, processed, and stored as per Laboratory Manual and will then be transferred to a specialty bioanalytical laboratory for analysis.

Vaginal swab samples will be collected by the investigator as indicated on the “Vaginal Swab for PK Samples (Multiple Timepoints)”, and “Vaginal Swab for PK Samples” eCRFs and tested in the PK/ADA specialty laboratory.

Concentrations that are BLQ will be set to zero for the computation of descriptive statistics. Excluded data and the reason(s) for exclusion will be described in a by-subject data listing.

4.6.2 Exploratory Endpoint 2 Analysis – ADA Response Analysis (Vaginal Fluid)

The objective is to determine the local ADA response when BNT331 is administered as SAD in healthy women and MAD in women diagnosed with BV.

The endpoint for each treatment group of BNT331 include:

- Part A and Part B: Vaginal ADA presence as signal to noise (S/N) ratio against BNT331 before study treatment (pre-dose) and at seven days post-dose (Part A) and 21 to 30 days after the first dose (Part B).

Vaginal swab samples will be collected by the investigator as indicated on the “Vaginal Swab for Anti-Drug Antibody” eCRF and tested in the PK/ADA specialty laboratory. Anti-drug antibody S/N ratio analyses will be summarized by part, cohort (and treatment group), and visit as appropriate based on the Biomarker Set.

4.6.3 Exploratory Endpoint 3 Analysis - Change in the Overall Vaginal Microbiome

The objective is to characterize the pharmacodynamics of BNT331 by evaluating the changes in the overall vaginal microbiome of subjects following treatment with BNT331, and to compare the change in vaginal microbiome with the clinical outcomes.

The endpoint for each treatment group of BNT331 and combined placebo group from 5 mg cohort and 20 mg cohort in Part A and Part B include:

- Change of relative abundance of bacterial taxa by shotgun metagenomic sequencing (such as identification of BV associated bacteria including *Gardnerella spp.* and *Lactobacillus spp.*) from baseline (before treatment) until seven days post-dose (Part A) and until 120 days after the first dose (Part B).
Relative abundance will be set to 0 at the respective visit / time point if there is a valid measurement of the sample, but no bacteria of the considered taxon were detected by molecular method.
- Change of Nugent score from baseline (before treatment) until 7 days post-dose (Part A) and until 120 days after the first dose (Part B).

Vaginal swabs for microbiome will be analyzed at the timepoints listed in the schedule of assessments for the study ([Appendix 1](#)). Vaginal swab samples for microbiome will be collected as indicated by the investigator on the “Vaginal Swab for Microbiome Assessment” eCRF and tested in the central laboratory.

The individual results of abundance in the vaginal microbiome analysis from baseline (before treatment) until seven days post-dose (Part A) and until 120 days after the first dose (Part B) by shotgun metagenomic sequencing will be presented in the stagger bar plot for each part, by cohort (and treatment group), and visit for the Biomarker Set.

Vaginal swabs for Nugent score will be analyzed at the timepoints listed in the schedule of assessments for the trial ([Appendix 1](#)). Vaginal swab samples for Nugent score will be collected by the investigator as indicated on the “Vaginal Swab for Nugent Score” eCRF and tested in the central laboratory.

Results of change in the Nugent score analysis from baseline (before treatment) until seven days post-dose (Part A) and until 120 days after the first dose (Part B) will be summarized for each part, by cohort (and treatment group) for the ITT Set (Part A) and for the mITT Set (Part B).

Evaluating the pharmacodynamics of BNT331 is a core element of the study setup. To evaluate to which extent the intended pharmacodynamical effect of BNT331 (killing *Gardnerella*) leads to changes in the clinical outcomes, three correlation analyses will be performed:

- Correlation of change in abundance of *Gardnerella* (V3 minus baseline, by whole genome sequencing [WGS]) against abundance of *Gardnerella* at baseline (by WGS), one point per subject in a correlation graph with linear axes, one panel each for 5 mg, 20 mg and placebo.
Relative abundance will be set to 0 at the respective visit / time point if there is a valid measurement of the sample, but no bacteria of the considered taxon were detected by molecular method.
- Correlation of change in Nugent Score (V3 minus baseline) against Nugent Score at baseline, one point per subject in a correlation graph with linear axes, one panel each for 5 mg, 20 mg and placebo.

- Dual-y-axis graph with change in Nugent score (V3 minus baseline) on left y-axis and change in abundance of *Gardnerella* (V3 minus baseline, by WGS) on right y-axis against abundance of *Gardnerella* at baseline (by WGS), two points per subject (one for each y-axis with identical x-value) in a correlation graph with linear axes.
Relative abundance will be set to 0 at the respective visit / time point if there is a valid measurement of the sample, but no bacteria of the considered taxon were detected by molecular method.

4.6.4 Exploratory Endpoint 4 Analysis - Change in the Levels of *Gardnerella* and *Lactobacillus* spp.

The objective is to characterize the pharmacodynamics of BNT331 by evaluating the changes in the levels of *Gardnerella* and *Lactobacillus* spp. from baseline (before treatment).

The endpoints for each treatment group of BNT331 and combined placebo group from 5 mg cohort and 20 mg cohort in Part A and Part B include:

- Quantitate the levels of *Gardnerella* and *Lactobacilli* from baseline (before treatment) to seven days post-dose (Part A) and to 21 to 30 days after the first dose (Part B).
Level of respective bacterial taxon will be set to LLOQ (if available, e.g. 6.25 copies/μL for *Gardnerella*, *Lactobacillus* pan and *Lactobacillus iners*), or 1 copy/μL (if LLOQ is not available), at the respective visit / time point if there is a valid measurement of the sample, but no bacteria of the considered taxon were detected by molecular method. Details will be documented in the study data specifications.

Vaginal swab samples for qPCR will be collected as indicated by the investigator on the “Vaginal Swab for qPCR” eCRF and tested in the specialty laboratory to quantitate *Gardnerella* and *Lactobacillus* levels.

The results in total and live bacteria, *Gardnerella*, *Lactobacillus* and *Lactobacillus iners* levels will be presented as a box plot, for each part, cohort (and treatment group) by visit for the Biomarker Set.

To evaluate the intended pharmacodynamic effect of BNT331 (killing of *Gardnerella*) at the two doses applied, the change in the *Gardnerella* load (Visit 3 vs. baseline, Part B) will be correlated to the *Gardnerella* load at baseline. The correlation will be performed for DNA copies/μL of both total and viable *Gardnerella* determined by qPCR and viability qPCR, respectively. In total, 3*2 correlation plots (two doses and placebo for *Gardnerella* load by qPCR and viability qPCR) will correlate the change in *Gardnerella* (Visit 3 vs. baseline) and the load of *Gardnerella* at baseline, one point per subject in a correlation graph with linear axes. Linear regression will be performed, and the slope and R squared reported.

A by-subject listing will be provided for the Safety Set.

4.6.5 Exploratory Endpoint 5 Analysis – Long-term Treatment Outcome Analysis (Part B Only)

For Part B, the objective is to evaluate the long-term treatment outcome of subjects with diagnosed BV who meet the definitions of clinical cure at Visit 3 (ToC) following treatment with BNT331.

The endpoint for each treatment group of BNT331 and combined placebo groups include:

- Part B: Occurrence of subjects reporting new BV symptoms and episodes until 120 days after the first dose.

New BV symptoms and episodes will be identified based on a change from normal to abnormal results of any individual Amsel's criteria or Nugent scores.

In case of BV recurrence at any time after Visit 3 (ToC visit) and before Visit 5 (Late Follow-up visit), Visit 4 (Follow-up/Recurrence visit) should be performed prior to administration of standard of care treatment. For Part B subjects, BV recurrence (including method of assessment, and administration of local standard care treatment) will be summarized by treatment group and visit as applicable using numbers and percentages on subjects who meet the definitions of clinical cure at Visit 3 (ToC).

Long-term treatment outcome analysis as indicated on the "BV Recurrence" eCRF will be performed based on the mITT Set.

4.6.6 Exploratory Endpoint 6 Analysis – Long-term Psychosocial Outcome Analysis (Part B Only)

For Part B, the objective is to measure the long-term psychosocial outcome of subjects with diagnosed BV.

Endpoints include patient-reported outcome measure (PROM) scores before and after treatment.

For each treatment group of BNT331 and combined placebo groups, including subjects after taking rescue medication:

- Change over time in scores of thirty-six-item short form health survey (SF-36) ([Ware and Sherbourne 1992](#)) and Changes in Sexual Functioning Questionnaire (CSFQ-F) ([Clayton et al. 1997](#)) at three timepoints (baseline, 21 to 30 days after the first dose, and 120 days after the first dose).

Details of the SF-36 questionnaire are provided in [Appendix 3](#). The change from baseline in the eight subscale domains and each summary score will be summarized descriptively at Visit 3 (ToC) and Visit 5 (Late follow-up) for Part B subjects by treatment group, and visit as applicable for the mITT Set.

Details of the CSFQ-F questionnaire is provided in to [Appendix 3](#). The sum of scores in the five domains and change from baseline for each of the five domains and the composite total

score will be summarized descriptively at Visit 3 (ToC), and Visit 5 (Late follow-up) for Part B subjects by treatment group, and visit as applicable for the mITT Set.

If not all questions are answered within a domain, no imputation for missing data will be applied and the domain will be presented as missing. In the event of missing domains for a subject the composite total score will also be presented as missing for that subject.

4.7 Other Analyses

Not applicable.

4.8 Safety Analyses

4.8.1 Extent of Exposure

Administration of study treatment will be presented as indicated on the “Study Drug Administration” eCRF (Part A), on the “End of Treatment” eCRF (Part A and B), and in the e-diary (Part B). Dose modifications are not allowed.

The following will be listed by part, cohort (and treatment group), and visit as applicable for the Screened Set:

- Subjects receiving study drug administration.
- Subjects not receiving study drug administration.
- Reasons for not receiving study drug administration.

For Part A, only a by-subject data listing will be provided for the exposure information based on the Screened Set.

For Part B subjects, the number and percentage by treatment group will be summarized for the Safety Set as following, based on the e-diary entries:

- Received dose at Day 1.
- Received dose at Day 2.
- Received dose at Day 3.
- Received dose at Day 4.
- Received dose at Day 5.
- Received ≥ 3 doses.
- Received ≥ 4 doses.
- Received all five doses.

Medication errors will be included in a by-subject data listing as indicated on the “Medication Errors” eCRF for the Safety Set.

4.8.2 Adverse Events

Adverse events will be presented as indicated on the “Adverse Events”, and “Serious Adverse Events” eCRFs. Any progression of the underlying disease during study participation will not be considered as an AE.

Adverse events with onset before the start of study treatment and AEs with onset more than 7 days (Part A) and 115 days (Part B) after the last dose of study treatment for each study part will be excluded from TEAE summaries.

Treatment-emergent adverse events are those AEs with an onset on or after the date and time of study treatment or events which were present before treatment, and which worsened after treatment. Only TEAEs will be included in summaries. See [Appendix 2](#) for handling of partial dates. In the case where it is not possible to define an AE as treatment-emergent or not, the AE will be classified by the worst case (i.e., treatment-emergent). Section [4.1.3](#) contains the data handling rules for uncoded events.

Treatment-emergent adverse events will be summarized overall and by treatment group for Part A and Part B, respectively. When summarizing events and relationship or severity to study treatment is missing for a TEAE it will be imputed as the worst case and assumed to be related or severe.

Individual summaries (number of subjects, percentages, number of events) of TEAEs will be presented by SOC and PT:

- Any TEAE.
- Treatment-related TEAE.
- Study-required procedure-related TEAE.

An overall TEAE summary table will present the number and percentage of subjects with at least one event in the above specified categories, as well as serious criteria and will include:

- Grade ≥ 3 TEAE.
- Grade ≥ 3 treatment-related TEAE.
- Grade ≥ 3 study-required procedure-related TEAE.
- Serious TEAE.
- Serious treatment-related TEAE.
- Serious study-required procedure-related TEAE.
- TEAE resulting in death.
- TEAE leading to treatment discontinuation.

In addition, a summary of TEAEs, including number of subjects and percentages, will be presented by SOC, PT, and maximum grade.

For presentation and derivation purposes, the following statistics are defined:

- n: Number of subjects with the measurement available (at least one TEAE in each category). Subjects with multiple TEAEs in each category are counted only once in each category.
- N: Number of subjects in the analysis set by treatment group.
- %: Percentages are based on the number of subjects in the analysis set by treatment group.

Adverse events leading to death are selected as all AEs with an outcome recorded as “Fatal” and will be included in a by-subject data listing and additional details as indicated on the “Death” eCRF will be provided where relevant.

4.8.3 Laboratory Data

Clinical laboratory data to be summarized includes hematology, blood chemistry, and urinalysis as indicated on the “Laboratory Tests – Central Hematology”, “Laboratory Tests – Central Blood Chemistry”, and “Laboratory Tests –Local Urinalysis” eCRFs.

The planned clinical laboratory parameters include:

- Hematology: hematocrit, hemoglobin, red blood cell count (red cell distribution width [RDW], mean corpuscular volume [MCV], mean corpuscular hemoglobin concentration [MCHC] and mean corpuscular hemoglobin [MCH]), white blood cell count and absolute differentials (neutrophils, lymphocytes, monocytes, eosinophils, basophils), platelet count.
- Blood chemistry: alkaline phosphatase (ALP), alanine transaminase (ALT), aspartate transaminase (AST), blood urea nitrogen (BUN), creatinine, non-fasting glucose, lactate dehydrogenase and total bilirubin.
- Dipstick urine analysis: bilirubin, glucose, ketones, leukocyte esterase, nitrite, pH, specific gravity, total protein, urobilinogen.

Reference ranges established by the central laboratory will be used to determine whether laboratory values are lower (L), within (N) or higher (H) than the reference range for that parameter. The low and high indicators will also be used to combine with any clinically significant abnormality for the purpose of creating abnormal summaries. A summary of abnormal incidences will be presented displaying normal, and abnormal (clinically significant or not) results for each parameter for each part by cohort (and treatment group), and visit as applicable for the Safety Set.

Listings of abnormal hematology, blood chemistry, and urinalysis laboratory test results will be provided and values that are below or above the normal ranges will be flagged. Determination of clinical significance will be made by the investigator and included in the by-subject data listing for the Safety Set.

4.8.4 Vital Signs

Vital sign parameters to be summarized include systolic blood pressure (mmHg), diastolic blood pressure (mmHg), body temperature (°C), and pulse rate (beats/minute), and will be presented as indicated on the “Vital Signs”, and “Vital Signs (Pre- and Post-Dose)” eCRFs.

Height (cm) and weight (kg) will be assessed at Screening only and will be presented as described in Section 4.3.1.

On dose administration days, vital signs will be recorded pre-dose and post-dose. On days without administration, vital signs will only be obtained once.

Vital sign test results (actual and change from baseline to each post-baseline visit and timepoint) will be summarized with descriptive statistics for each parameter for each part by cohort (and treatment group), visit and timepoint as applicable for the Safety Set.

Assessment of vital signs (including determination of clinical significance made by the investigator) will be presented in a by-subject data listing for the Safety Set.

4.8.5 Other Safety Assessments

4.8.5.1 Physical Examination

A complete physical examination will be conducted at Screening and a symptom-directed examination will be conducted at subsequent visits as specified in the schedule of assessments for the study ([Appendix 1](#)). All abnormalities will be listed for the Safety Set as indicated on the “Physical Examination - Traditional”, and “Physical Examination” eCRFs.

4.8.5.2 Pelvic Examination

Subjects will undergo a pelvic examination at Screening and at each scheduled site visit as indicated on the “Pelvic Examination”, and “Pelvic Examination (Pre- and Post-Dose)” eCRFs for assessment of local safety/tolerability and for required sampling.

On dose administration days, pelvic examination will be recorded pre-dose and post-dose. On days without administration, pelvic examination will only be performed once.

The number and percentage of subjects who have abnormal (clinically significant or not) pelvic examination results and abnormal findings (clinically significant or not) for the pelvic examination will be summarized by part, cohort (and treatment group), visit and timepoint as applicable based on the Safety Set.

4.8.5.3 Trichomonas rapid test (Part B only)

For Part B, subjects will undergo a rapid test at Screening for detection of parasitic infection as indicated on the “Trichomonas Rapid Test” eCRF. The number and percentage of subjects based on the results of detection of parasitic STI from the trichomonas rapid test will be summarized for the Safety Set.

4.8.5.4 Pregnancy Test

Urine and/or serum pregnancy test results will be presented in a by-subject data listing for the Safety Set as indicated on the “Pregnancy Test and Contraceptive Counseling” eCRF.

4.8.5.5 Local Tolerability

Tolerability in Part A will be assessed as indicated on the “Local Tolerability Investigator Assessment”, and “Subject Phone Call - Part A” eCRFs. The Investigator will perform a cervix examination and will assess signs of erythema, oedema, and excoriation on the vaginal mucosa approximately four hours (± 15 minutes) post-dosing at Visit 1. A telephone interview will be conducted three to four days post-dose to ask about any complaints of dysuria, vulvo-vaginal symptoms such as redness, itching, pain, smell/odor, discharge, and/or bleeding from the subjects to assess the local tolerability of study treatment. If during the telephone call, a subject voices genital symptoms, an in-person appointment will be made to see and examine the subject.

The number and percentage of subjects who have experienced the following symptoms (per the investigator’s assessment at Visit 1, and telephonic interview at three to four days post-dose) will be summarized by treatment group based on the Safety Set.

Investigator’s assessment during the screening will include:

- Subjects experiencing erythema.
- Subjects experiencing petechiae.
- Subjects experiencing oedema.
- Subjects experiencing excoriation.
- Subjects experiencing other symptoms.

Telephonic interview will include:

- Subjects experiencing dysuria.
- Subjects experiencing vulvo-vaginal symptoms including redness, itching, pain, smell/odor, discharge, and/or bleeding.

For Part B, the local tolerability will be collected daily through e-diary and summarized by treatment group and timepoint as appropriate based on the Safety Set.

4.9 Subgroup Analyses of Interest

Refer to Section 4.1.1 for subgroup definitions.

5. INTERIM ANALYSIS AND ANALYSIS SEQUENCE

An IA will be performed using the data of Part A and Part B after the last subject in Part B completes Visit 3 (ToC visit) to analyze all available primary and secondary endpoints, including quantification of *Gardnerella spp.* in vaginal fluid, which is an exploratory endpoint.

6. SAMPLE SIZE DETERMINATION

Planned enrollment is 96 subjects (24 subjects in Part A and 72 subjects in Part B), for an adequate evaluation of safety and tolerability.

The sample size is driven by typical small dose-escalation designs for early detection of potential safety and tolerability events, and is considered adequate to support the study objectives, while minimizing the number of study subjects exposed to a new investigational medicinal product (IMP). With 18 subjects receiving active drug in Part A, the probability of detecting at least one AE is 85% if the underlying AE rate is 10%. In Part B, assuming the same AE rate of 10%, with 24 subjects receiving active drug in each treatment group, the probability of detecting at least one AE is 92%.

7. CHANGES FROM AND ADDITIONS TO METHODS PLANNED IN THE CLINICAL TRIAL PROTOCOL

Not applicable.

8. REFERENCES

Clayton AH, McGarvey EL, Clavet GJ. (1997) The Changes in Sexual Functioning Questionnaire (CSFQ): development, reliability, and validity. *Psychopharmacol Bull.* 1997;33(4):731-745.

DAIDS 2017. Division of AIDS (DAIDS) Table for Grading the Severity of Adult and Pediatric Adverse Events. Version 2.1, published, July 2017.

Nugent RP, Krohn MA, Hillier SL. (1991). Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *J Clin Microbiol.* 1991;29(2):297-301.

Ware JE Jr, Sherbourne CD. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992 Jun;30(6):473-483.

9. SUPPORTING DOCUMENTATION

Supporting documentation is provided in appendices.

Appendix 1 Schedule of Procedures/Assessments

The schedule of assessments for the study is presented below. The schedule of assessments for Part A is shown in [Table 4](#) and for Part B it is shown in [Table 5](#).

Table 4: Schedule of Activities – Part A

Description	Screening	Dosing	End of therapy	Telephone interview	Follow-up
Visit ID	Visit 0	Visit 1	Visit 2 ⁱ	(No visit)	Visit 3
Time relative to dosing (incl window)	Within 30 d prior Visit 1	NA	1 d post-dose (±0 d)	3 d post-dose (+1 d)	7 d post-dose (+5 d) ^m
Informed consent	X				
Inclusion/exclusion criteria	X	X (confirm)			
Urinalysis ^a	X				X
Pregnancy test and counseling for WOCBP ^b	X	X (pre-dose)			X
Medical history incl. menstrual cycle	X	X (update)			
Demography	X				
Prior and concomitant medications	X	X	X	X	X
Physical examination ^c	X				X
Pelvic examination ^d	X	X (pre- and post-dose)	X		X
Amsel's criteria ^e	X				
Vital signs ^f	X	X (pre- and post-dose)	X		X
Blood draw for clinical laboratory tests ^g (15 mL)	X	X ⁿ (pre-dose)			X
Randomization		X (pre-dose)			
Treatment instructions and dispensing		X (pre-dose)			
Administer study treatment ^p		X			
Blood sampling for BNT331 active substance concentration (PK) in serum ^h (5 mL)		X (pre- and post-dose)	X		X
Blood sampling for ADA response in serum (5 mL)		X (pre-dose)			X

Description	Screening	Dosing	End of therapy	Telephone interview	Follow-up
Visit ID	Visit 0	Visit 1	Visit 2 ⁱ	(No visit)	Visit 3
Time relative to dosing (incl window)	Within 30 d prior Visit 1	NA	1 d post-dose (±0 d)	3 d post-dose (+1 d)	7 d post-dose (+5 d) ^m
Vaginal swab sampling for local PK levels of BNT331 active substance ^j		X (pre-and post-dose)	X		X
Vaginal swab sampling for ADA response (investigator)		X (pre-dose)			X
Only in subjects in 5 mg and 20 mg DL cohort: Vaginal swab sampling for microbiome assessment ^q		X (pre-dose)			X
Vaginal swab sampling for Nugent score	X ^o	X ^o (pre-dose)			X
Local tolerability assessment (investigator) ^k		X			
Local tolerability interview ^l				X	
Record any AEs and SAEs since last visit	Start	=>	=>	=>	End
Record pregnancies (until 60 d after Visit 3)	Start	=>	=>	=>	=>
Total blood volume collected per visit	~15 mL	~45 mL	~5 mL		~25 mL
<p>a. Routine urinalysis by centrally provided dipstick.</p> <p>b. Rapid and highly sensitive urine pregnancy tests for βhCG will be performed using kits provided by sponsor and all subjects will be counseled about the need for highly effective methods of contraception (see CTP Section 10.5.2).</p> <p>c. Complete examination at Visit 0 and a brief (symptom-directed) physical examination at the Follow-up Visit (see CTP Section 8.4.1).</p> <p>d. Pelvic examination (see CTP Section 8.4.8) after 4 h and 24 h, also checking for dissolution of vaginal insert.</p> <p>e. Subjects need to be negative (0/4) for Amsel's criteria (see CTP Section 8.3.2).</p> <p>f. Vital signs: systolic/diastolic blood pressure, pulse rate, and oral body temperature. Height and body weight only at Screening (see CTP Section 8.4.2).</p> <p>g. Clinical laboratory (see CTP Section 10.3).</p> <p>h. Before study treatment, at 1 h (±15 minutes), 2 h (±15 minutes), 4 h (±15 minutes), 6 h (±15 minutes), 24 h (±2 h) after dosing, and at the Follow-up Visit.</p> <p>i. The timing of this visit should be arranged to allow the +24 h (±2 h) PK sampling.</p> <p>j. Before study treatment, at 4 h (±15 minutes), 24 h (±2 h) after dosing, and at the Follow-up Visit.</p>					

Description	Screening	Dosing	End of therapy	Telephone interview	Follow-up
Visit ID	Visit 0	Visit 1	Visit 2 ⁱ	(No visit)	Visit 3
Time relative to dosing (incl window)	Within 30 d prior Visit 1	NA	1 d post-dose (±0 d)	3 d post-dose (+1 d)	7 d post-dose (+5 d) ^m
<p>k. Examination of cervix and assessing signs of erythema, oedema, and excoriation on the vaginal mucosa. To be assessed 4 h (±15 minutes) post-dosing.</p> <p>l. A study site staff will call the subjects and ask about any complaints of dysuria, vulvo-vaginal symptoms such as redness, itching, pain, smell/odor, discharge, and/or bleeding. If during the telephone call, a subject voices genital symptoms, an appointment will be made to see and examine the subject.</p> <p>m. Investigations should ideally be performed on 7 d post-dose, while the +5 d time window should only be used in case of specific events such as unexpected onset of menstruation.</p> <p>n. If the screening blood sample was taken >3 d before, a new sample will be taken on the day of dosing, results may be pending at start.</p> <p>o. Normal Nugent score should be ascertained once at early Screening (up to 30 d prior). A new sample will be collected again on the day of dosing, corresponding to the microbiome baseline sample for central analysis (Section 8.3.3).</p> <p>p. Directly after dosing, subjects need to remain in a lying position for 6 h to allow serial sampling. Interruptions, such as short walks to the procedure room or the bathroom should be documented as non-important protocol deviations.</p> <p>q. Microbiome assessment includes both whole genome sequencing and qPCR.</p> <p>Note: The total volume of blood drawn from each study subject in this study will be up to ~90 mL.</p>					
<p>ADA = Anti-drug antibody. AE = Adverse event. βhCG = Beta human chorionic gonadotropin. CTP = Clinical trial protocol. d = day(s). DL = Dose level. h = hour(s). NA = Not applicable. PK = Pharmacokinetics. qPCR = quantitative polymerase chain reaction. SAE = Serious adverse event. WOCBP = Women of childbearing potential.</p>					

Table 5: Schedule of Activities – Part B

Description	Screening	Randomization	Treatment Phase	Early Response (End of Therapy)	Call	ToC	Follow-up/Recurrence ⁿ or Early termination	Late Follow-up
Visit ID	Visit 0	Visit 1	No visit	Visit 2	No visit	Visit 3 ^l	Visit 4	Visit 5
Day no.	-2 to 1 ^o	-2 to 1 ^p	1 to 5	6 ± 0 ^p	8 to 14	21 to 30	60 ± 7	120 ± 7
Informed consent	X							
Inclusion/exclusion criteria	X	X						
Urinalysis ^a	X			X		X		
Pregnancy test and counseling for WOCBP ^b	X	X						X
Medical history incl. BV and menstrual cycle	X	X (update)						
Demography	X							
Prior and concomitant medications	X	X		X	X	X	X	X
Physical examination ^c	X	X		X		X		
Pelvic examination ^d	X	X		X		X	X	X
Amsel's criteria ^e	X			X		X	X	X
Vaginal swab sampling for Nugent score ^f	X	X		X		X	X	X
Vital signs ^g	X	X		X		X		
Blood draw for clinical laboratory ^h (15 mL)	X			X		X		
Trichomonas rapid test	X							
Vaginal swabs for microbiome analysis		X		X		X	X	X

Description	Screening	Randomization	Treatment Phase	Early Response (End of Therapy)	Call	ToC	Follow-up/ Recurrence ⁿ or Early termination	Late Follow-up
Visit ID	Visit 0	Visit 1	No visit	Visit 2	No visit	Visit 3 ^l	Visit 4	Visit 5
Day no.	-2 to 1 ^o	-2 to 1 ^p	1 to 5	6 ± 0 ^p	8 to 14	21 to 30	60 ± 7	120 ± 7
Treatment instructions and dispensing		X						
Return unused study treatment				X				
Blood sampling for ADA response in serum (5 mL)		X		X		X	X	
Vaginal swab sampling for ADA response		X				X		
Administer study treatmentⁱ			X-X-X-X-X					
Blood sampling for BNT331 active substance concentration (PK) in serum ^j (5 mL)		X		X		X		
Vaginal swab sampling for local PK levels of BNT331 active substance ^j		X		X		X		
PROM scores ^m		X				X	X	X
Issue/conclude subject e-diaries ^k		X				X		
Period of daily completion and review of subject e-diaries			Start at Day 1	=>	=>	End		
Record AEs and SAEs since last visit	Start	=>	=>	=>	=>	=>	=>	End
Record pregnancies (until 60 d after Visit 3)	Start	=>	=>	=>	=>	=>	=>	=>

Description	Screening	Randomization	Treatment Phase	Early Response (End of Therapy)	Call	ToC	Follow-up/Recurrence ⁿ or Early termination	Late Follow-up
Visit ID	Visit 0	Visit 1	No visit	Visit 2	No visit	Visit 3 ^l	Visit 4	Visit 5
Day no.	-2 to 1 ^o	-2 to 1 ^p	1 to 5	6 ± 0 ^p	8 to 14	21 to 30	60 ± 7	120 ± 7
Total blood volume collected per visit	~25 mL			~25 mL		~25 mL	~5 mL	

- a. Routine urinalysis by centrally provided dipstick.
- b. Rapid and highly sensitive urine pregnancy tests for β hCG will be performed using kits provided by sponsor and all subjects will be counseled about the need for highly effective method of contraception (see CTP Section 10.5.2).
- c. Complete examination at Visit 0, and a brief (symptom-directed) physical examination at all later visits (see CTP Section 8.4.1).
- d. Pelvic examination (see CTP Section 8.4.8), on Day 6 (Visit 2) also checking for dissolution of vaginal insert.
- e. Amsel's criteria (see CTP Section 8.3.2).
- f. Nugent scoring system (see CTP Section 8.3.3), results may be pending at randomization/dosing.
- g. Vital signs: systolic/diastolic blood pressure, pulse rate, and oral body temperature. Height and body weight only at Screening (see CTP Section 8.4.2).
- h. Clinical laboratory tests (see CTP Section 8.4.4), results may be pending at randomization/dosing.
- i. Following Visit 1 (Day 1 until Day 5), treatment will be self-administered by the subject at home, at night before going to bed.
- j. Before study treatment and at 8-24 h after the last dose (at Early Response and ToC Visit).
- k. Daily recording of treatment adherence (until Day 5), local tolerability/discomfort (until ToC Visit: dysuria, vulvo-vaginal symptoms such as redness, itching, pain, smell/odor, discharge, and/or bleeding).
- l. After the end of the next menstruation, if applicable.
- m. PROM scores, SF-36 and CSFQ-F (see CTP Section 8.4.6).
- n. In case of early termination or BV recurrence at any time after ToC Visit and before the Late Follow-up Visit, the Follow-up/Recurrence Visit assessments should also be performed prior to administration of standard of care treatment and recorded as unscheduled visits, respectively.
- o. Screening and Randomization visit can occur on the same day. In this case, overlapping assessments and procedures as per SoA, may be performed only once.
- p. The first dose of IMP will be inserted at home at bedtime on the day of randomization. Randomization and first dose would be on Day 1. Exceptionally, dosing can be started on one of the next 2 days (within 72 h of screening) at investigators discretion if that allows Visit 2 to fall on a Monday. Randomization would be on Day -1, or Day -2 respectively.
- q. Microbiome assessment includes both whole genome sequencing and qPCR.

Note:

Description	Screening	Randomization	Treatment Phase	Early Response (End of Therapy)	Call	ToC	Follow-up/ Recurrence ⁿ or Early termination	Late Follow-up
Visit ID	Visit 0	Visit 1	No visit	Visit 2	No visit	Visit 3 ^l	Visit 4	Visit 5
Day no.	-2 to 1 ^o	-2 to 1 ^p	1 to 5	6 ± 0 ^p	8 to 14	21 to 30	60 ± 7	120 ± 7

The total volume of blood drawn from each study subject in this study will be up to ~80 mL.

ADA = Anti-drug antibody; AE = Adverse event; βhCG = Beta human chorionic gonadotropin; BV = Bacterial vaginosis; CSFQ-F = Changes in Sexual Functioning Questionnaire; CTP = Clinical trial protocol; d = day(s); PROM = Patient-reported outcome measure; qPCR = quantitative polymerase chain reaction; SAE = Serious adverse event; SF-36 = Thirty-six-item short form health survey; SoA = Schedule of activities; ToC = Test of Cure; WOCBP = Women of childbearing potential.

Appendix 2 Date Imputation Rules

Missing or Partial Date Variable	Imputation Rule
First dose of study treatment date	No imputation will be performed.
Last dose of study treatment date	<p>If the last date of study treatment is completely missing and there is no End of Treatment eCRF page and no death date, the subject should be ongoing and use the cutoff date for the analysis as the last dosing date. Note: Confirm that the subject is actively receiving dose at the time of the data cutoff.</p> <p>If the last date of study treatment is completely or partially missing and there is EITHER an End of Treatment eCRF page OR a death date available (within the data cutoff date), then use this date as the last dose date:</p> <p>If only Year is available and Year < Year of min (EOT date, death date), then use 31DECYYYY,</p> <p>If both year and month are available and Year = Year of min (EOT date, death date) and Month < Month of min (EOT date, death date), then use = last day of the month,</p> <p>For all other conditions, use min (EOT date, death date).</p>
Death date	<p>If the day is missing but month and year are available, the death date will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the 1st day of the month and year of death.</p> <p>If both the month and day are missing, the death date will be imputed as the maximum of the full (non-imputed) day after the date of last contact and January 1st of the year of death.</p> <p>If the entire date is missing, the death date will be imputed as the day after the date of last contact.</p>
Adverse event start date	<p>If the day is missing but month and year are available, the AE start date will be imputed using the first day of the month, except when one of the following conditions apply:</p> <ul style="list-style-type: none"> If this date is earlier than the first dose date, then the first dose date will be used instead; If this date is later than the stop date (possibly imputed), then the stop date will be used instead. <p>If day and month are missing, the AE start date will be imputed using June 15, except when one of the following conditions apply:</p> <ul style="list-style-type: none"> If this date is earlier than the first dose date, then the first dose date will be used instead; If this date is later than the stop date (possibly imputed), then the stop date will be used instead. <p>If the start date of an AE is completely missing, then it is imputed with the first dose date.</p>
Adverse event stop date	<p>If the day is missing but month and year are available, the AE stop date will be imputed using the last day of the month.</p> <p>If day and month are missing, the AE stop date will be imputed using December 31.</p> <p>EXCEPTION: If the imputed date is after the date of death (possibly imputed), then the date of death will be used instead.</p>
Medication/therapy start date	If the day is missing but month and year are available, the medication/ therapy start date day will be imputed using:

Missing or Partial Date Variable	Imputation Rule
	<ul style="list-style-type: none"> 15 for the day if <ul style="list-style-type: none"> the start date year is before the study treatment start date year the start date year is equal to the study treatment start date year, but the month is before the study treatment start date month 1 for the day if <ul style="list-style-type: none"> the start date year is after the study treatment start date year. the start date year is equal to the study treatment start date year, but the month is after the study treatment start date month 1 day prior to study treatment start date if the start date month and year are equal to the study treatment start date month and year <p>If day and month are missing, the medication/therapy start date will be imputed using:</p> <ul style="list-style-type: none"> July 1 if the start date year is before the study treatment start date year; January 1 if the start date year is after the study treatment start date year. 1 day prior to study treatment start date if the start date year is equal to the study treatment start date year. <p>If the start date year is missing, then the date is imputed as 1 day prior to study treatment start date.</p> <p>EXCEPTION: If complete end date is available and the imputed start date is greater than the end date, then the assumed start date should be set to the end date.</p>
Medication/therapy end date	<p>If the day is missing but month and year are available, the medication/ therapy end date day will be imputed using the minimum of study treatment end date and the last day of the month:</p> <p>If day and month are missing, the medication/therapy end date day and month will be imputed using is the minimum of study treatment end date and the end of the year (31DECYYYY).</p> <p>If the end date year is missing, then the end date will not be imputed, but for classification into prior or concomitant, the medication will be considered as {prior/concomitant}.</p> <p>EXCEPTION: If the imputed end date is less than the start date, then the end date should be set to the start date.</p>

Appendix 3 Health Reported Outcome Surveys and Questionnaires

Thirty-six-item Short Form Health Survey (SF-36)

The SF-36 is a short form health survey consisting of 36 questions ([Ware and Sherbourne 1992](#)), yielding the following eight health-related quality of life subscale domains:

- Physical functioning.
- Role-physical.
- Bodily pain.
- General health.
- Vitality.
- Social function.
- Role-emotional.
- Mental health.

as well as a psychometrically-based physical component score (PCS) and a mental component score (MCS). Scoring of the SF-36 will be conducted by a third-party vendor and the eight subscale domains plus two overall summary scores (PCS and MCS) will be derived by the vendor. A higher score represents better quality of life.

Sexual Functioning Questionnaire – Female Version (CSFQ-F)

The CSFQ-F ([Clayton et al. 1997](#)) is a self-administered or interviewer-assisted questionnaire designed to measure illness and medication-related changes in female genital diseases and pregnancy complications. In this study the questionnaire will be self-administered in the local language.

The questionnaire produces scores for five aspects/domains of sexual functioning, in addition to providing a composite total score. The five domains are: pleasure, desire/frequency, desire/interest, arousal/excitement, and orgasm/completion.

The analysis of the domains and composite total score are as follows:

Aspect(s)/Domain(s)	Item(s)
Pleasure	1
Desire/Frequency	2 + 3
Desire/Interest	4 + 5 + 6
Arousal/Excitement	7 + 8 + 9
Orgasm/Completion	11 + 12 + 13
Total	Sum (1 to 14)

Note: Item 14 is only included as part of the total; it is not included as a sub-category.

A higher composite total score indicates greater sexual function.