NOVARTIS

Clinical Development

AAA405

Clinical Trial Protocol CAAA405A12301 / NCT04838613

Phase III study for evaluation of the diagnostic performance of [^{18}F]CTT1057 PET imaging in patients with prostate cancer with rising PSA levels [biochemical recurrence (BCR)] (GuidePath)

Statistical Analysis Plan (SAP)

Document type:	SAP Documentation
Document status:	Final
Release date:	26-Dec-2023
Number of pages:	56

Template Version 6.0, Effective from 23-Nov-2022

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
10-Oct-2021	After FPFV	Creation of final version	N/A - First version	NA
03-Oct-2022	After PA01, before original IND submission	Incorporating changes made in the protocol amendment v01.	1. Clarification added to the targeted population, i.e. prostate cancer patients diagnosed with biochemical recurrence (BCR) after initial definitive therapy (with either RP or curative intent RT). 2. Clarification of the different CTS levels will be applied hierarchically at lesion level for efficacy analyses added. 3. Addition of a new option named "Radiation alone with change in radiation treatment plan" for questionnaire 2. 4. Clarification added that the secondary endpoint (based on central imaging reads) and [redacted] for concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of PSMA-positive lesions (in terms of number of lesions detected and their location) will be assessed at lesion level, not patient level. 5. Updated the definition of efficacy analysis set to align with PA01. 6. Updated the 2 subgroups for subgroup analyses to patients with prior RP and patients with prior curative intent RT. 7. Modified analyses to be performed regarding patient	Section 1.1, 1.2, 2.1.1, 2.2, 2.3.1, 2.5.6, 2.6, 2.7, 2.12, 3 and Section 5.

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			disposition output to fully align with SC meeting. 8. Added details regarding a supplementary analysis for the primary endpoints. 9. Removed the sentence about aggregating results among readers for the intra-reader variability calculation. 10. Added COVID-19 related AE summaries and two required tables by ClinicalTrials.gov. 11. Added 3-day time window for post-baseline lab and vital sign assessments. [redacted] 13. Minor typographical corrections, removal of duplicated wording and minor rewording implemented across the document.	
20-Dec-2023	After CSR dry-run review and before DBL.	Incorporating changes identified during the previous SC meeting, CSR TFLs amendment review, and dry-run review.	1. Added clarification to include [^{18}F]CTT1057 ([^{68}Ga]Ga-PSMA-11) treatment-related AEs that occur beyond 14-days after the injection of [^{18}F]CTT1057 ([^{68}Ga]Ga-PSMA-11) as [^{18}F]CTT1057 ([^{68}Ga]Ga-PSMA-11) TEAE as well and moved such definition to Section 2.7.1. Adverse events. 2. Added a listing to display the information about analysis sets for each participant. [redacted] 4. Updated the patient disposition section to include analysis set (if different from FAS), and to clarify that treatment discontinuation (with reasons) and treatment completion	Section 1.2.1, 2.1.1, 2.2, 2.2.1, 2.3.1, 2.3.2, 2.4.1, 2.4.2, 2.5.2, 2.5.5, 2.5.6, 2.6, 2.6.1, 2.6.2, 2.7, 2.7.1, 2.7.2, 2.7.3, 2.7.4, 2.12, 4, 5.1, 5.4.1 and 5.5.

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			will be displayed by period 1/2 and by assigned treatment sequence. 5. Added grouping information when tabulating primary tumor clinical stage, biopsy Gleason score, radical prostatectomy Gleason score (if applicable) per comments received from SC meetings. 6. Description added to include an additional summary table of time (in days) between two PET scan dates. 7. Added clarification to the definition of concomitant medications/therapies/procedures for each PET imaging agent. 8. Added description that patient-level PPV will be repeated by CTS levels. 9. Added the weighted estimator approach for region-level CLR under the sensitivity analyses. 10. Instructions added regarding how to display the output for covariate-adjusted logistic (random-effects) model for the co-primary endpoints. 11. Added clarification about the analysis set for secondary endpoints, wherever applicable. 12. Added description about displaying shift table about participants' treatment plans. 13. Added detailed instruction on the computation of concordance between two PET imaging agents on lesion-level, overall, and by region and M1 regions.	

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			14. Removed duplicated information about pre-/on-/post-treatment periods under the safety analyses section. 15. Changed the name from AESI to STI to align with RLT program standard. 16. Added clarification that on-treatment death will be tabulated for F-SAF and Ga-SAF separately, while all deaths analyses will be carried out for SAF only. 17. Added clarification to data that will be analyzed for labs and VS. [redacted] [redacted] 20. Added clarification that subgroup analyses will be applied to co-primary endpoints and secondary efficacy endpoints only. 21. Added minor update to the imputation rule when imputed end date is < start date, the end date will be set to start date. 21. Added modelling and programming details of logistic random effects model for the region-level CLR to the appendix.	

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
			22. Minor updates to the estimand wording were made as compared to the protocol for clarification.  25. Replaced "subject" and "patient" with "participant" throughout the document, wherever applicable and appropriate. 26. Updated SAP template version. 27. Minor typographical corrections, removal of duplicated wording and minor rewording implemented across the document.	

Table of contents

List of figures

List of tables

List of abbreviations

AE	Adverse Event
CRF	Case Report Form
CSR	Clinical Study Report
DMS	Document Management System
FAS	Full Analysis Set
F-SAF	[^{18}F]CTT1057 Safety Set
FDG	Fluorodeoxyglucose
FN	False Negative
FP	False Positive
Ga-SAF	[^{68}Ga]Ga-PSMA-11 Safety Set
HR	Heart Rate
IA	Interim Analyses
ICF	Informed Consent Form
MBq	Mega-Becquerel
mCi	Millicurie
mCRPC	metastatic-Castration Resistant Prostate Cancer
mg	milligram(s)
min	minute(s)
mL	milliliter(s)
MRI	Magnetic Resonance Imaging
MedDRA	Medical Dictionary for Drug Regulatory Affairs
p.i.	post-injection
PCa	Prostate Cancer
PCWG3	Prostate Cancer Working Group 3
PET	Positron Emission Tomography
PLN	Pelvic Lymph Node
PPA	Positive percent agreement
PPV	Positive Predictive Value
PSA	Prostate Specific Antigen
PSMA	Prostate Specific Membrane Antigen
pTNM	pathological stage
PK	Pharmacokinetics
PRO	Patient-reported Outcomes
RAP	Reporting & Analysis Process
RECIST	Response Evaluation Criteria In Solid Tumors
RP	Radical Prostatectomy
RR	Respiratory Rate
RT	Radiation Therapy
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SBP	Systolic Blood Pressure

SC	Steering Committee
SoC	Standard of Care
SOP	Standard Operating Procedure
SoT	Standard of Truth
STI	Safety Topic of Interest
SUSAR	Suspected Unexpected Serious Adverse Reaction
[redacted]	[redacted]
[redacted]	[redacted]
TEAE	Treatment Emergent Adverse Event
TN	True Negative
TP	True Positive
TFLs	Tables, Figures, Listings
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the Clinical Study Report (CSR) of study CAAA405A12301, a Phase III study for evaluation of the diagnostic performance of [^{18}F]CTT1057 PET imaging in patients with prostate cancer with rising PSA levels.

The content of this SAP is based on protocol CAAA405A12301 version 01. All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock of the study data.

1.1 Study design

This is a prospective, open-label, multi center, single-arm Phase III study to evaluate the diagnostic performance of [^{18}F]CTT1057 as a PET imaging agent for detection and localization of PSMA positive tumors in prostate cancer patients diagnosed with biochemical recurrence (BCR) after initial definitive therapy with either radical prostatectomy (RP) or curative intent radiation therapy (RT), using a composite truth standard (CTS) as reference.

Approximately 190 participants will be enrolled to ensure at least 152 participants are evaluable for the co-primary endpoints. [^{18}F]CTT1057 PET/CT scan imaging will be centrally read by 3 independent nuclear medicine physicians who will be blinded to any other participant data. See study protocol Section 4.1 and Section 8.3.1 for central read details.

The CTS to be used as reference will be hierarchical in nature, with 3 levels of Standard of Truth (SoT) procedures, that will be applied as follows:

Level 1

- **Histopathology** if available for a lesion (from prospective biopsy or salvage surgery performed within 8 weeks after the [^{18}F]CTT1057 PET/CT scan); OR in case that histopathology is not available for a lesion, inconclusive (i.e. if biopsy/surgery tissue sample is not evaluable for histopathology assessment due to sample quality or inadequate quantity or any other technical causes) or negative (for biopsy only):

Level 2

- **Imaging** diagnostic procedures performed on each participant as clinically indicated per SoC, which must include the [^{68}Ga]Ga-PSMA-11 PET/CT scan that is mandatory for all participants randomized in the study and at least high resolution CT scan (covering at least chest, abdomen, and pelvis and any other region as clinically indicated) with intravenous and oral contrast (or MRI if CT scan with contrast is medically contraindicated). The high resolution CT scan and any other imaging diagnostic procedure clinically indicated as per SoC must be acquired within the 8 weeks either prior to or following the [^{18}F]CTT1057 PET/CT scan. Three-month follow-up imaging (from baseline) will also be used as part of the CTS Level 2 in cases where it is clinically required for the diagnosis of particular lesion(s); OR if neither of the two above are feasible or deemed appropriate or they are inconclusive (i.e. images are not interpretable due to poor or inadequate image quality that precludes reliable assessment):

Level 3

- 50% or greater decline in PSA following radiation therapy (as long as no concomitant androgen deprivation therapy (ADT) is given) as per PCWG3 criteria (Scher et al 2016).

In the cases where pathology will be available (CTS Level 1), assessments by the local pathologists will be performed as per SoC, and results should be available within 2 weeks after surgery. Pathologists will be blinded to the [^{18}F]CTT1057 PET/CT data. In this case, only pathology will be used as SoT for the concerned lesion(s).

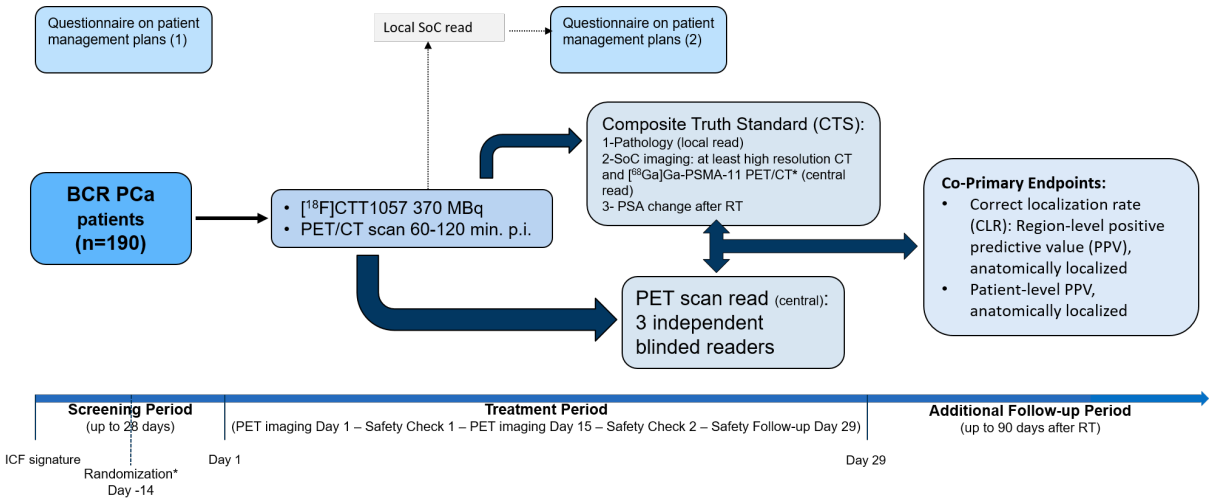
In the cases where pathology will not be available for a lesion, the central read results of the [^{68}Ga]Ga-PSMA-11 PET/CT scan (performed for all participants randomized) and other imaging diagnostic procedures to be performed on each participant for the CTS Level 2 will be used as SoT for the co-primary endpoint calculations. See study protocol Section 4.1 for central read details.

All participants will undergo a [^{68}Ga]Ga-PSMA-11 PET/CT as part of the study, for both the CTS Level 2 (in case it is required as SoT) and the secondary endpoint of assessment of concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of lesions (as detailed in the study protocol Section 2).

In addition to central review, local review of SoC images (including [^{68}Ga]Ga-PSMA-11) will be performed to be used by the treating physician/Clinical Study Investigator for participant management decisions and overall assessment. See study protocol Section 8.3.1 for assessment details and Section 6.4 for blinding requirements.

A questionnaire on the planned participant management will be filled in by the treating physician/Clinical Study Investigator before (questionnaire 1) and within 14 days after (questionnaire 2) knowing the results of the [^{18}F]CTT1057 PET/CT scan. A local review of [^{18}F]CTT1057 PET/CT images will also be performed by a local nuclear medicine physician or radiologist with expertise in reading oncology PET/CT scans and the results will be provided to the treating physician/Clinical Study Investigator for completion of questionnaire 2. Options will be given in the questionnaire to capture possible management plan, such as Surgery, Radiation alone, Radiation alone with change in radiation treatment plan (only applicable for questionnaire 2), Radiation plus ADT, ADT alone, Observation/surveillance, Other (free text box). Any change in patient management plan between the questionnaire 1 and questionnaire 2 should not be based only on [^{18}F]CTT1057 PET scan results since this is an investigational diagnostic imaging product. Other diagnostic procedures should be performed as per SoC in order to confirm and implement the changed management plan. See Figure 1-1 for study design schema.

Figure 1-1 Study Design



Screening period

Written informed consent form (ICF) must be obtained prior to any screening procedures. The participant must be registered in the Interactive Response Technology (IRT) for screening. All procedures described in the Assessment Schedule as per Table 8-1 in the study protocol must be carried out, prioritizing laboratory assessments to allow time to obtain the results at least 14 days prior the planned first PET imaging day (Day 1). Eligibility must then be confirmed at the latest on Day -14. The screening period should last up to 28 days.

Once eligibility is confirmed, the participants will be randomized in IRT to be assigned to one of the following two PET/CT scan sequences at random in a 1:1 ratio:

- Sequence 1: [^{18}F]CTT1057 on Day 1 (investigational imaging agent) followed by [^{68}Ga]Ga-PSMA-11 at least 14 days apart (as part of CTS if required, and for secondary endpoint)
- Sequence 2: [^{68}Ga]Ga-PSMA-11 (as part of CTS if required, and for secondary endpoint) on Day 1 followed by [^{18}F]CTT1057 (investigational imaging agent) at least 14 days apart

PET Imaging days

The 2 PET imaging procedures will be done at least 14 days apart. The day of the first PET imaging agent injection will be considered study Day 1. Please refer to the study protocol Section 8.3.1 for the steps that will take place on each PET imaging day.

Both visits will also be considered as the End of Treatment (EOT) visits.

Safety Checks

All treated participants will be contacted by phone within 24 to 72 hours following each PET scan in order to capture potentially occurring Adverse Events.

Safety Follow-up

Participants will come back to the hospital 14 days (+ 3 day window) after each PET scan day, for a safety visit including vital signs, lab and urine analysis. For participants undergoing surgery, radiation therapy or initiating any antineoplastic medication (provided all CTS procedures have been completed before this initiation) prior to planned safety visit, these safety follow-up assessments will need to be performed before the surgery/radiation therapy/initiation of antineoplastic medication.

Additional Follow-up

Additional follow up period will be applicable where additional diagnostic procedures are to be performed. These may include SoC pathology, imaging or other modality disease assessments (CTS Levels 1 or 2) up to 8 weeks after [^{18}F]CTT1057 PET scan or three-month follow-up imaging (from baseline) where it is clinically required for the diagnosis of particular lesion(s).

In addition, in cases where the CTS Level 2 is not feasible/appropriate, the CTS Level 3 will apply and will include the assessment of PSA levels prior to starting radiotherapy (which should commence within 8 weeks of completing [^{18}F]CTT1057 PET scan) and at approximately 90 days from last day of radiation.

PSA response per PCWG3 is defined as a 50% or greater decline in PSA on the post-RT timepoint compared to PSA prior to RT. If a participant has PSA response, in the absence of concomitant androgen deprivation therapy administration, then the irradiated lesion (recorded in the concomitant antineoplastic – radiotherapy CRF page) will be considered as a confirmed prostate cancer lesion.

1.2 Study objectives, endpoints and estimands

Table 1-1 Objectives and related endpoints

Objective(s)	Endpoint(s)
Primary objective(s)	**Endpoint(s) for primary objective(s)**
• Evaluate the **region-level correct localization rate** (CLR) of [^{18}F]CTT1057	• Proportion of regions containing at least one TP lesion (anatomically localized correspondence between PET imaging and the reference standard), regardless of any co-existent FP findings within the same region, out of all regions containing at least one PET-positive finding by central assessments. See Section 1.2.1 for Primary Estimand
• Evaluate the **patient-level positive predictive value** (with anatomical localization) of [^{18}F]CTT1057	• Proportion of patients who have at least one TP lesion (anatomically localized correspondence between PET imaging and the reference standard), regardless of any co-existent FP findings, out of all patients who are PET/CT scan positive by central assessments. See Section 1.2.1 for Primary Estimand
Secondary objective(s)	**Endpoint(s) for secondary objective(s)**
• Evaluate the patient-level **sensitivity** of [^{18}F]CTT1057	• Proportion of patients that test positive on [^{18}F]CTT1057 and CTS (TP) among those that are CTS positive (TP or FN)

Objective(s)	Endpoint(s)
• Evaluate the patient-level **specificity** of [^{18}F]CTT1057	• Proportion of patients that test negative on [^{18}F]CTT1057 and CTS (TN) among those that are CTS negative (TN or FP)
• Evaluate the patient-level **negative predictive value** of [^{18}F]CTT1057	• Proportion of patients who are both [^{18}F]CTT1057 and CTS negative (TN) among those who test negative on [^{18}F]CTT1057 (TN or FN)
• Evaluate the patient-level **accuracy** of [^{18}F]CTT1057	• Proportion of patients that are CTS and [^{18}F]CTT1057 positive (TP) and negative (TN) among those patients that identified on [^{18}F]CTT1057 (TP, TN, FP or FN)
• Evaluate correct detection rate (CDR)	• Proportion of patients who have at least one TP lesion (exactly localized correspondence between PET imaging and the reference standard), regardless of any co-existent FP findings, out of all patients who are scanned
• Evaluate detection rate	• Proportion of patients who have at least one PET positive lesion, regardless of TP or FP findings, out of all patients who are scanned
• Evaluate the region level **sensitivity** of [^{18}F]CTT1057	• Proportion of regions that test positive on both [^{18}F]CTT1057 and CTS (TP) among those regions that are CTS positive (TP or FN)
• Evaluate the region level **specificity** of [^{18}F]CTT1057	• Proportion of regions that test negative on both [^{18}F]CTT1057 and CTS (TN) among those regions that are CTS negative (FP, or TN)
• Evaluate the region level **negative predictive value** of [^{18}F]CTT1057	• Proportion of regions that are CTS and [^{18}F]CTT1057 negative (TN) among those regions that test negative on [^{18}F]CTT1057 (TN or FN)
• Evaluate the region level **accuracy** of [^{18}F]CTT1057	• Proportion of regions that are CTS and [^{18}F]CTT1057 positive (TP) and negative (TN) among those regions that identified on [^{18}F]CTT1057 (TP, TN, FP and FN)
• Evaluate the patient-level positive predictive value related to PSA levels	• Percentage of patients who have at least one TP lesion (exactly anatomically localized correspondence between [^{18}F]CTT1057 PET imaging and the reference standard), regardless of any co-existent FP findings, out of all patients who are [^{18}F]CTT1057 positive, stratified by PSA levels
• Characterize the safety and tolerability of [^{18}F]CTT1057	• Incidence of AEs. Treatment emergent adverse event (TEAE) rate within 14 days after each PET tracer administration
• [^{18}F]CTT1057 scan inter-reader variability	• Inter-reader agreement of [^{18}F]CTT1057 images
• [^{18}F]CTT1057 scan intra-reader variability	• Intra-reader agreement of [^{18}F]CTT1057 images

Objective(s)	Endpoint(s)
• Concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of lesions at lesion level using central reads	• Concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of PSMA-positive lesions (location and number) using central reads
• Evaluate the change in patient management plans attributed to the PET/CT scan	• Percentage of patients who underwent a change in intended treatment plan attributed to the PET/CT scan as assessed by pre and post imaging questionnaires
• Assess all the above primary and secondary objectives independently in the subgroup of patients with prior radical prostatectomy and the subgroup of patients with prior curative intent radiation therapy	• All the above primary and secondary endpoints independently in the subgroup of patients with prior RP and the subgroup of patients with prior curative intent radiation therapy

[redacted]

1.2.1 Primary estimand(s)

The primary clinical question of interest is: What is the probability that a positive [^{18}F]CTT1057 PET/CT scan truly detects and localizes (PSMA-expressing) tumor recurrence in BCR PCa patients?

The justification for the primary estimands is that we wish to assess the diagnostic performance of [^{18}F]CTT1057 PET/CT in detecting and localizing PSMA positivity in PCa patients diagnosed of biochemical recurrence (BCR), which represents a stage of prostate cancer where there is an increase in PSA levels after initial treatment. Correct identification of disease location and extent guides treatment decisions. It is relevant, therefore, to detect smaller lesions as early as possible. Sensitivity and specificity of [^{18}F]CTT1057 PET/CT in detecting and localizing PSMA positivity using histopathology as SoT will be assessed in an independent study in newly diagnosed high-risk PCa patients candidate to surgery. However, in this primary staging population a low prevalence and high variability of metastatic lymph nodes has been reported (from 4-58%) (Petersen and Zacho 2020). Therefore, an additional study in the BCR setting is needed, since this is a relevant patient population which is expected to have an adequate representation of metastasis-positive and metastasis-negative patients by any available reference standard. In this population, a CTS is needed for reference (for CTS details see protocol Section 3 and Section 4.1**)**. Pathology is available in only a limited number of cases, and morphological imaging is negative in a substantial number of cases. This makes the identification of true and false negative results a challenge, thus complicating the evaluation of sensitivity and specificity. Therefore, the region-level **correct localization rate (CLR)** defined as the region-level PPV with anatomical localization and the patient-level PPV with anatomical localization are more appropriate co-primary endpoints than sensitivity and specificity in the BCR setting. The patient-level PPV with anatomical localization, as the conditional probability of positivity given the result of the PET/CT scan is more meaningful than the correct detection rate (CDR) defined as the percentage of true-positive patients in the whole population (unconditional probability of positivity). Moreover, the CDR has reported to be dependent of the PSA levels in the BCR population using another [^{18}F]-PSMA labeled PET agent (Chantadisai et al 2020).

The primary estimands are described by the following attributes:

Primary estimand 1:

1. Population: Patients with biopsy proven prostate adenocarcinoma and rising PSA after definitive therapy with RP or curative-intent radiation therapy (external beam or brachytherapy), diagnosed of biochemically recurrent prostate cancer. Further details about the population are provided in the protocol Section 5.
2. Variable: Regional-level correct localization rate (CLR), defined as the proportion of regions containing at least one true positive lesion (anatomically localized correspondence between [^{18}F]CTT1057 PET imaging and the reference standard), regardless of any co-existent false positive findings within the same region, out of all regions containing at least one [^{18}F]CTT1057 PET-positive finding, using central reads.
3. Treatment of interest: [^{18}F]CTT1057 injected as a single intravenous dose of approximately 370 MBq and subsequent PET/CT scan. Further details about the investigational treatment is provided in the protocol Section 6.
4. Intercurrent events: Participants who received the investigational imaging agent [^{18}F]CTT1057 but did not undergo/complete the PET scan for any reasons (e.g. consent withdrawal, PET camera failures, etc.). Details on how to handle intercurrent events are provided in Section 2.5.3.
5. Summary measure: Region-level CLR along with two-sided 95% confidence interval estimated using logistic random-effects models taking into account the correlation among regions within one participant.

Primary estimand 2:

1. Population: the same as primary estimand 1
2. Variable: Patient-level positive predictive value (PPV), defined as the proportion of participants who have at least one true positive lesion (anatomically localized correspondence between [^{18}F]CTT1057 PET imaging and the reference standard), regardless of any co-existent false positive findings, out of all PET positive participants, using central reads.
3. Treatment of interest: the same as primary estimand 1
4. Intercurrent events: the same as primary estimand 1
5. Summary measure: Patient-level PPV along with two-sided exact binomial 95% confidence interval.

2 Statistical methods

2.1 Data analysis general information

The final analysis will be performed by Novartis. SAS version 9.4 or later will be used to perform all data analyses and to generate tables, figures and listings.

Data included in the analysis

Only one final analysis is planned for the co-primary efficacy endpoints. All statistical analyses will be performed using all data collected in the database up to the data cut-off date. The cut-off date for the final analysis of study data will be established after all randomized participants have completed EOS visit (Last Patient Last Visit).

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. No center effect will be assessed.

Qualitative data (e.g., race, etc.) will be summarized by tables; a missing category will be included as applicable. Percentages will be calculated using the number of participants in the relevant population or subgroup as the denominator.

Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, 1st and 3rd quartiles, minimum, and maximum).

2.1.1 General definitions

Investigational PET imaging agents: [^{18}F]CTT1057

Study Treatment: the administration of at least one of the two PET imaging agents: [^{18}F]CTT1057, as a single injection intravenously (i.v.) of approximately 370 MBq (range 266 – 407 MBq), and [^{68}Ga]Ga-PSMA-11 as a single injection intravenously (i.v.) of approximately 150 MBq (range 111 – 185 MBq), independently of whether the PET/CT scans were acquired or not. In this study, the investigational imaging agent of interest is the radioligand imaging compound [^{18}F]CTT1057. [^{68}Ga]Ga-PSMA-11 will be used as a component of the CTS Level 2 if required, and in all participants for the secondary endpoint of assessment of concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of lesions (as detailed in the study protocol Section 2).

Date of administration of investigational PET imaging agents

The date of administration of investigational PET imaging agent is defined as the date when a non-zero dose of investigational PET imaging agent is administered and recorded on the study treatment (e)CRF. The date of administration of study treatment will also be referred as start of investigational PET imaging agent.

Study day

The study day describes the day of the event or assessment date, relative to the reference start date.

The study day is defined as:

- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference date for all assessments (safety, efficacy, etc) is the start of study treatment.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Time unit

A year length is defined as 365.25 days. A month length is 30.4375 days (365.25/12). If duration is reported in months, duration in days will be divided by 30.4375. If duration is reported in years, duration in days will be divided by 365.25.

Baseline

For safety evaluations, the last available assessment on or before the date/time of start of study treatment [^{18}F]CTT1057 is defined as "baseline" assessment for [^{18}F]CTT1057 safety.

The last available assessment on or before the date/time of start of study treatment [^{68}Ga]Ga-PSMA-11 is defined as "baseline" assessment for [^{68}Ga]Ga-PSMA-11 safety.

If participants have no value as defined above, the baseline result will be set to missing.

Efficacy evaluations will not include any comparisons between baseline and post-baseline values, and therefore a baseline definition for efficacy evaluations does not apply.

On-treatment assessment/event and observation periods

For adverse event reporting, the overall observation period will be divided into three mutually exclusive segments:

1. ***pre-treatment period***: from day of participant's informed consent to the day/time before administration of study treatment
2. ***on-treatment period***: from date of first administration of study treatment to 14 days after date of actual administration of any study treatment (including the administration date). The on–treatment period will be divided into the F-on-treatment period and Ga-on-treatment period for the respective study treatment of [^{18}F]CTT1057 or [^{68}Ga]Ga-PSMA-11
    - The F-on-treatment period is from the date of [^{18}F]CTT1057 administration to 14 days after the date of [^{18}F]CTT1057 administration.
    - The Ga-on-treatment period is from the date of [^{68}Ga]Ga-PSMA-11 administration to 14 days after the date of [^{68}Ga]Ga-PSMA-11 administration.
3. ***post-treatment period***: starting 15 days after the date of last administration of study treatment.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for death including on-treatment and

post-treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize ***treatment-emergent*** AEs only (see Section 2.7.1).

However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

Windows for multiple assessments

In order to summarize weight, vital signs, and laboratory tests collected over time (including unscheduled visits), the assessments will be time slotted. The following general rule will be applied in creating the assessment windows: If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the earlier of the 2 assessments will be used. If there are multiple assessments on the same date then the worst case value will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

For participants randomized into sequence 1: [^{18}F]CTT1057 - [^{68}Ga]Ga-PSMA-11

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline for [^{18}F]CTT1057	On or before Study Day 1*	≤ Study Day 1, and time prior to [^{18}F]CTT1057 injection
Day 1 for [^{18}F]CTT1057	Study Day 1	Date of [^{18}F]CTT1057 injection
Day 15 for PET for [^{18}F]CTT1057	Study Day 15	Study Day 15 - 22, and time prior to [^{68}Ga]Ga-PSMA-11 injection
Baseline (Day 1) for [^{68}Ga]Ga-PSMA-11	On or before [^{68}Ga]Ga-PSMA-11 injection and after [^{18}F]CTT1057 injection	≥Study Day 15 and ≤ Study Day 22, time prior to [^{68}Ga]Ga-PSMA-11 injection
Day 1 for [^{68}Ga]Ga-PSMA-11	On [^{68}Ga]Ga-PSMA-11 injection	Date of [^{68}Ga]Ga-PSMA-11 injection

Day 15 for [^{68}Ga]Ga-PSMA-11	the date of [^{68}Ga]Ga-PSMA-11 injection + 14 days	Study Day 29- 39

*Study Day 1 = the date of [^{18}F]CTT1057 injection

For participants randomized into sequence 2: [^{68}Ga]Ga-PSMA-11 - [^{18}F]CTT1057

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline for [^{68}Ga]Ga-PSMA-11	On or before Study Day 1*	≤ Study Day 1, and time prior to [^{68}Ga]Ga-PSMA-11 injection
Day 1 for [^{68}Ga]Ga-PSMA-11	Study Day 1	Date of [^{68}Ga]Ga-PSMA-11injection
Day 15 for PET for [^{68}Ga]Ga-PSMA-11	Study Day 15	Study Day 15 - 22, time prior to [^{18}F]CTT1057 injection
Baseline for [^{18}F]CTT1057	On or before [^{18}F]CTT1057 and after [^{68}Ga]Ga-PSMA-11	≥Study Day 15 and ≤ Study Day 22, time prior to [^{18}F]CTT1057 injection
Day 1 for [^{18}F]CTT1057	On [^{18}F]CTT1057 injection	Date of [^{18}F]CTT1057 injection
Day 15 for [^{18}F]CTT1057	the date of [^{18}F]CTT1057 injection + 14 days	Study Day 29-39

*Study Day 1 = the date of [^{68}Ga]Ga-PSMA-11 injection

For vital signs, post-injection/pre-imaging, and post-imaging assessment will be compared to the baseline (pre-injection assessment) in addition to its Day 15 assessment compared to baseline. Post-injection/pre-imaging and post-imaging will be performed on the day of imaging.

2.2 Analysis sets

The Full Analysis Set (FAS) includes all randomized participants.

The Efficacy Analysis Set (EFF) includes all randomized participants who receive the dose of investigational treatment (i.e. [^{18}F]CTT1057), have both an evaluable [^{18}F]CTT1057 PET/CT scan imaging and at least one evaluable CTS assessment, and have not received any prohibited systemic antineoplastic therapy before the completion of PET/CTs and CTS procedures.

The [^{18}F]CTT1057 Safety Set (F-SAF) includes all participants who received the investigational treatment (i.e. [^{18}F]CTT1057).

The [^{68}Ga]Ga-PSMA-11 Safety Set (Ga-SAF) includes all participants who received [^{68}Ga]Ga-PSMA-11. Participants will be analyzed according to the study treatment received.

The **Safety Set (SAF)** includes all participants who received at least one of [^{18}F]CTT1057 or [^{68}Ga]Ga-PSMA-11.

The number (%) of participants in each analysis set will be summarized and information about analysis sets for each participant will be displayed in a listing.

2.2.1 Subgroup of interest

The primary and secondary efficacy endpoints will be summarized by initial definitive therapy received:

- prior radical prostatectomy
- prior curative intent radiation therapy

The primary efficacy co-primary endpoints will also be summarized by PSA level at baseline (from central laboratory):

- PSA $\leq$ 0.5 ng/mL
- 0.5 ng/mL<PSA$\leq$1 ng/mL
- 1 ng/mL<PSA$\leq$2 ng/mL
- 2 ng/mL<PSA$\leq$5 ng/mL
- PSA > 5 ng/mL

2.3 Patient disposition, demographics and other baseline characteristics

2.3.1 Patient disposition

Enrollment by country or subdivision and center will be summarized for all screened participants. The number (%) of screen failure and the reasons for screening failure will be displayed. For participants screened more than once, the data from the last screening visit will be used in the summaries. The following summaries will be provided (with % based on the total number of FAS participants unless otherwise specified):

- Number (%) of participants who were screened (based on the 'Disposition' page and with % based on all screened participants)
- Number (%) of participants who completed the screening phase and were randomized (based on the 'Disposition' page and with % based on all screened participants)

- Number (%) of participants who did not complete screening (were not randomized) with reason (based on the 'Disposition' page and with % based on all screened participants)
    - Number (%) of participants who did not complete screening related to COVID-19 and primary relationships to COVID-19 (based on the 'Disposition' page and with % based on all screened participants)

The following summaries will be displayed by assigned treatment sequence and overall:

- Number (%) of participants who completed treatment period 1 (based on the 'Disposition' page)
- Number (%) of participants who discontinued treatment period 1 with reason (based on the 'Disposition' page)
    - Number (%) of participants who discontinued treatment period 1 related to COVID-19 and primary relationships to COVID-19 (based on the 'Disposition' page)
- Number (%) of participants who completed treatment period 2 (based on the 'Disposition' page)
- Number (%) of participants who discontinued treatment period 2 with reason (based on the 'Disposition' page)
    - Number (%) of participants who discontinued treatment period 2 related to COVID-19 and primary relationships to COVID-19 (based on the 'Disposition' page)
- Number (%) of participants who completed the study (based on the 'Disposition' page)
- Number (%) of participants who discontinued the study and the reason for discontinuation (based on the 'Disposition' page)
    - Number (%) of participants who discontinued the study related to COVID-19 and primary relationships to COVID-19 (based on the 'Disposition' page)

Protocol deviations

The number (%) of participants in the FAS with any confirmed protocol deviation will be tabulated by deviation category (as specified in the study Data Handling Plan). All protocol deviations will be listed. For all important protocol deviations, the relationship to COVID-19 will also be captured.

2.3.2 Demographics and other baseline characteristics

The FAS and EFF will be used for all baseline and demographic summaries unless otherwise specified. Listings will be provided using the FAS.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed. Categorical data (e.g. Age groups: < 65 and ≥ 65 years, race, ethnicity, ECOG performance

status, others as applicable) will be summarized by frequency counts and percentages; the number and percentage of participants with missing data will be provided. Continuous data (e.g. age, weight, height, body mass index, PSA level at screening from central laboratory for subgroup participants who had prior RP and participants who had prior curative intent RT, separately) will be summarized by descriptive statistics (N, mean, median, 1st and 3rd quartile, standard deviation, minimum and maximum).

Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following for all participants: PSA level at time of initial cancer diagnosis (ng/mL), primary tumor clinical stage (i.e. T2c or less, T3, T3a, T3b, T4 and Tx), regional lymph node clinical stage, distant metastasis clinical stage, biopsy predominant histology/cytology, percentage of each histological type/pattern, biopsy Gleason score (i.e. ≤ 6, 7 (3+4), 7 (4+3), 8, and 9 or 10), initial definitive therapy received, time since initial diagnosis to first injection (in months), time from initial diagnosis to biochemical recurrence diagnosis (in months), the time since primary definitive therapy to diagnosis of BCR, time since biochemical recurrence diagnosis to the first injection (in days), PSA level at biochemical recurrence diagnosis (ng/mL). Note: biopsy Gleason score will be displayed along with the score details for each score combination in the listing.

For participants who underwent radical prostatectomy: primary tumor pathological stage, regional lymph node pathological stage, distant metastasis pathological stage, radical prostatectomy Gleason score (i.e. ≤ 6, 7 (3+4), 7 (4+3), 8, and 9 or 10), radical prostatectomy residual disease status. Note: Radical prostatectomy Gleason score will be displayed along with the score details for each score combination in the listing.

For participants who underwent curative intent radiation therapy: Nadir PSA level (ng/mL)

Derivations:

Time since initial diagnosis to the first injection (in months) = (date of the first injection – date of first prostate cancer diagnosis + 1) / 30.4375

Time from initial diagnosis to biochemical recurrence diagnosis (in months) = (biochemical recurrence diagnosis - date of first prostate cancer diagnosis + 1) / 30.4375

Time since biochemical recurrence to the first injection (in days) = date of the first injection – biochemical recurrence + 1

ECOG Performance Status

ECOG performance status at baseline will be presented in a summary table and will be listed.

Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on (e)CRF will be summarized and listed. Separate summaries will be presented for ongoing and historical medical conditions. The summaries will be presented by primary system

organ class (SOC), preferred term (PT). Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

The following exposure information of [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11, as reported in the study treatment page of the eCRF, will be summarized for the F-SAF and Ga-SAF :

- Radioactivity dose administered (MBq)
- Number (%) of participants who were dosed but with extravasation during administration
- Number of participants with dose administered outside the planned range:
  - for [^{18}F]CTT1057 : < 266 MBq or > 407 MBq
  - for [^{68}Ga]Ga-PSMA-11 : < 111 MBq or > 185 MBq
- Time (in minutes) from injection to full-body PET imaging acquisition start time
- Number (%) of participants with additional pelvic PET imaging acquisition, with time from injection to additional pelvic PET imaging acquisition start time if applicable

An additional summary of the time (in days) between the [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 PET scan dates will be provided using the SAF.

Participant level listings of doses administered or not, along with reasons if not administered will be produced using the FAS.

2.4.2 Prior, concomitant and post therapies

Prior anti-cancer therapy

The number and percentage of participants who received any prior anti-neoplastic medications, prior anti-neoplastic radiotherapy or prior anti-neoplastic surgery will be summarized. Prior anti-neoplastic medications will be summarized by total number of regimens, setting (e.g. adjuvant, metastatic, etc.) and also by lowest ATC class and preferred term.

Prior anti-neoplastic radiotherapies will be summarized by method (i.e. external beam vs. internal beam) and setting.

Prior anti-neoplastic surgeries will be summarized by procedure (i.e. radical prostatectomy vs. other) and residual disease (e.g. yes, no, etc.)

Separate listings will be produced for prior anti-neoplastic medications (including regimen number, setting and intent), radiotherapy, and surgery. Anti-neoplastic medications will be coded using the WHO Drug Dictionary (WHO-DD); anti-neoplastic surgery will be coded using MedDRA. Details regarding MedDRA and WHO-DD version will be included in the footnote in the tables/listings.

The above summary tables will be performed using the FAS and EFF and listings will be provided using the FAS.

Concomitant medications

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a participant coinciding with the study treatment period. Concomitant therapy includes medications (other than study drugs)/therapies/procedures starting on or after the start date of study treatment or medications/therapies/procedures starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term. These summaries will include:

1. Medications/therapies/procedures starting within 14 days of [^{18}F]CTT1057 (or [^{68}Ga]Ga-PSMA-11) as long as they are prior to [^{68}Ga]Ga-PSMA-11 (or [^{18}F]CTT1057) dose (if applicable)
2. Medications/therapies/procedures starting prior to start of study treatment and continuing after the start of study treatment.

Number of participants (%) who received diuretics (e.g. Furosemide) as concomitant medication on the day of PET/CT scans will also be summarized.

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of first dose of study treatment or starting more than 14 days after the last date of study treatment will be flagged in the listing. Concomitant medication summary tables will be provided using the F-SAF and Ga-SAF separately, and listings will be provided for the SAF.

2.5 Analysis supporting primary objective(s)

The primary objectives of the study are to evaluate the region-level CLR of [^{18}F]CTT1057, and to evaluate the patient-level PPV (with anatomical localization) of [^{18}F]CTT1057.

2.5.1 Primary endpoint(s)

The co-primary endpoints evaluated for participants in the EFF are the following:

Region-level CLR: Proportion of regions containing at least one TP lesion (anatomically localized correspondence between PET imaging and the reference standard), regardless of any co-existent FP findings within the same region, out of all regions containing at least one PET-positive finding by central assessments. See Section 1.2.1 for Primary Estimand.

Patient-level PPV: Proportion of participants who have at least one TP lesion (anatomically localized correspondence between PET imaging and the reference standard), regardless of any co-existent FP findings, out of all participants who are PET/CT scan positive by central assessments. See Section 1.2.1 for Primary Estimand.

Centralized imaging assessments (CTS Level 2) will be used as reference standard if pathology assessment (CTS Level 1) is not available for a lesion, inconclusive or negative (for biopsy only). [redacted]

2.5.2 Statistical hypothesis, model, and method of analysis

For region-level CLR, the observations (regions positive/negative status) are clustered within a participant. Observations within participants (clusters) are correlated, which may lead to biased estimates of region-level CLR when the correlated nature of the data within each cluster is ignored. More important, an analysis ignoring correlation often yields a misleading small estimated standard error (and 95% CI) for CLR because it erroneously counts all observations as independent observations. Region-level CLR adjusted for correlation among observations within participants can be modeled by using a logistic regression model with random effects (Gatsonis CA 1995, Verbeke G, Molenberghs G 2000).

When calculating CLR, the number of TP regions (y_i), follows the binomial distribution with parameter (n_i, p_i),

$$y_i|b_i \sim \text{Binomial}(n_i, p_i)$$

p_i denotes the probability of TP among all positive PET regions (TP or FP) for participant i, and n_i corresponds to the denominator in the CLR definition, i.e. the number of all regions containing at least one PET-positive finding by central assessments (TP or FP regions) for participant i.

The number of TP and FP regions overall (i.e. across all five regions, which are prostate region, PLN, skeletal, extra-pelvic lymph nodes and visceral) and by region across all participants in the EFF will be presented, along with region-level CLR and two-sided 95% CIs estimated using the following logistic regression model with random effects.

$$\log\left(\frac{p_i}{1-p_i}\right) = \mu + b_i$$

where p_i is the region-level CLR for participant i, μ is the fixed effect of overall mean, b_i is the random effect for participants (clusters). b_i follows a normal distribution: $b_i \sim N(0, \sigma_b^2)$.

The lower bound of the 95% CI should be greater than 0.50 to attain the first co-primary endpoint.

For patient-level PPV, the number of TP and FP participants, as well as the number of positive and negative (and inconclusive, if applicable) [^{18}F]CTT1057 PET/CT scans, will be presented, along with patient-level PPV with two-sided exact binomial 95% CIs (Clopper and Pearson 1934).

The lower bound of the 95% CI for patient-level PPV should be greater than 0.20 attain the second co-primary endpoint.

The same analysis for patient-level PPV will be repeated by CTS Level.

An individual PET reader will be considered successful if he/she meets the predefined thresholds for both co-primary endpoints, and at least two of three readers should be successful for overall study positivity.

2.5.3 Handling of intercurrent events

The primary analyses will account for intercurrent events as explained in the following:

Participants who receive the investigational imaging agent [^{18}F]CTT1057 but do not undergo/complete the PET/CT scan for any reason (e.g. consent withdrawal, PET camera failures, etc.) will not be used for the co-primary endpoint analyses. Note that safety data for these participants will be collected after the administration of the investigational imaging agent and will be summarized.

2.5.4 Handling of missing values not related to intercurrent event

This is a diagnostic study, where the co-primary endpoints are calculated on the basis of a one-time imaging assessment. Missing data will not be imputed.

2.5.5 Sensitivity analyses

For participants where CTS Level 1 or 2 is used as the truth standard for the primary analysis, the next lower level of CTS will be used to verify positivity or negativity for the sensitivity analysis, if available. For instance, if pathology (CTS Level 1) is used for a lesion as the truth standard for the primary estimand, the imaging component of CTS (Level 2) will be used as truth standard for sensitivity analysis for the primary estimand. Similarly, results of CTS Level 2 will be replaced by CTS Level 3 if applicable. However, for those using changes in PSA level (CTS Level 3) as the truth standard for primary estimand, their positivity status cannot be replaced by another component of CTS for sensitivity analysis.

Region-level CLR along with its 95% CI will also be computed based on the weighted estimator approach (Lee and Dubin 1994) for each central reader as following:

$$\hat{p}_w \pm 1.96\hat{\sigma}/\sqrt{n}.$$

$\hat{p}_w$is the point estimator with $\hat{p}_w = \sum_{i=1}^{n}\hat{p}_i\,/n$ and $\hat{p}_i = y_i/n_i$, where y_i and n_i are the same as defined in Section 2.5.2 above and n represents the number of participants with at least one PET-positive finding by central assessments. $\hat{\sigma}$ is the estimated standard deviation of $\hat{p}_i$.

2.5.6 Supplementary analyses

Supplementary analyses for region-level CLR and patient-level PPV will be performed for the following subgroups:

- Initial definitive therapy received:
  - prior radical prostatectomy
  - prior curative intent radiation therapy

- PSA level at baseline (from central laboratory):

- PSA ≤ 0.5 ng/mL
- 0.5 ng/mL<PSA≤1 ng/mL
- 1 ng/mL<PSA≤2 ng/mL
- 2 ng/mL<PSA≤5 ng/mL
- PSA > 5 ng/mL

In addition, estimates of region-level CLR and corresponding 95% CI will be obtained from the logistic random-effects regression model, adjusting for covariates and taking into account the correlation among regions within one participant; while estimates of patient-level PPV and corresponding 95% CI will be obtained from the logistic regression model adjusting for covariates. Candidate covariates include categorized PSA level at baseline (from central laboratory) and the initial definitive therapy received (with either RP or curative-intent RT) as given above. The p-values from the Type III tests for each candidate covariate in the full model will be provided. The region-level CLR and patient-level PPV (and their 95% CIs) for each level of the covariates estimated from the full model including all candidate covariates, as well as the final model including only statistically significant covariates (p-value < 0.05), will be displayed. In the case there are no statistically significant covariates identified from the full model, the region-level CLR (and 95% CI) from the unadjusted logistic random-effects regression model and the patient-level PPV (and 95% CI) from the unadjusted logistic regression model will be provided.

2.6 Analysis supporting secondary objectives

The following other secondary efficacy objectives are defined for the EFF unless otherwise specified:

- Evaluate the patient-level sensitivity of [^{18}F]CTT1057
- Evaluate the patient-level specificity of [^{18}F]CTT1057
- Evaluate the patient-level negative predictive value of [^{18}F]CTT1057
- Evaluate the patient-level accuracy of [^{18}F]CTT1057
- Evaluate correct detection rate (CDR)
- Evaluate detection rate
- Evaluate the region level sensitivity of [^{18}F]CTT1057
- Evaluate the region level specificity of [^{18}F]CTT1057
- Evaluate the region level negative predictive value of [^{18}F]CTT1057
- Evaluate the region level accuracy of [^{18}F]CTT1057
- Characterize the safety and tolerability of [^{18}F]CTT1057
- [^{18}F]CTT1057 scan inter-reader variability
- [^{18}F]CTT1057 scan intra-reader variability
- Concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of lesions using central reads
- Evaluate the change in patient management plans attributed to the PET/CT scan
- Assess all the above secondary objectives independently in the subgroup of patients with prior radical prostatectomy (RP) and the subgroup of patients with prior curative intent radiation therapy (RT)

2.6.1 Secondary endpoint(s)

For the following secondary endpoints, centralized imaging assessments (CTS Level 2) will be used as reference standard if pathology assessment (CTS Level 1) is not available for a lesion, inconclusive or negative (for biopsy only).

Patient-level sensitivity is defined as the proportion of TP participants among those who are TP or FN

Patient-level specificity is defined as the proportion of TN participants among those who are TN or FP

Patient-level negative predictive value is defined as the proportion of TN participants among those who are TN or FN

Patient-level accuracy is defined as the proportion of TP and TN participants among all scanned participants (regardless of TP, TN, FP, FN)

Patient-level correct detection rate is defined as the proportion of TP participants among all scanned participants (regardless of TP, TN, FP, FN)

Patient-level detection rate is defined as the proportion of TP and FP participants among all scanned participants (regardless of TP, TN, FP, FN)

where TP, TN, FP, FN participants are defined as follows:

- PSMA TP participants are those who have at least one TP lesion (with anatomically localized correspondence between [^{18}F]CTT1057 PET imaging and the CTS).
- PSMA TN participants are those who do not show any pathological [^{18}F]CTT1057 uptake and will be confirmed as NOT having any lesions with the CTS.
- PSMA FP participants are those who show at least one pathological [^{18}F]CTT1057 uptake but will be verified as NOT having any lesions with the CTS or no lesions are correctly localized in anatomical location by the CTS.
- PSMA FN participants are those who do not show any pathological [^{18}F]CTT1057 uptake but will be confirmed as having at least one lesion with the CTS.

Number of TP, TN, FP, FN participants will be presented and used to calculate the diagnostic performance parameters.

Two-sided exact binomial 95% CIs will be calculated for patient-level diagnostic performance parameters.

[^{18}F]CTT1057 scan inter-reader variability is defined as agreement among three readers determinations and will be assessed by Fleiss' Kappa statistic.

[^{18}F]CTT1057 scan intra-reader variability is defined as within-reader agreement for two different time points and will be assessed by Cohen's Kappa statistic. A subset of imaging reading will be performed at two different time points by each reader. See details in imaging charter.

Change in patient management plans attributed to the [^{18}F]CTT1057 PET/CT scan is defined as the percentage of participants who undergo a change in intended treatment plan attributed to

the PET/CT scan as assessed by pre-imaging questionnaire 1 and post-imaging questionnaire 2.

The number and percentage of participants from the F-SAF for each of the following categories of treatment plan in each of the two questionnaires will be calculated: surgery, radiation alone, radiation alone with change in radiation treatment plan (only applicable for questionnaire 2), radiation plus ADT, ADT alone, observation/surveillance, other (free text box).

Treatment plan change is defined as a change in the pair of categories (one category of questionnaire 1, one category of questionnaire 2).

The number and percentage of participants from the F-SAF with treatment plan change as defined above will be calculated. Shift tables comparing treatment plans between questionnaire 1 and 2 will be displayed. The number (%) of participants with and without a change in treatment plans by [^{18}F]CTT1057 local read results (i.e. positive, negative, and inconclusive) will be tabulated.

Region-level sensitivity is defined as the proportion of TP regions among those that are TP or FN

Region-level specificity is defined as the proportion of TN regions among those that are TN or FP

Region-level negative predictive value is defined as the proportion of TN regions among those that are TN or FN

Region-level accuracy is defined as the proportion of TP and TN regions among those that are identified on [^{18}F]CTT1057 (regardless of TP, TN, FP, FN)

where TP, TN, FP, FN regions are defined as follows:

- PSMA TP regions are regions containing at least one TP lesion (anatomically localized correspondence between [^{18}F]CTT1057 PET imaging and the CTS).
- PSMA TN regions are regions containing no lesions with pathological [^{18}F]CTT1057 uptake and confirmed by CTS.
- PSMA FP regions are regions containing lesions with pathological [^{18}F]CTT1057 uptake but are verified negative by CTS or none is correctly localized in anatomical location by the CTS.
- PSMA FN regions are regions containing no lesions with pathological [^{18}F]CTT1057 uptake but are confirmed as having at least one lesion with the CTS.

Number of TP, TN, FP, FN regions will be presented and used to calculate the diagnostic performance parameters.

Two-sided 95% CIs for (overall) region-level diagnostic performance parameters will be calculated using a logistic random-effects models taking into account the correlation among regions within a participant.

For different regions (i.e. prostate region, PLN, extra-pelvic lymph nodes, skeletal, and visceral), two-sided exact binomial 95% CIs for region-level diagnostic performance parameters will be calculated.

Concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for detection of lesions and positive percent agreement will be calculated to assess the concordance.

2.6.2 Statistical hypothesis, model, and method of analysis

For patient-level diagnostic performance parameters:

- Patient-level sensitivity
- Patient-level specificity
- Patient-level negative predictive value
- Patient-level accuracy
- Patient-level correct detection rate
- Patient-level detection rate

The number and percentage of TP, TN, FP, and FN participants will be presented and used to calculate the diagnostic performance parameters.

Two-sided exact binomial 95% CIs for patient-level diagnostic performance parameters will be calculated.

[^{18}F]CTT1057 scan inter-reader variability will be assessed by Fleiss' Kappa statistic (Fleiss 1971) defined below:

$$\kappa_1 = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}.$$

Let N represent the total number of participants who have evaluable PET/CT, n the number of central readings per participant received ($n = 3$), and k the number of outcome categories ($k = 2$) into which assignments are made. Participants are indexed by $i = 1, \ldots, N$ and categories are indexed by $j = 1, \ldots, k$. Let n_{ij} represent the number of readers who assigned the ith participant to the jth category. Then $P_i = \frac{1}{n(n-1)} \sum_{j=1}^{k} n_{ij}(n_{ij} - 1)$ and $\bar{P} = \frac{1}{N} \sum_{i=1}^{N} P_i$;

$$p_j = \frac{1}{Nn} \sum_{i=1}^{N} n_{ij} \quad \text{and} \quad \bar{P}_e = \sum_{j=1}^{k} p_j^2.$$

A contingency table of N row (N is the number of participants who have evaluable PET/CT) by 2 (the number of outcome categories) will be obtained to summarize the frequency of agreements among the 3 readers for each case and each patient-level outcome (i.e. positive, negative). The counts across each of the N rows will sum up to 3 (number of readers).

Based on the contingency table as described above, a Fleiss's Kappa statistic and corresponding 95% confidence interval will be estimated based on asymptotic estimation of the standard error (i.e.

$\widehat{se}(\kappa_1) = \frac{\sqrt{2}}{\sum_{j=1}^{k} p_j(1-p_j)\sqrt{Nn(n-1)}} \sqrt{\left(\sum_{j=1}^{k} p_j\ (1-p_j)\right)^2 - \sum_{j=1}^{k} p_j(1-p_j)(1-2p_j)}$) (Fleiss, Nee, Landis 1979; Fleiss J, Levin B, Paik MC 2003) and the normality assumption (i.e. 95% CI can be calculated as: $\kappa_1 \pm 1.96\widehat{se}(\kappa_1)$). An additional table will be presented to show the

distribution of agreements with the number (%) of scans agreed by two readers and the number (%) of scans agreed by all three readers. Inter-reader variability will be calculated using the F-SAF.

[^{18}F]CTT1057 scan intra-reader variability will be assessed by Cohen's Kappa statistic.

To assess the intra-reader variability, each of the three readers will re-read the same 20 randomly selected cases (with evaluable PET/CT) but each case will be presented to each reader in a random order. The readers will not reference any previous read results nor discuss any of these cases at any time during the read process. Each reader will use the same method to read the randomly presented cases as they did in the previous reads. Intra-reader agreement will first be evaluated separately for each of the 3 readers using Cohen's kappa statistic (κ_2) as described below:

Step 1: A 2 by 2 contingency table (e.g. Table 2-1) will be calculated to summarize the frequency of concordance and discordance for patient-level positivity or negativity. The counts in the table will add up to 20 in total (i.e. a + b+ c + d = 20).

Table 2-1 Example 2 by 2 Contingency Table for Each Reader

		New Reads	
		Positive	Negative
Original Reads	Positive	a	b
	Negative	c	d

Step 2: Based on the contingency tables described above, a simple agreement rate, i.e. $\frac{a+d}{a+b+c+d}$ and its corresponding 95% confidence interval (Agresti A, Coull BA 1998) will be estimated based on normal approximation.

Step 3: Finally, a Cohen's kappa statistic (κ_2) and its corresponding 95% confidence interval will be estimated (Cohen J 1960; Cohen J 1968; Fleiss J, Levin B, Paik MC 2003) using the formula $\kappa_2 \pm 1.96SE$, where $SE = \sqrt{\frac{p_o(1-p_o)}{N(1-p_e)^2}}$ (McHugh M 2012). Let N denote the number of re-reads ($N = 20$), k the number of categories ($k = 2$), n_{i1} the number of times the original reads were assigned to i ($i = 1, \dots, k$) and $n_{i2}$the number of times the re-read were assigned to i ($i = 1, \dots, k$). Then Cohen's kappa can be calculated as $\kappa = \frac{p_o - p_e}{1-p_e}$, where p_o is the relative agreements among the original and re-reads ($p_o = \frac{a+d}{a+b+c+d}$ in Table 2-1) and $p_e = \sum_{i=1}^{k} \frac{n_{i1}}{N}\frac{n_{i2}}{N}$ ($p_e = \frac{a+b}{N}\frac{a+c}{N} + \frac{c+d}{N}\frac{b+d}{N}$ in Table 2-1). Intra-reader variability will be calculated for the F-SAF.

For region-level diagnostic performance parameters:

- **Region-level sensitivity**

- **Region-level specificity**
- **Region-level negative predictive value**
- **Region-level accuracy**

Number of TP, TN, FP, FN regions will be presented and used to calculate the diagnostic performance parameters.

Two-sided 95% CI's for (overall) region-level diagnostic performance parameters will be calculated using a logistic random-effects models taking into account the correlation among regions within a participant.

For different regions (i.e. prostate region, PLN, skeletal, extra-pelvic lymph nodes and visceral), two-sided exact binomial 95% CI's for region-level diagnostic performance parameters will be calculated.

Concordance between [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 For detection of PSMA-positive lesions, positive percent agreement (PPA) will be calculated to assess the concordance. Logistic random-effects models will be used taking into account the correlation among lesions within a participant overall and by region. When calculating PPA, the number of positive lesions identified and anatomically co-localized by both [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 for participant i (see [redacted]), y_i, follows the binomial distribution with parameter (n_i, p_i),

$$y_i|b_i \sim \text{Binomial}(n_i, p_i)$$

p_i denotes the PPA for participant i, n_i corresponds to the denominator in the PPA definition, i.e. the number of all positive lesions identified by [^{68}Ga]Ga-PSMA-11 for participant i.

PPA along with two-sided 95% CIs will be estimated by the logistic regression model with random effects below:

$$\log\left(\frac{p_i}{1-p_i}\right) = \mu + b_i$$

where p_i is the lesion-level PPA for participant i, μ is the fixed effect of overall mean, b_i is the random effect of participants (clusters). b_i follows a normal distribution : $b_i \sim N(0, \sigma_b^2)$.

The same analysis will be repeated by region and combining skeletal, extra-pelvic lymph nodes, and visceral regions (M1).

The number of PSMA-positive lesions identified on [^{18}F]CTT1057 PET/CT, on [^{68}Ga]Ga-PSMA-11 PET/CT, on [^{18}F]CTT1057 PET/CT but not on [^{68}Ga]Ga-PSMA-11, on [^{68}Ga]Ga-PSMA-11 but not on [^{18}F]CTT1057 PET/CT, and the number of PSMA-positive lesions identified on both[^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11 PET/CT will be displayed overall, by region and by combined M1 region.

The above analyses will be provided by each central reader using the F-SAF.

2.6.3 Handling of intercurrent events

Though no secondary estimands are defined for the study, intercurrent events defined in the primary estimands will impact secondary endpoints analysis. Thus they will be handled in the same way as in primary estimands analyses. Participants who receive the investigational imaging agent [^{18}F]CTT1057 but do not undergo/complete the PET/CT scan for any reasons (e.g. consent withdrawal, PET camera failures, etc.) will not be used for the secondary efficacy endpoint analyses, but safety data for these participants will be collected after the administration of the investigational imaging agent and will be summarized.

2.6.4 Handling of missing values not related to intercurrent event

This is a diagnostic study, where the secondary endpoints are calculated on the basis of a one-time imaging assessment. Missing data will not be imputed.

2.6.5 Sensitivity analyses

No sensitivity analyses are planned for the secondary endpoints.

2.6.6 Supplementary analyses

Supplementary analyses for patient-level and region-level diagnostic performance parameters defined in Section 2.6.2 will be performed for the following subgroup:

- PSA level at baseline (from central laboratory):
  - PSA ≤ 0.5 ng/mL
  - 0.5 ng/mL<PSA≤1 ng/mL
  - 1 ng/mL<PSA≤2 ng/mL
  - 2 ng/mL<PSA≤5 ng/mL
  - PSA > 5 ng/mL

2.7 Safety analyses

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). Safety summary tables will be presented using the F-SAF and Ga-SAF separately, and listings will be provided for the SAF, unless otherwise specified.

The overall observation period will be divided into three mutually exclusive segments (see Section 2.1.1 for details).

2.7.1 Adverse events (AEs)

Adverse events are coded using MedDRA terminology. The latest MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings. AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. Summary tables for adverse events (AEs) will include only AEs that started or worsened during the on-treatment period (see Section 2.1.1). All AEs that started or worsened within 14 days after the [^{18}F]CTT1057 dose and as long as they were prior to the [^{68}Ga]Ga-

PSMA-11 dose (applicable for Sequence 1), and adverse events that were reported as related to [^{18}F]CTT1057 irrespective of time of onset will be regarded as [^{18}F]CTT1057 **treatment-emergent** AEs. Similarly, all AEs that started or worsened within 14 days after the [^{68}Ga]Ga-PSMA-11 dose and as long as they were prior to the [^{18}F]CTT1057 dose (applicable for Sequence 2), and adverse events that were reported as related to [^{68}Ga]Ga-PSMA-11 irrespective of time of onset will be regarded as [^{68}Ga]Ga-PSMA-11 **treatment-emergent** AEs. [^{18}F]CTT1057 TEAEs and [^{68}Ga]Ga-PSMA-11 TEAEs will be summarized.

AEs will be summarized by the number and percentage of participants having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT) using MedDRA coding. A participant with multiple occurrences of an AE will be counted only once in the respective AE category. A participant with multiple CTCAE grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries, the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency.

The following adverse event summaries will be produced: overview of adverse events and deaths (number and % of participants who died, with any AE, any SAE, any dose change, drug withdrawn etc.), AEs by SOC, PT and maximum grade, by relationship (all AEs and AEs related to study treatment), seriousness (SAEs and non-SAEs), leading to treatment discontinuation, leading to dose change, requiring additional therapy, and leading to fatal outcome. In addition, a summary of serious adverse events with number of occurrences will be produced (an occurrence is defined as >1 day between start and prior end date of record of same preferred term). To help evaluate the impact of the COVID-19 on the safety, the incidence of COVID-19 related adverse event preferred terms will be presented. All COVID-related AEs will be included in the listings.

For the legal requirements of ClinicalTrials.gov and EudraCT, two required tables for on-treatment emergent adverse events which are not serious adverse events with an incidence of greater than 5% and on-treatment emergent serious adverse events and SAE suspected to be related to study treatment will be provided by SOC and PT for the SAF.

If for the same participant, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

All AEs collected in the AE (e)CRF page will be listed along with the information collected on those AEs, e.g. AE relationship to study drug, AE outcome, etc. All AEs, deaths, and serious adverse events (including those from the pre- and post-treatment periods) will be listed and those collected outside of the on-treatment period will be flagged.

2.7.1.1 Safety topics of interest / grouping of AEs

A safety topic of interest (STI) is a grouping of adverse events that are of scientific and medical concern specific to the tracer [^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. STI will be defined at the project level and may be regularly updated. The grouping of AEs in STI according to project standards will be specified in the electronic Case Retrieval Sheet (eCRS) for each specific tracer ([^{18}F]CTT1057 and [^{68}Ga]Ga-PSMA-11). For each specified STI, the number and percentage of participants with at least one event of the STI occurring during the on-treatment period will be summarized.

STIs will be summarized by grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, hospitalization, death etc. A listing of all grouping levels down to the MedDRA preferred terms used to define each STI will be generated. The analyses related to STIs for [^{18}F]CTT1057 will be carried out using the F-SAF and the analyses related to STIs for [^{68}Ga]Ga-PSMA-11 will be carried out using the Ga-SAF.

2.7.2 Deaths

Separate summaries for on-treatment and all deaths (including post-treatment death) will be produced. On-treatment death analyses will be presented for the F-SAF and Ga-SAF separately, by system organ class and preferred term. All deaths analyses will be presented for the SAF, by system organ class, preferred term, and treatment sequence.

All deaths will be listed, and post-treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened participants.

2.7.3 Laboratory data

Laboratory data from all sources (central and local laboratories) will be combined for the analysis. The summaries will include all assessments available for the lab parameter collected prior to (baseline) and within 14 days (+ 3 day-window) after each imaging agent administration (and prior to the subsequent imaging agent administration, if applicable).

All laboratory data will be listed by participant and visit/time and if normal ranges are available, abnormalities will be flagged. Summary statistics will be provided by visit/time. Shift tables

using the low/normal/high/ (low and high) classification will be used to compare baseline to the worst on-treatment value.

Grading of laboratory values will be assigned programmatically as per National Cancer Institute (NCI) CTCAE version 5.0. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used.

For laboratory tests where grades are not defined by CTCAE version 5.0, results will be categorized as low/normal/high based on laboratory normal ranges.

The following listings/summaries will be generated separately for hematology and biochemistry tests:

- Listing of all laboratory data with values flagged to show the corresponding CTCAE version 5.0 grades if applicable and the classifications relative to the laboratory normal ranges

For laboratory tests where grades are defined by CTCAE version 5.0:

- Worst post-baseline CTCAE grade (regardless of the baseline status). Each participant will be counted only once for the worst grade observed post-baseline.
- Shift tables using CTCAE version 5.0 grades to compare baseline to the worst on-treatment value

For laboratory tests where grades are not defined by CTCAE version 5.0:

- Shift tables using the low/normal/high/ (low and high) classification to compare baseline to the worst on-treatment value.

2.7.4 Other safety data

2.7.4.1 ECG and cardiac imaging data

The number and percentage of participants with an abnormal ECG (clinically significant) at screening will be presented. A listing of all ECG assessments will be produced. These analyses will be performed for the FAS.

2.7.4.2 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters were collected: body temperature, blood pressure, heart rate, respiratory rate.

Data handling

Vital signs collected prior to (baseline), and within 14 days (+ 3 day-window) after each imaging agent administration (and prior to the subsequent imaging agent administration, if applicable) will be summarized. Values measured outside of the on-treatment period will be flagged in the listings.

Data analysis

For analysis of vital signs, the clinically notable vital sign criteria are provided in Table 2-2 below.

Table 2-2 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	Increase	**Decrease**
Systolic blood pressure (mmHg)	>=180 with increase from baseline of >=20	<=90 with decrease from baseline of >=20
Diastolic blood pressure (mmHg)	>=105 with increase from baseline of >=15	<=50 with decrease from baseline of >=15
Body temperature (°C)	>= 39.1	-
Pulse rate (bpm)	> 100 with increase from baseline of > 25%	< 60 with decrease from baseline > 25%
Respiratory rate	> 22 breaths per minute	<14 breaths per minute

The number and percentage of participants with notable vital sign values (high/low) will be presented. Descriptive statistics will be tabulated for baseline, at each post-baseline time point and changes from baseline at each post-baseline time point for each vital sign measure.

A listing of all vital sign assessments will be produced by treatment sequence and notable values will be flagged. In the listing, the assessments collected outside of on-treatment period will be flagged.

2.8 Pharmacokinetic endpoints

No pharmacokinetic endpoints are defined in the study.

2.9 PD and PK/PD analyses

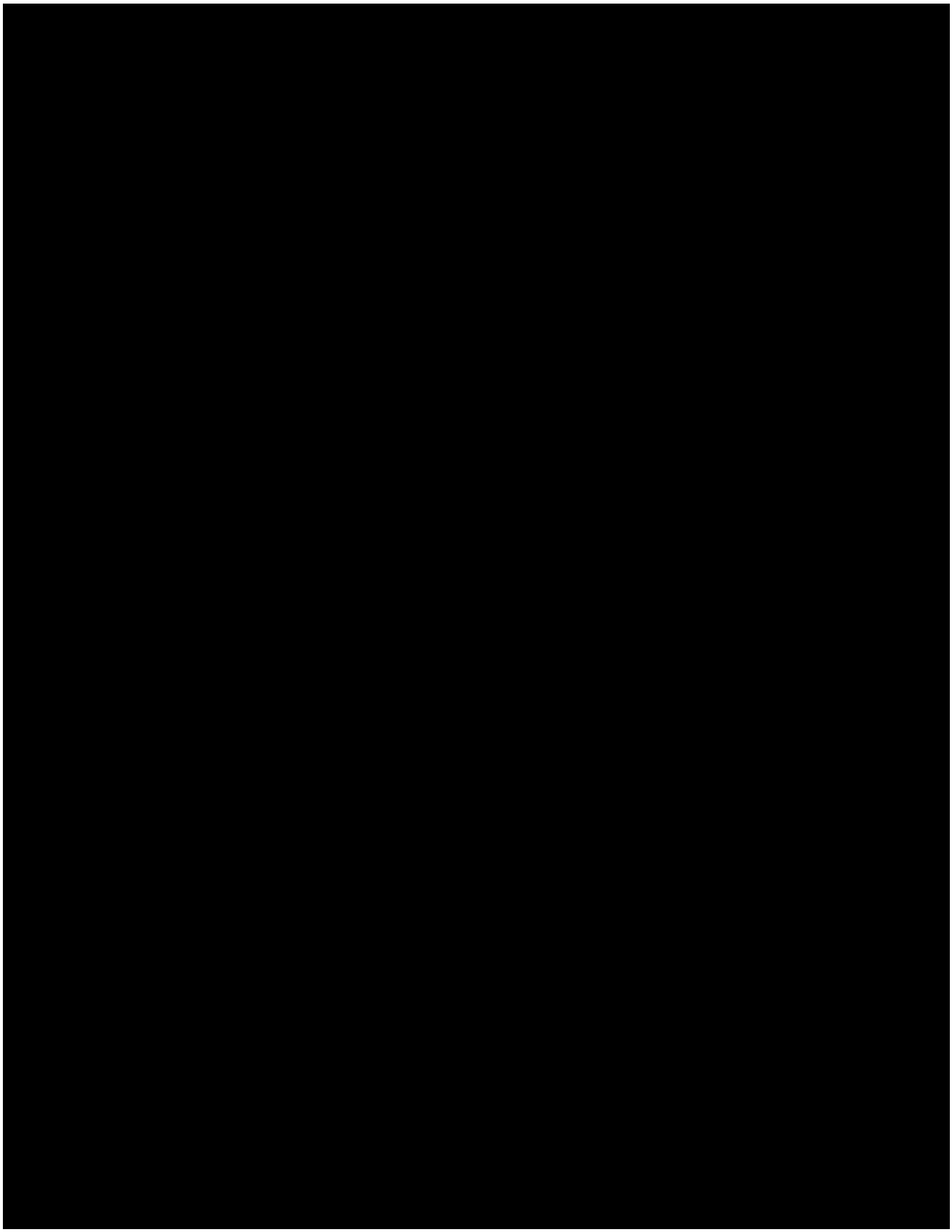
No PD and PK/PD analyses are planned in the study.

2.10 Patient-reported outcomes

No patient-reported outcomes analyses are planned in the study.

2.11 Biomarkers

No biomarker analyses are planned in the study.



2.13 Interim analysis

No interim analysis is planned for this study.

3 Sample size calculation

Overall, approximately 190 participants will be enrolled to ensure at least 152 participants are evaluable (i.e. those who have both an evaluable [18F]CTT1057 PET/CT scan imaging, and at least one evaluable CTS assessment and have not received any prohibited systemic antineoplastic therapy before the completion of PET/CTs and CTS procedures), which will be required for the calculation of the co-primary endpoints. These calculations were made using the software PASS 11 and R 3.6.1.

3.1 Primary endpoints(s)

This study is planned to recruit approximately 190 participants. The sample size calculation is based on the co-primary endpoints of region-level CLR and the patient-level PPV of [^{18}F]CTT1057.

The assumptions for sample size calculations are based on the estimates from data available from prior studies.

The primary endpoint is to evaluate region-level CLR (prostate region, PLN, extra-pelvic lymph nodes, skeletal, and visceral) and patient-level PPV of [^{18}F]CTT1057 PET for detection of tumor location confirmed by CTS. Based on the results of previous studies, the following distribution of disease across regions are anticipated: Prostate bed: 30%, pelvis: 15%; extrapelvic soft tissue: 20%; bone metastases: 35% (Eiber et al 2015, Ceci et al 2015, Van Leeuwen et al 2016). It is anticipated that the CLR for the five regions and for all regions combined using conventional imaging ranges from 30-60%. An overall CLR for [^{18}F]CTT1057 PET of at most 50% will be considered as unacceptably low. Hence, the null hypothesis that the CLR is at most 50% will be tested against the alternative hypothesis that the CLR is greater than 50%.

The BCR population targeted in this study is expected to have similar PSA values as prior studies (Morris et al 2020) of approved diagnostic PET agents for PCa such as [^{18}F]-fluciclovine, which had a 27% PPV for patients with PSA values that were less than 1.78, as noted in their summary basis of approval. Thus, a patient-level PPV of 20% will be considered as unacceptably low. The [^{18}F]CTT1057 positive patient rate (detection rate) for the study population is 80% based on [^{18}F]-fluciclovine studies. The sample size calculation will be performed based on the null and alternative hypotheses as follows.

Region-level CLR

Correlated binary test results occur in this situation where each experimental unit (participant) consists of five correlated regions. Lee and Dubin developed a sample size formula for clustered binary data (Lee and Dubin 1994). They propose estimating the binary proportion by assigning equal weights to clusters regardless of their sizes to simplify the derivation of their sample size

formula. The weighted estimator for region-level CLR is approximately Gaussian distributed with a variance of a beta distribution (beta-binomial model).

The null hypothesis (H0: region-level CLR p = 0.50) will be tested against the alternative hypothesis (H1: p > 0.50). Assuming a CLR of 0.58 under the alternative hypothesis, a total sample size of 172 enrolled participants would achieve 90% statistical power to detect a change in CLR of 0.08 using a one-sided test at a target significance level of 2.5%. This sample size includes an adjustment for a dropout rate of 20%.

The sample size was determined as n to ensure a type I error using a one-sided significance level of 2.5% and with 90% power to reject H0 if the alternative H1 is true.

By the standard normal approximation, the sample size is given as follows:

$$n = \frac{\sigma^2 \left(Z_\alpha - Z_{1-\beta}\right)^2}{(p_1 - p_0)^2}$$

where σ^2 is the variance of a beta distribution, Z_a is the upper quantile of the standard normal distribution, p_0 is a CLR of 0.50 under the null hypothesis, and p_1 is a CLR of 0.58 under the alternative hypothesis.

Suppose there is no information on the within-cluster correlation of test results. In this case, a uniform distribution seems reasonable, setting parameters in the beta distribution of alpha=1 and beta=1. Then σ^2 is 1/12.

Under these assumptions, 137 participants will be necessary to ensure 90% power. Taking into account a 20% dropout rate, a total of 172 participants will be recruited for this endpoint of the study.

The lower bound of the 95% CI for CRL should be greater than 0.50 to be considered a success.

Patient-level PPV

The null hypothesis (H0: patient-level PPV p = 0.20) will be tested against the alternative hypothesis (H1: p > 0.20). Assuming a patient-level PPV of 0.33 under the alternative hypothesis, a total sample size of 190 enrolled participants (which includes 152 [^{18}F]CTT1057 scanned participants based on 80% [^{18}F]CTT1057 positive patient rate) would achieve 90% power to detect a change in patient-level PPV of 0.13 using a one-sided binomial test at a targeted significance level of 2.5%. This sample size includes an adjustment for a dropout rate of 20%.

The lower bound of the 95% CI for patient-level PPV should be greater than 0.20 to be considered a success.

A total sample size of 190 participants can ensure 92.7% statistical power for region-level CLR and 90% statistical power for patient-level PPV resulting in an overall study statistical power of at least 83.4% (0.927*0.90).

4 Change to protocol specified analyses

[redacted]

The subgroup analyses on participants with prior RP and participants with prior curative intent RT will be applied to co-primary and secondary efficacy endpoints only.

Minor updates to the estimand wording were made as compared to the protocol for clarification.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The complete date of dose administration is required for this study. Completely or partially missing date will not be imputed and should be considered as a data issue and the statistician should contact the data manager of the study.

5.1.2 AE, ConMeds and safety assessment date imputation

Table 5-1 Imputation of start dates (AE, CM) and assessments (LB, EG, VS)

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates

day, month	• If available year = year of study treatment start date then
	  ○ If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY
	  ○ Else set start date = study treatment start date.
	• If available year > year of study treatment start date then 01JanYYYY
	• If available year < year of study treatment start date then 01JulYYYY

Missing Element	Rule
day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY. o Else set start date = study treatment start date. • If available month and year > month and year of study treatment start date then 01MONYYYY • If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates (AE, CM)

Missing Element	Rule (*=last study treatment date plus 15 days not > (death date, cut-off date, withdrawal of consent date))
day, month, and year	• Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	• If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	• If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

If imputed end date < start date, set end date = start date.

Any AEs and ConMeds with partial/missing dates will be displayed as such in the data listings.

Any AEs and ConMeds which are continuing as per data cut-off will be shown as 'ongoing' rather than the end date provided.

5.1.3 Other imputations

For incomplete date of initial diagnosis of cancer and date of biochemical recurrence diagnosis missing day is defaulted to the 15th of the month and missing month and day is defaulted to 01-Jan.

5.2 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

CTCAE grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death. This grading system inherently places a value on the

importance of an event, although there is not necessarily proportionality among grades (a grade 2 is not necessarily twice as bad as a grade 1).

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 (specify version used in the RAP). The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters. The latest available version of the document based on the underlying CTCAE version 5.0 at the time of analysis will be used. For laboratory tests where grades are not defined by CTCAE v5.0, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

Imputation Rules

CTC grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of WBC.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X.

The following rules will be applied to derive the WBC differential counts when only percentages are available for a xxx differential

xxx count = (WBC count) * (xxx %value / 100)

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

Corrected Calcium (mmol/L) = Calcium (mmol/L) + 0.02 [40 (g/L) - Albumin (g/L)]

For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

CTC grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) and corrected calcium will be assigned as described above for grading

5.4 Statistical models

5.4.1 Analysis supporting primary objective(s)

For region-level CLR, statistical model details are presented in Section 2.5. The GLIMMIX procedure in SAS will be used to implement the logistic regression model with random effects.

This analysis will be restricted to participants with at least one PET positive region (TP or FP) by central assessments in the EFF. In case of problems with fitting the model and estimating parameters, the default estimation technique (i.e. residual pseudo-likelihood with a subject expansion (METHOD = RSPL)) will be used. This analysis will be applied to each of the three central readers via the (BY READER) statement.

For this model, the SAS code will be:

```
proc glimmix data=dataset;
by reader;
   class subject;
   model tpcount / poscount= / solution;
   random intercept / subject=subject;
run;
```

- *tpcount* refers to the number of TP regions per participant
- *poscount* refers to the number of PET positive regions (TP or FP) per participant

The estimated fixed effect $\hat{\mu}$ and its standard error SE on the logit scale can be obtained from the output. And the (subject-specific) region-level CLR (conditional on the random effect when $b_i = 0$) along with the two-sided 95% CI can be computed by applying the inverse link function:

Point estimate of the region-level CLR: $E[p_i|\widehat{b_i = 0}] = E\left[\frac{y\_i}{n\_i}|\widehat{b_i = 0}\right] = \frac{\exp(\hat{\mu})}{1+\exp(\hat{\mu})}$

$$CI_{Lower} = \frac{\exp(\hat{\mu} - z_{0.975}SE)}{1 + \exp(\hat{\mu} - z_{0.975}SE)}$$

$$CI_{Upper} = \frac{\exp(\hat{\mu} + z_{0.975}SE)}{1 + \exp(\hat{\mu} + z_{0.975}SE)}$$

For patient-level PPV, an exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way tables) will be calculated (Clopper and Pearson 1934)

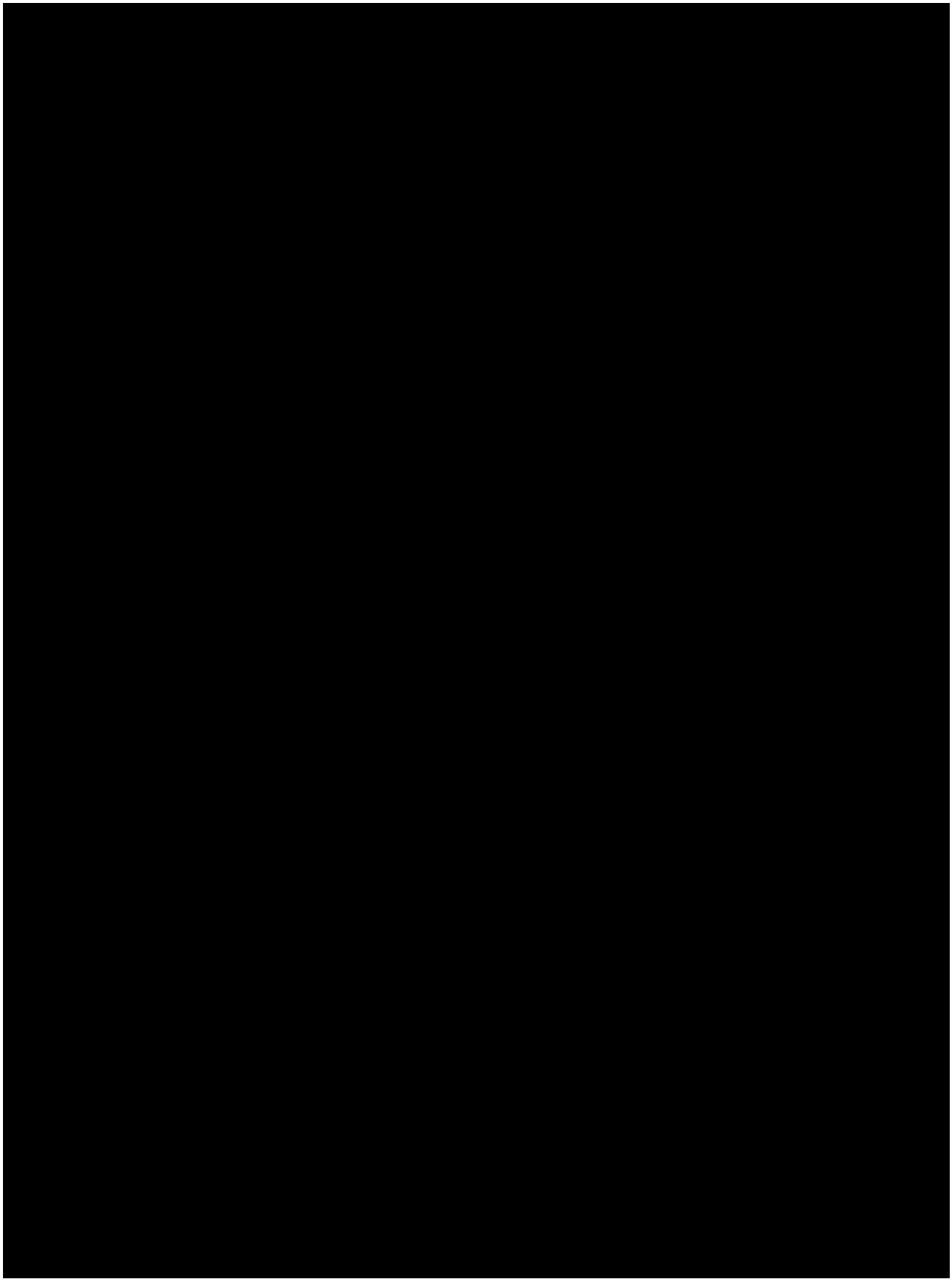
Efficacy endpoints be summarized in terms of percentage rates with 95% CIs.

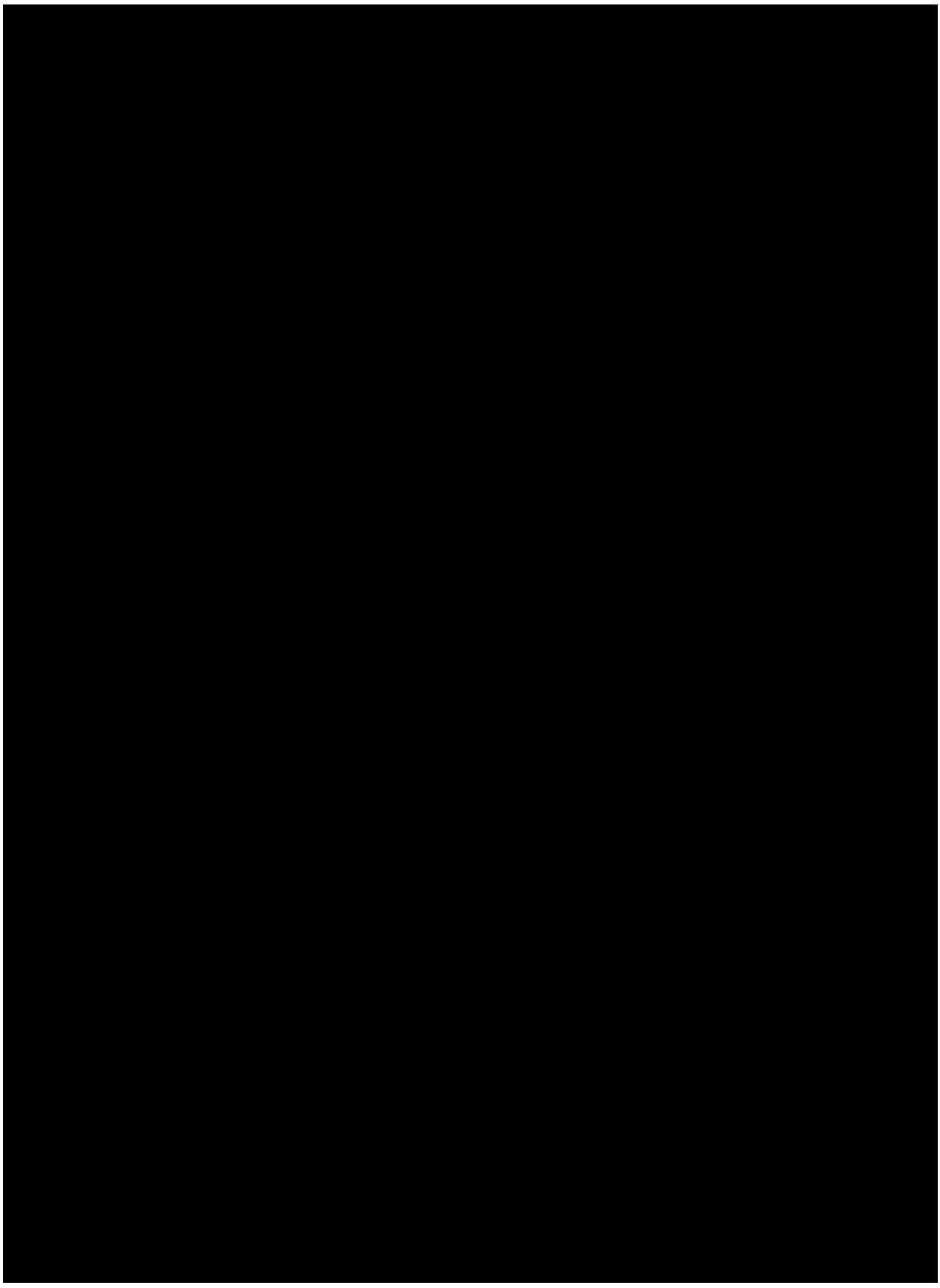
5.4.2 Analysis supporting secondary objective(s)

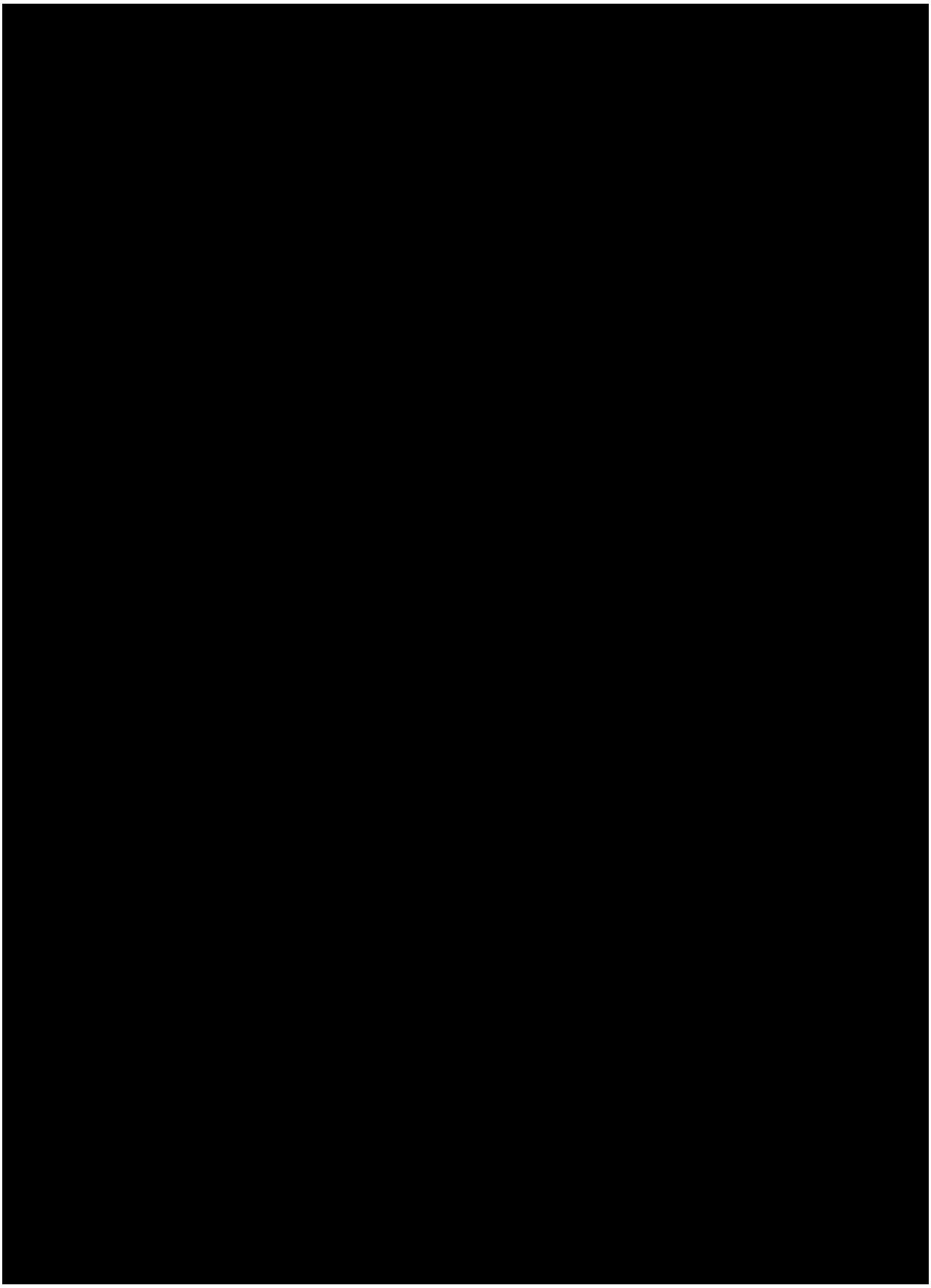
For region-level efficacy endpoints, statistical model details are presented in section 2.6. The GLIMMIX procedure in SAS will be used to implement the logistic regression model with random effects.

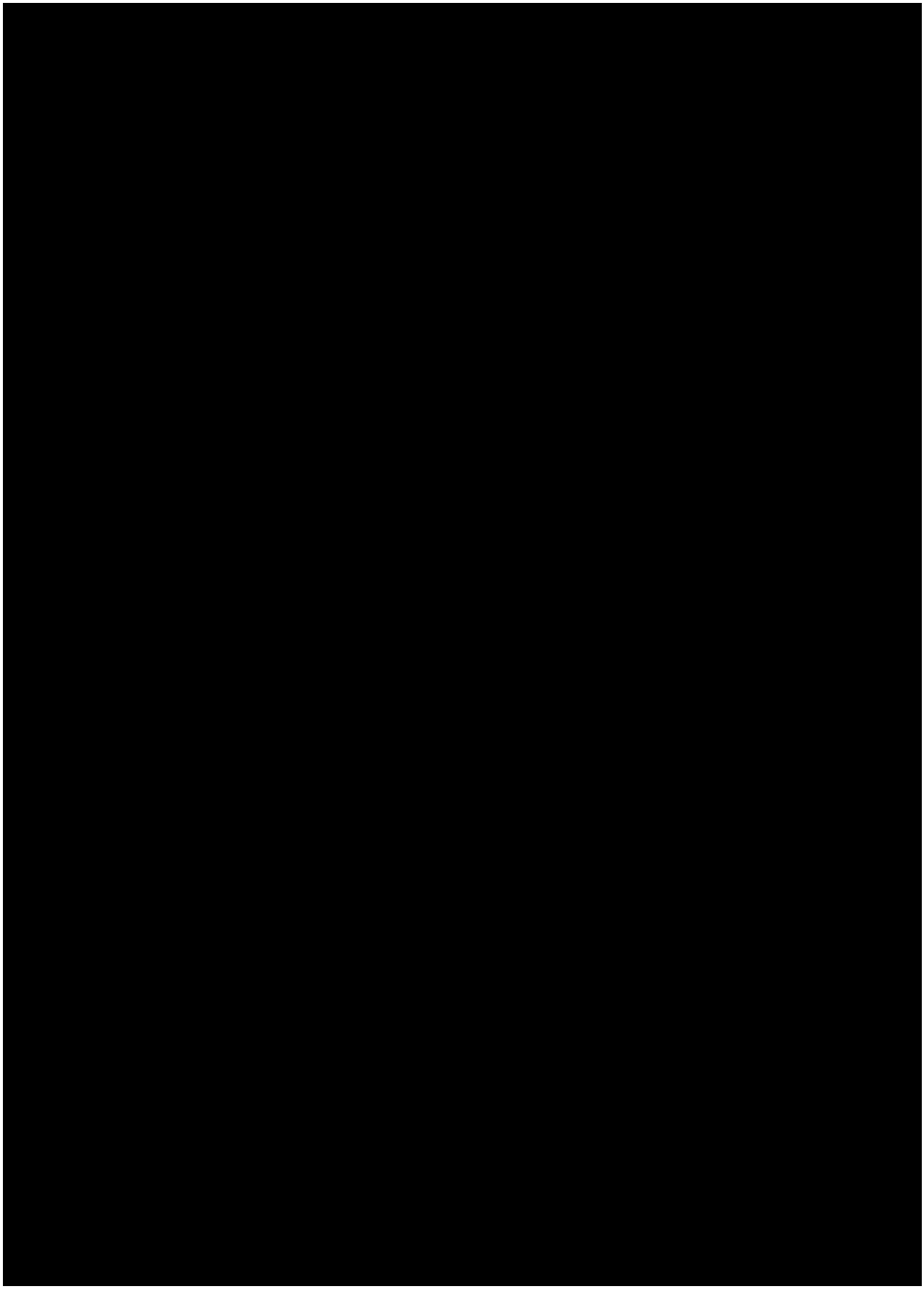
For patient-level efficacy endpoints, an exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way tables) will be calculated (Clopper and Pearson 1934)

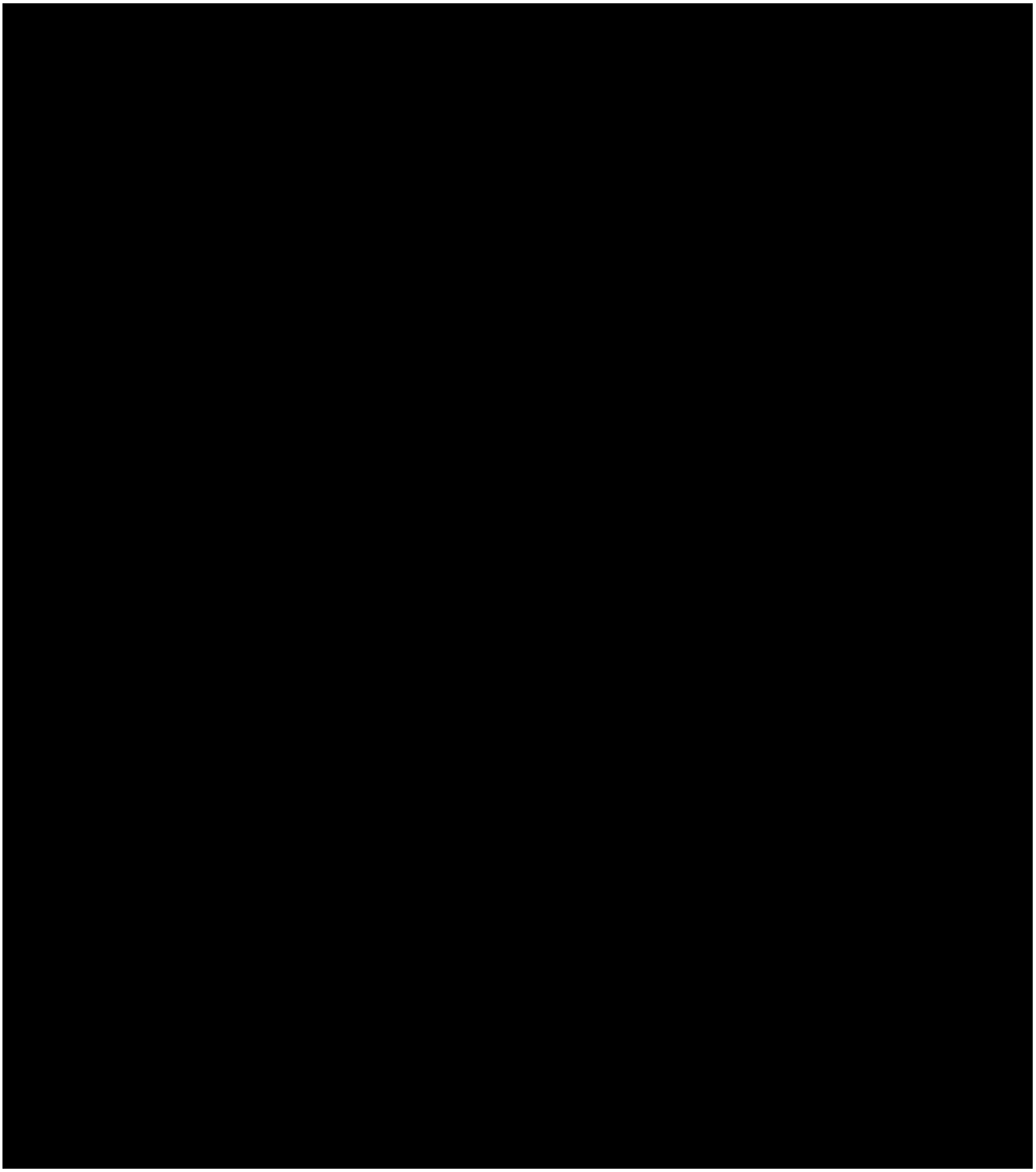
Efficacy endpoints be summarized in terms of percentage rates with 95% CIs.











6 Reference

Agresti A, Coull, BA (1998) Approximate is better than "Exact" for interval estimation of binomial proportions. The American Statistician: 52(2):119-26.

Ceci F, Uprimny C, Nilica B, et al (2015) (68)Ga-PSMA PET/CT for restaging recurrent prostate cancer: which factors are associated with PET/CT detection rate? Eur J Nucl Med Mol Imaging; 42:1284-94.

Chantadisai M, Buschner G, Krönke M, et al (2020) Positive predictive value and correct detection rate of 18F-rhPSMA-7 PET in biochemically recurrent prostate cancer validated by composite reference standard. J Nucl Med; 12:1-38.

Clopper, C., & Pearson, E. (1934). The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika, 26(4), 404-413.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement; 20(1): 37-46.

Cohen, J. (1968). Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychological bulletin; 70(4): 213.

Eiber M, Maurer T, Souvatzoglou M, et al (2015) Evaluation of Hybrid ^{68}Ga-PSMA Ligand PET/CT in 248 Patients with Biochemical Recurrence After Radical Prostatectomy. J Nucl Med; 56(5):668-74.

Fleiss J. L. (1971). Measuring nominal scale agreement among many raters. Psychological bulletin; 76(5): 378-382.

Fleiss J, Levin B, Paik MC (2003) Statistical Methods for Rates and Proportions, In: 3rd Ed, Wiley Series in Probability and Statistics.

Fleiss J, Nee J, Landis J (1979) The large sample variance of kappa in the case of different sets of raters. Psychological bulletin; 86(5): 974-977.

Gatsonis CA. Random-effects models for diagnostic accuracy data. Acad Radiol 1995;2(suppl1):S14–S21; discussion S57–S67, S61–S64.

Lee EW, Dubin N (1994) Estimation and sample size considerations for clustered binary responses. Stat Med; 13:1241-52.

McHugh, M (2012) Interrater reliability: the kappa statistic. Biochemia Medica; 22(3): 276-282.

Morris MJ, Carroll PR, Saperstein L, et al (2020) Impact of PSMA-targeted imaging with 18F-DCFPyL-PET/CT on clinical management of patients (pts) with biochemically recurrent (BCR) prostate cancer (PCa): Results from a phase III, prospective, multicenter study (CONDOR) [abstract]. J Clin Oncol; 38 (15 Suppl):5501A.

Petersen LJ, Zacho HD (2020) PSMA PET for primary lymph node staging of intermediate and high-risk prostate cancer: an expedited systematic review. Cancer Imaging; 20:10.

Scher HI, Morris MJ, Stadler WM, et al (2016) Trial Design and Objectives for Castration-Resistant Prostate Cancer: Updated Recommendations From the Prostate Cancer Clinical Trials Working Group 3. J Clin Oncol; 34(12):1402-18

Van Leeuwen PJ, Stricker P, Hruby G, et al (2016) (68) Ga-PSMA has a high detection rate of prostate cancer recurrence outside the prostatic fossa in patients being considered for salvage radiation treatment. BJU Int; 117:732-9.

Verbeke G, Molenberghs G. Linear mixed models for longitudinal data. New York, NY: Springer, 2000.