STATISTICAL ANALYSIS PLAN (SAP)
BNT151-01

Version: 3.0 **Date:** 10 JUN 2024

Sponsor: BioNTech SE

Protocol number:	BNT151-01
Protocol title:	Phase I/IIa, first-in-human, open-label, dose escalation trial with expansion cohorts to evaluate safety, pharmacokinetics, pharmacodynamics, and preliminary efficacy of BNT151 as a monotherapy and in combination with other anti-cancer agents in patients with solid tumors
Short Title / Acronym	BNT151 as a monotherapy and in combination with other anti-cancer agents in patients with solid tumors
Trial Phase	I/IIa
Protocol version:	8.0
Protocol date:	13 DEC 2023
Compound:	BNT151
SAP version:	3.0
SAP date:	10 JUN 2024
Contract Research Organization name:	Syneos Health
SAP author:	PPD Staburo GmbH, on behalf of BioNTech SE

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 SAP APPROVAL

This SAP has been prepared, reviewed, and approved in accordance with the sponsor's standard operating procedures (SOP). Documentation of this process is filed in the trial master file.

For documentation of approval see e-signatures on the additional page at the end of this document.

2 VERSION HISTORY

Table 1: SAP version history summary

SAP version	Approval date	Change	Rationale
1.0	18 APR 2023	-	Final version
1.1	Final draft (not formally approved) 06 DEC 2023	• Removal of Immunogenicity Analyses (moved to PK & PD analysis plan) • Exploratory endpoint analyses highlighted as optional (PFS, OS, PK parameters) • Clarifications for baseline derivation, DTL evaluation period derivation, imputation of partial dates, survival follow-up period derivation. • Clarifications of censoring windows for PFS and how to apply them to DOR • Minor adjustment to current eCRF (ATC level 4 instead of ATC level 5, physical examination not by body system) • Minor wording fixes	-
2.0	24 MAY 2024	• Delete Part 2 related analyses. • Clarification of imputation dates for most recent progression disease and prior systemic cancer therapy • Clarification of time points in shift tables. • Change of ATC level 4 to standardized medication name in the summary of concomitant medication • Remove KM curve for DOR • Minor clarifications and fixes	Final version
3.0	10 JUN 2024	• Formatting changes and change of signature page • Clarification of number of cycles for pre-conditioning cohort	Formatting changes for final version

3 INTRODUCTION

This is a Phase I/IIa, first-in-human, open-label, dose escalation trial with expansion cohorts to evaluate safety, pharmacokinetics (PK), pharmacodynamics (PD) and preliminary efficacy trial of BNT151 as monotherapy and in combination with other anti-cancer agents in patients with solid tumors.

This SAP describes the detailed procedures for the planned statistical analysis for protocol BNT151-01 to support the completion of the clinical trial report.

Safety Review Committee (SRC) and the analyses to support SRC data reviews are not a part of this SAP.

The purpose of this SAP is to define and describe the statistical methodologies to ensure complete and appropriate analyses and allow valid conclusions regarding the trial objectives.

This SAP is only applicable to Part 1 of the trial.

Syneos Health will perform the statistical analyses (using Statistical Analysis Software [SAS®] software version 9.4 or higher, and/or other statistical software as required) and is responsible for the production and quality control of all tables, figures and listings for all deliveries (draft, main, final).

PK analyses will be performed using Phoenix WinNonlin® version 8.0 or higher according to a separate analysis plan and validated by Syneos Health.

Templates for each unique table, figure and patient listing are provided in a separate document “Data Presentation Plan” (DPP).

3.1 Objectives and endpoints

Trial objectives and endpoints are listed in Table 2.

Table 2: Objectives and endpoints

Objectives	Endpoints
Primary objectives	**Endpoints**
For Part 1, identify the maximum tolerated dose (MTD) and/or recommended Phase II dose (RP2D) of IMP based on the occurrence of dose limiting toxicities (DLTs) using the following definitions: • MTD defined as the highest tolerated dose where less than 1/3 of the patients experience a DLT. • RP2D based on integrated evaluation of safety, tolerability, clinical benefit, PK, and selected PD markers, for all dose levels tested.	• Occurrence of DLTs within a patient during the DLT evaluation period.
Assess the safety and tolerability of IMP.	• Occurrence of TEAE within a patient including Grade ≥3, serious, fatal TEAE by relationship.

Objectives	Endpoints
	• Occurrence of dose reduction and discontinuation of IMP within a patient due to TEAE.
Secondary objectives	**Endpoints**
Evaluate anti-tumor activity of IMP according to RECIST 1.1.	• The ORR is defined as the proportion of patients in whom a CR or PR (per RECIST 1.1) is observed as best overall response. • The DCR is defined as the proportion of patients in whom a CR or PR or SD (per RECIST 1.1, SD assessed at least 6 weeks after first dose) is observed as best overall response. • The DOR is defined as the time from first objective response (CR or PR per RECIST 1.1) to first occurrence of objective tumor progression (progressive disease per RECIST 1.1) or death from any cause, whichever occurs first.
Exploratory objectives	**Endpoints**
Evaluate efficacy of IMP in terms of PFS and OS.	• PFS defined as the time from first dose of IMP to first objective tumor progression, or death from any cause, whichever occurs first. • OS defined as the time from first dose of IMP to death from any cause.[1]
Characterize the PK profile of translated IL-2 variant.	• PK parameters (including but not limited to AUC, C_{max}, t_{max}, and $t_{1/2}$).[1]
Identify potentially predictive or other exploratory pharmacodynamic markers.	• Changes in selected cytokines and other soluble innate and adaptive immune system activation markers compared to baseline.[1] • Changes in systemic and intra-tumoral immune response in blood and tumor tissue compared to baseline (e.g., immunophenotyping of immune cells and tumor microenvironment analysis). [1] • Correlate potential predictive biomarkers in tumor and periphery with antitumor response. [1]
Examine potential incidence of immunogenicity by measuring ADAs against translated proteins derived from the IMPs or against PEG lipids.	• Evaluate pre-existing (pre-treatment) and post-treatment ADAs and examine the immunogenicity incidence with treatment. [1]

ADAs = anti-drug antibodies; AUC = area-under-the-concentration-time curve; C_{max} = maximum observed serum concentration; CR = complete response; DCR = disease control rate; DLTs = dose limiting toxicities; DOR = duration of response; IL = interleukin; IMP = investigational medicinal product; MTD = maximal tolerated dose; ORR = objective response rate; OS = overall survival; PD = pharmacodynamics; PFS = progression-free survival; PK = pharmacokinetic; PR = partial response; RECIST = Response Evaluation Criteria in Solid Tumors; RP2D = recommended Phase II dose; SD = stable disease; $t_{1/2}$ = half-life; t_{max} = time to C_{max}; TEAE = treatment emergent adverse event.

1 Analyses of these exploratory endpoints are out of scope of this SAP and may be described in a separate document.

3.2 Trial design

Trial design	This is an open-label, multicenter Phase I/IIa dose escalation, safety, PK and PD trial of BNT151 with expansion cohorts in various solid tumor indications. The monotherapy dose escalation and biomarker cohort (Part 1) of this clinical trial will enroll patients with various solid tumors that are metastatic or advanced unresectable stage for whom there is no available standard therapy likely to confer clinical benefit, or patients who are not candidates for such available therapy. During combination dose escalation (Part 2A), patients with squamous cell carcinoma of head and neck (SCCHN), and hepatocellular carcinoma (HCC) will be enrolled and treated with a combination of BNT151 and pembrolizumab. Once Part 2A of SCCHN and HCC is completed, patients with renal cell carcinoma (RCC), triple negative breast cancer (TNBC), and non-small cell lung cancer (NSCLC) will be enrolled and treated with a combination of BNT151 and the respective standard of care (SoC). The trial consists of Part 1, Part 2A and Part 2B with adaptive design elements: • **Part 1** will be a monotherapy dose escalation in patients with advanced solid malignancies until the MTD and/or recommended Phase II dose (RP2D) of BNT151 as monotherapy are defined. Dose escalation and RP2D regimen may include a pre-conditioning dose, which would be implemented under the rules defined in protocol Section 4.1.4 Adaptive trial design elements. The Part 1 of the trial also plans to implement a dedicated biomarker cohort in BNT151 monotherapy: • The Biomarker Cohort will recruit patients at selected sites in the United States (US) only. The objective of the cohort is to observe PD activity and drug-induced changes in the blood and tumor. The gathered data are expected to inform on the drugs mechanism of action and further refine selection of monotherapy dose. This cohort will only enroll patients that are capable and willing to donate serial biopsies. Patients will only be dosed at dose levels at the RP2D level or lower, which have been cleared safe in the monotherapy dose escalation, and where PD activity is expected. Approximately, 20 patients will be enrolled in this cohort.

	• **Part 2** will start once Part 1 monotherapy dose escalation is completed (exception: Biomarker Cohort in Part 1 can continue enrolling as long as a BNT151 monotherapy dose is established in the dose escalation). It consists of 5 expansion cohorts starting with BNT151 in combination with pembrolizumab in patients with SCCHN (Cohort 1) and HCC (Cohort 2). Safety evaluation of BNT151 in combination with pembrolizumab in Cohorts 1 and 2 will be done at the same time, and data generated in both patient populations will be used to assess safety. Once the safety of the combination is confirmed by the SRC (for details see the SRC Charter), the 2 cohorts will then enroll independently in expansion and a further 3 cohorts evaluating BNT151 in combination with SoC will be opened for enrollment: RCC (Cohort 3), NSCLC (Cohort 4) and TNBC (Cohort 5). Further cohorts may be opened based on sponsor's decision. Each Part 2 cohort will consist of 2 sub-parts: • **Part 2A** where the dose of BNT151 will be established for each combination using either a safety run-in or abbreviated dose finding based on predefined criteria described in protocol Section 4.1.4 Adaptive trial design elements. • **Part 2B** where a predefined number of patients will be treated with the confirmed RP2D of BNT151 in combination with respective SoC. In case the same dosing regimen is used in Parts 2A and 2B (i.e., in case option of safety run-in is adopted), the patients enrolled in Part 2A are eligible for efficacy evaluation in Part 2B. In case a different dosing regimen is used for Part 2A compared to Part 2B (i.e., in case option of abbreviated dose finding is adopted), efficacy generated in Part 2A will be used as supporting data.

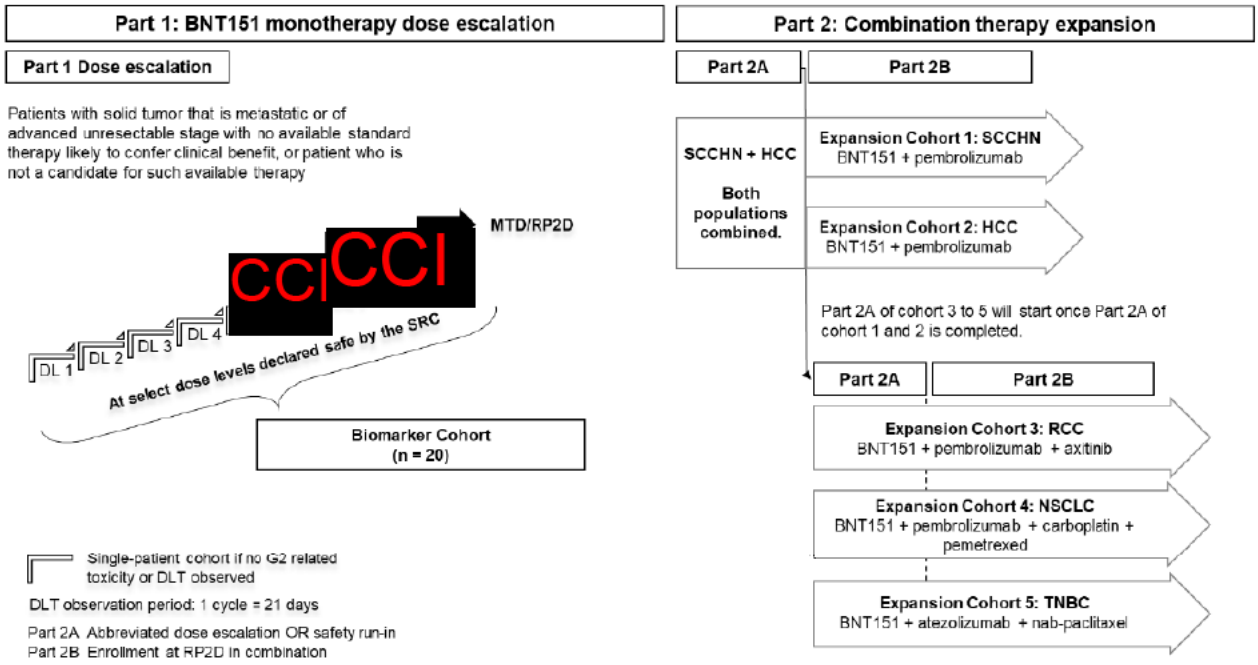

Figure 1: Trial design

The monotherapy dose escalation (Part 1) of this clinical trial will enroll patients with various solid tumors that are metastatic or advanced unresectable stage and for whom there is no available standard therapy likely to confer clinical benefit, or patient who is not a candidate for such available therapy. During combination dose escalation (Part 2A), patients with SCCHN, and HCC will be enrolled and treated with a combination of BNT151 and pembrolizumab. Once Part 2A of SCCHN and HCC is completed, patients with RCC, TNBC, and NSCLC will be enrolled and treated with a combination of BNT151 with the respective SoC.

DL = dose level; DLT = dose limiting toxicity; G2 = Grade 2; HCC = hepatocellular carcinoma; MTD = maximum tolerated dose; NSCLC = non-small cell lung cancer; RCC = renal cell carcinoma; RP2D = recommended Phase II dose; SCCHN = squamous cell carcinoma of head and neck; SoC = standard of care; TNBC = triple negative breast cancer

Trial population	The monotherapy dose escalation (Part 1) of this clinical trial will enroll patients with histologically confirmed solid tumors that are metastatic or unresectable stage for whom there is no available standard therapy likely to confer clinical benefit, or patients who are not candidates for such available therapy. During combination dose escalation (Part 2A), patients with SCCHN, and HCC will be enrolled and treated with a combination of BNT151 and pembrolizumab. Once Part 2A of SCCHN and HCC is completed, patients with RCC, NSCLC, and TNBC will be enrolled and treated with a combination of BNT151 and the respective SoC.
Trial centers and geographic regions	Approximately 7 investigational sites in up to 3 countries including the US, United Kingdom and Spain.
Investigational medicinal product (IMP):	

IMP:	BNT151
Composition:	The IMP is a preservative-free, sterile ribonucleic acid (RNA)-lipid nanoparticle (LNP) dispersion in an aqueous cryoprotectant buffer for intravenous (IV) administration. For further information on the IMP, refer to the pharmacy manual for BNT151-01.
Administration:	IV bolus injection/IV infusion (depending on the administered volume)
Dosage regimen:	In Part 1, BNT151 will be administered IV on Day 1 of each treatment cycle (21 days; or 7 days for the pre-conditioning cycle) after all required procedures and assessments before administration have been completed. Once eligibility is confirmed, administration of BNT151 can be delayed for up to 7 days unless otherwise approved by the sponsor medical monitor. Patients will be administered BNT151 according to dose levels (0.4 µg/kg starting dose to be escalated to CCI µg/kg). BNT151 will be administered as a body weight-based dosing. A cap of dosing for patients weighing 120 kg and more will be implemented in the proposed clinical trial. Patients enrolled in dose escalation of monotherapy with pre-conditioning will start with a pre-conditioning dose (equals either maximum tolerated dose [MTD] for monotherapy without pre-conditioning or the recommended pre-conditioning dose) on Day 1 of pre-conditioning cycle and then receive an intended dose (may follow either the CCI main dose levels or the predefined CCI [redacted] at the discretion of the SRC) on Day 1 of the following treatment cycles.
Duration of treatment:	The Part 1, monotherapy part of the trial will be preceded by an enrollment and screening period of 13 months. Patients are screened at a Screening Visit within 3 weeks prior to the beginning of treatment. During the treatment and dose limiting toxicity (DLT) observation period, patients will be treated with treatment cycles lasting 21 days until progression, or treatment discontinuation of BNT151 due to other factors. Following discontinuation, and a safety follow-up (FU) period of 60 days, patients will be followed up for survival every 12 weeks until death.
Planned number of patients	In Part 1, the sample size will be up to 54 DLT-valuable patients depending on the number of DLTs which may occur, with the possible enrollment of up to 10 additional patients if efficacy is seen in a specific tumor type. In addition, approximately 20 patients will be enrolled in the Biomarker Cohort.

Randomization and blinding	This is a non-randomized open-label trial.
Tumor assessment schedule	Anti-tumor activity assessed by computer tomography or magnetic resonance imaging will be evaluated according to Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 (Eisenhauer et al. 2009) (see Protocol Section 12). Efficacy will be assessed by on-treatment imaging at Week 6 (+7 days), every 6 weeks (±7 days) for 48 weeks, and every 12 weeks (±7 days) thereafter until disease progression is assessed by the investigator (unless the investigator elects to continue treatment), withdrawal of consent, trial termination by the sponsor, or death, whichever occurs first. The RECIST 1.1 criteria will be used for secondary endpoint response evaluation including progression-free survival (PFS) (Eisenhauer et al. 2009). All images obtained must be submitted to the central imaging vendor. Patients who discontinue treatment for reasons other than radiographic disease progression (e.g., toxicity) will continue scheduled tumor assessments at the same frequency as would have been followed if the patient had remained on trial treatment.

3.3 Schedule of activities

See Protocol Section 1.3 for the schedule of activities (SoA).

4 STATISTICAL HYPOTHESES

The primary objectives of this trial Part 1 are to assess the safety profile and to identify the MTD and/or RP2D of BNT151. The Biomarker Cohort examines the exploratory markers for the mechanism of action in the tumor and periphery. Hence, no statistical hypothesis is under test for trial Part 1.

5 INTERIM ANALYSES AND ANALYSIS SEQUENCE

No formal interim analysis is planned for Part 1. However, data will be reviewed after each cohort and at the end of dose escalations.

The main analysis of the trial will be performed based on all available data from a clinical data cut-off that will occur when all patients have been followed up for at least 6 months of treatment or discontinued before.

A final analysis will be performed when the last patient has discontinued from the trial or when the trial is closed by Sponsor.

6 SAMPLE SIZE DETERMINATION

The sample size for Part 1 of the trial is driven by accelerated titration design of single-patient cohorts with a switch to the 3+3 trial design. In Part 1, the sample size is expected to be up to 54 DLT-evaluable patients depending on the number of DLTs which may occur, with the possible enrollment of up to 10 additional patients if efficacy is seen in a specific tumor type.

In addition, approximately 20 patients are planned to be enrolled in the Biomarker Cohort with paired biopsies. This sample size is based on the proportion of patients who have achieved at least 50% increase from baseline in lymphocytes (total lymphocytes, CD4+, CD8+, T_{regs}, and NK cells) upon treatment with BNT151. Twenty patients are estimated to have >80% chance to rule out a proportion of 10% at a 2-sided significance level of 0.05 if the true proportion is 40%.

7 ANALYSIS SETS AND SUBGROUPS

7.1 Analysis sets

The following analysis sets are defined:

Table 3: Definitions of analysis sets

Analysis set	Description
Screened set	The screened set is defined as all patients who signed informed consent.
Treated set	The treated set is defined as all patients who received investigational medicinal product (IMP) (i.e., at least one dose of BNT151).
Safety set	The safety set is defined as all patients who received IMP (i.e., at least one dose of BNT151).
Efficacy evaluable set	The efficacy evaluable set is defined as all patients who are assigned to IMP and have a baseline and at least one on-treatment/post-treatment tumor response assessment. This analysis set may only be used in Part 2 of the trial.
DLT evaluable set	The DLT evaluable set includes all patients from the safety set who are enrolled in dose escalation cohorts and either experienced a DLT during the DLT evaluation period (Cycle 1) or completed the DLT evaluation period. Patients who do not experience any DLT during the DLT observation period are considered to be evaluable if they have been observed for minimum 21 days following the first dose and are considered to have sufficient safety data to conclude that a DLT did not occur.
Pharmacodynamic set	The pharmacodynamic set is defined as all patients with baseline and at least one valid on-treatment/post-treatment follow-up PD assessment.
Pharmacokinetic set	The pharmacokinetic set is defined as all patients with baseline and at least one valid on-treatment/post-treatment follow-up pharmacokinetic assessment.
Immunogenicity evaluable set	The Immunogenicity set is defined as all patients who received IMP and have at least one valid and determinate on-treatment/post-treatment follow up immunogenicity result and have no important protocol deviations as determined by the clinician.

DLT = dose limiting toxicity; IMP = investigational medicinal product; PD = pharmacodynamic.

The DLT evaluable set will be used for the evaluation of DLTs in order to assess the MTD and RP2D. The safety set will be used for all other safety analyses. The treated set will be used for demographics, disease characteristics, medical history, prior and concomitant medication, other medical procedures and efficacy analyses. The screened set and treated set will be used for disposition as applicable.

7.2 Protocol deviations

Protocol deviation management for this trial is detailed in the trial specific Protocol Deviation and Non-compliance Management Plan. According to this plan, protocol deviations or site non-compliance are documented concisely in the Medidata Clinical Trail Management System (CTMS) and periodically reviewed as part of the project oversight by a wider trial team.

Protocol deviations are failures to adhere to the inclusion/exclusion criteria and protocol requirements and will be classified into important protocol deviations and non-important protocol deviations.

- Important protocol deviations are a subset of protocol deviations that may significantly affect the completeness, accuracy, and/or reliability of the trial data or that may significantly affect a patient's rights, safety, or well-being. For example, important protocol deviations may include enrolling patients in violation of key eligibility criteria designed to ensure a specific patient population or failing to collect data necessary to interpret primary endpoints, as this may comprise the scientific value of the trial.
- Non-important protocol deviations are any non-compliance that does not adversely affect the study or process (i.e. does not meet the definition of important).
- Important protocol deviations will be identified by trial team review during the data review prior to database lock.

Additional important protocol deviations may be defined at the Data Review Meeting (DRM) prior to database lock. During the DRM, protocol deviations as specified in this section of the SAP and all protocol deviations as reported in the CTMS system will be evaluated.

7.2.1 Protocol deviations due to COVID-19

Protocol deviations related to COVID-19 are documented in the CTMS with the preface "COVID-19" in the deviation description.

- COVID-19 related protocol deviations are, for example:
- COVID-19, Visit xx on date xx of patient not conducted due to site facility re-organization related to COVID-19 disruption.
- COVID-19, missing Visit xx due to patient affected by COVID-19.

Important protocol deviations and protocol deviations due to COVID-19 will be summarized by treatment group, site and deviation type for the treated set.

All important and non-important protocol deviations will be listed. In addition, a listing of all inclusion and exclusion criteria which were not met, based on the screened set will be generated.

Further, the following listings will be generated for the treated set to report the impact of COVID-19 on patients enrolled in this study:

- Patients impacted by COVID-19 related trial disruption.
- Missing Visits due to COVID-19.

7.3 Subgroups

No subgroup analysis is planned for Part 1.

7.4 Replacement of patients

According to the protocol, patients are replaced in the following scenarios:

- Patient is not evaluable for DLTs (only applies to patients who do not experience a DLT).
- Patient was enrolled in the trial but did not receive a dose of BNT151.

8 STATISTICAL ANALYSES

8.1 General considerations

In general, statistical analysis will be performed by treatment group, Clinical data from dose escalation cohorts (including backfills will be pooled and analyzed for treatment groups defined by the IMP dose levels that are assigned and received on Cycle 1 Day 1 and with pre-conditioning or not.

Continuous variables will be summarized using the following descriptive statistics: number of patients (n), mean, standard deviation, median, minimum and maximum.

Categorical variables will be summarized using absolute and relative frequencies (n and %) of patients in each category. A “missing”’ category will be presented if there is one or more missing value, with exception of by-visit summaries.

Shift tables from baseline will summarize all cases, wherein at least one of time points is available. A “missing” category will be included where applicable. The percentages are calculated with respect to the available patients. Generally, only measurements at scheduled visits will be summarized and included in the tables by visit. Measurements at unscheduled visits will be included only in the listings, unless specified otherwise.

Rates of binary endpoints will be summarized with absolute and relative frequencies (n and %) along with the corresponding two-sided 95% Confidence Intervals (CI) using an exact method. Exact CI for binomial proportions will be derived using the Clopper-Pearson method (Clopper and Pearson, 1934).

Time-to-event-endpoints will be analyzed using Kaplan-Meier methodology (Kaplan and Meier, 1958) by treatment group. Time-to-event endpoints will be censored in accordance with the FDA Guidance "Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics".

The median survival time (including 95% CI) according to Brookmeyer and Crowley (1982) and the first and third quartile will be presented for each treatment group. Survival rates (including two-sided 95% CI based on Greenwood's formula) as well as the number of patients with events, censored and at risk will be displayed for selected time points (e.g., at 3, 6, 12 months, etc).

The time-to-event analysis will be illustrated using Kaplan-Meier plots.

Additional graphic illustrations for efficacy may include: swimmer plot for duration on treatment and overall response assessments, waterfall plot for best percent change of target lesions in sum of diameters from baseline and spider plot for percent change of target lesions in sum of diameters from baseline.

All data collected will be taken into account for the statistical analysis.

All relevant data will be listed. The listings will be sorted by treatment group, patient number and date/time of assessment/event (if applicable). Unscheduled measurements will be included in the listings.

Programming considerations of planned summary tables, figures and listings and templates for each unique table and patient listing are provided in DPP.

8.2 Key definitions

Date/time of first IMP dose:

Is defined as date/time first administration of BNT151.

DLT evaluation period:

The DLT evaluation period is defined as Cycle 1 (21 days). This definition also applies to patients receiving pre-conditioning treatment.

For detailed definition of DLT see protocol Section 6.5.1.

Baseline Definition:

Baseline is defined as last available value prior to the date and time of the first administration of IMP. Unscheduled measurements prior to the first dose of IMP will be considered in the derivation of the baseline values. Assessments performed on the day of the first administration of IMP will be considered for baseline value, unless there is evidence the assessment was performed after trial treatment had been administered.

Change and percent change from baseline:

Change from baseline will be calculated as follows:

- Change from baseline = post-baseline assessment value – baseline assessment value.

- Percent change from baseline = (post-baseline assessment value – baseline assessment value) / baseline assessment value ×100.

Age (in years) at date of informed consent:

Age is derived in the database based on the collected Year of birth as follows:

Age = (Informed consent Date – (01July + Year of birth)) / 365.25.

Age as derived in the database will be analyzed.

Duration:

Duration will be calculated as follows:

Duration (days) = last observation date – first observation date + 1

For conversion of days to months or years the following rules will be applied:

- 1 month = 30.4375 days
- 1 year = 365.25 days

Time from initial diagnosis to enrollment (in months):

Time from initial diagnosis to enrollment (months) will be calculated as:

(Informed consent date – imputed date of initial diagnosis) / 30.4375

Time from most recent disease progression to first IMP dose (in months):

Time from most recent disease progression to first IMP dose will be calculated as:

(Date of first IMP dose– imputed date of most recent disease progression) / 30.4375

Note: this formula will not be applicable to efficacy endpoints such as duration of response (DOR), progression-free survival (PFS) and overall survival (OS).

Trial day:

Trial day is defined as follows:

If date of assessment is on or after the date of first IMP dose:

- Date of assessment/event – date of first IMP dose + 1

If date of assessment is before date of first IMP dose:

- Date of assessment/event – date of first IMP dose

Laboratory values with “< xx” or “> xx”:

Any laboratory assessment values given as “< xx” or “> xx” in the database will be imputed with the numeric value of xx without the sign for the calculation of descriptive statistics and the changes from baseline (e.g., a value of < 1 will be imputed as 1 for the calculations).

For severity grading of the clinical laboratory assessment values, “<” and “>” will be considered when the numeric part “xx” is equal to a cut-off value.

Last date known to be alive:

Last date known to be alive is defined as last documented date in the database for the patient (at which the patient was alive), including all visit dates, assessment dates (if different from visit dates), trial treatment administration dates, tumor assessment dates, end of treatment dates, date of trial completion or early trial withdrawal date, start/stop dates of Aes, start/stop dates of concomitant medications, start/stop dates of new anti-cancer therapies, start/stop dates of radiotherapies/procedures/non-drug therapies, survival FU date (where patients survival status is alive), and data cut-off date (if patients are alive post the data cut-off date), whichever comes the last.

Start of Follow-up Period:

Patients discontinuing from treatment for any reason will have Safety FU visits 30 days (+5 days) and 60 days (±7 days) after the patient receives the last dose of BNT151. If the patient initiates new anti-cancer treatment within 60 days of the last dose of trial treatment, the safety FU visit should be performed prior to starting new anti-cancer treatment. Once new anti-cancer treatment is initiated, the patient will move into Survival FU.

For patients starting new anti-cancer treatment, max(date of start of new anti-cancer treatment, date of last Safety FU visit)+1 will be used as start date of Survival FU.

Survival FU starts 12 weeks after all other trial visits have been completed. Survival FU may be performed as telephone, email, or clinic visit. Planned time points for Survival FU are provided in the schedule of activities, see protocol Section 1.3. Survival FU will continue until the patient dies, or the sponsor closes the study, whichever occurs earlier.

Pooling of centers:

No pooling of centers is planned for this trial.

Conventions for imputing missing/partially missing dates:

When computations on dates are to be performed, incomplete/missing dates will be imputed using the following rules.

- Missing day, month/year present: If the month/year is the same as the month/year of the first IMP administration date, then impute missing start dates with first IMP dose date. Otherwise, for dates corresponding to a start date, impute with the first day of the month and, for dates corresponding to a stop date, impute with the last day of the month.
- Missing month/day, year present: If the year is the same as the year of the first IMP dose date, then impute missing start dates with first IMP dose date. Otherwise, for dates corresponding to a start date, impute with the first day of the year and, for dates corresponding to a stop date, impute with the last day of the year.
- Missing month/day/year: No imputation will be done in general; however, for AE onset date and prior/concomitant medication start date, it will be imputed as date of first IMP dose.
- Whenever imputed start date is after stop date, it should be set to stop date. Whenever imputed stop date is before start date, it should be set to start date.

This imputation method will be applied to the following dates:

- Onset date of AEs (for treatment-emergent purpose), incomplete AE end dates.
- Medical history start date to define prior medical history and concomitant diseases.
- Date of prior/concomitant medication (start dates) to define prior and concomitant medication, incomplete stop dates of prior/concomitant medications.
- Date of initial diagnosis to define time from initial diagnosis to enrolment.
- Date of most recent disease progression to define time from most recent disease progression.

Best Overall Response based on RECIST 1.1:

The following order of tumor response categories will be used to define the best overall response category based on RECIST 1.1, with “Complete Response” is the best category:

Complete Response (CR) – Partial Response (PR) – Stable Disease (SD) – Progressive Disease – Not Evaluable (NE) – Not Applicable (NA).

If no target lesion is present at baseline, the response “Non-CR / non-progressive disease” for non-target lesions post baseline results in the best overall response category “Non-CR / non-progressive disease”.

8.3 Missing data

All reasonable efforts will be made to obtain complete data for all patients. However, missing observations may occur due to patients lost to FU or to non-compliance with required trial visits and/or assessments. Missing data will not be imputed and data analysis will be performed based on the observed values, unless otherwise specified.

8.4 Visit windows

Every attempt should be made to perform evaluations at the designated time point/visit. Visit windows for visits per cycles and for FU visits are defined in the SoA tables. All visits will be summarized according to the nominal visit. For SFU-1 and SFU-2, the following assessments will be summarized with extended time windows (SFU-1 at 30d +15d, SFU-2 at 60d +/- 15d): body weight, physical examination, vital signs, electrocardiogram (ECG), ECOG performance status and laboratory assessments.

8.5 Patient disposition

For the screened set, the number of patients screened and re-screened, number and percentage of patients who are eligible to participate in the trial, screening failures (see protocol section 5.4 for detailed definition) and the main reason for non-eligibility will be summarized. The number of patients screened and the number and percentage of patients who are eligible and treated will be summarized by country and site as well.

For the treated set, the number and percentage of treated patients included in each analysis set (i.e., safety set, efficacy evaluable set, DLT evaluable set, PK set, and Immunogenicity

evaluable set), the number and percentage of patients being excluded and the reasons for exclusion will be summarized by treatment group.

For the treated set, the treatment status will be summarized using the number and percentage of patients having prematurely discontinued and the primary reason for premature treatment discontinuation of the IMP (e.g., AEs, disease progression, death, withdrawal of consent, lost to FU) for each treatment group.

Further, the number and percentage of patients who entered safety FU period and survival FU period as well as end of trial and the primary reason for discontinuation from the trial will be summarized by treatment group.

All patient disposition data will be listed. Patients excluded from any of the analysis sets and the reason for exclusion from the analysis sets will be listed as well.

8.6 Protocol deviations

Important protocol deviations and protocol deviations due to COVID-19 will be summarized by treatment group, site and deviation type for the treated set.

All important and non-important protocol deviations will be listed. In addition, a listing of all inclusion and exclusion criteria which were not met, based on the screened set will be generated.

Further, the following listings will be generated for the treated set to report the impact of COVID-19 on patients enrolled in this study:

- Patients impacted by COVID-19 related trial disruption.
- Missing visits due to COVID-19.

8.7 Baseline characteristics

8.7.1 Demographics and baseline characteristics

Demographic and baseline variables will be summarized for patients in the treated set. Age (years), weight (kg), height (cm), body mass index (kg/m^2), body surface area (BSA) (m^2) and glomerular filtration rate (mL/min/1.73 m^2) will be summarized as continuous data. Age (<65 years, ≥65 - <75 years, ≥75 years), sex (male or female), ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reportable, Unknown), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Not Reportable, Unknown, Other), body mass index (BMI) (<18 kg/m^2, ≥18 kg/m^2 - <25 kg/m^2, ≥25 kg/m^2 - <30 kg/m^2, ≥30 kg/m^2) and smoking history (has never smoked, ex-smoker, current smoker) will be summarized as categorical data.

A listing of demographics and baseline characteristics will be provided for the treated set.

8.7.2 Disease characteristics

Disease characteristics will be summarized for patients in the treated set. Time from initial diagnosis to enrollment (month) will be summarized as continuous data. Initial diagnosis; tumor stage (American Joint Committee on Cancer [AJCC], current version) at initial

diagnosis and at study entry; tumor, nodes and metastasis (TNM) classification at initial diagnosis and at study entry; time from most recent progression (as recorded on the cancer classification and staging case report form [CRF] page) to start of any trial treatment and ECOG PS at baseline will be summarized as categorical data.

Further, on the basis of the initial tumor assessment (based on RECIST 1.1) reported at the Screening visit, the following variables will be reported:

- Disease characteristics – Imaging: target lesions (absent/present), non-target lesions (absent/present), location of target lesions and sum of target lesion diameters.

Disease characteristics data will be listed for the treated set.

8.7.3 Medical history

Medical history data will be coded using Medical Dictionary for Regulatory Activities (MedDRA, version 24.0 or later version). The number and percentage of patients with a medical history (prior medical history and concomitant diseases) will be summarized by system organ class (SOC) and preferred term (PT). This summary will be done by treatment group in the treated set.

Medical history terms with ongoing ticked “No” will be considered as prior medical history, with ongoing ticked “Yes” as concomitant diseases.

A listing of medical history data will be provided for the treated set.

8.7.4 Prior anti-cancer treatment

Prior cancer surgeries will be coded using Medical Dictionary for Regulatory Activities (MedDRA version 24.0 or later version).

Prior systemic cancer therapies will be coded using the World Health Organization Drug Dictionary (WHO-DD) drug codes of the most recent version (WHO-DD Global B3 SEP 2021 or later version) resulting in Anatomical-Therapeutic-Chemical (ATC) codes indicating therapeutic classification.

A table “Prior Cancer Therapy Types” will be generated for the treated set, including the number of patients and percentages of

- Any prior cancer surgery, including information of treatment intent (palliative, curative, diagnostic, unknown) and residual disease (no residual tumor, microscopic residual tumor, macroscopic residual tumor, unknown, not applicable).
- Any prior cancer radiotherapy, including information of site, setting (adjuvant, neoadjuvant, metastatic, palliative, curative, unknown) and intent (palliative, curative, unknown).
- Any prior systemic cancer therapy, including information of setting (adjuvant, neoadjuvant, metastatic) and intent (palliative, curative, unknown).

A table “Prior Systemic Cancer Therapies” will be generated for the line (number) of prior systemic cancer therapies and the following information of the last therapy:

- setting (adjuvant, neoadjuvant, metastatic),

- best response (CR, PR, PD, SD/Non-PD, NE, Unknown),
- disease status at the end of therapy (CR, PR, PD, SD/Non-PD, NE, Unknown),
- reasons for termination of the therapy (Toxicity/Intolerance, Disease Progression, Other, Unknown, Completion of therapy)

The number and percentages of patients with any prior systemic cancer therapy will be summarized by ATC therapeutic class (ATC level 2), ATC pharmacological class (ATC level 3), and standardized medication name for each treatment group in the treated set.

The number and percentages of patients with any prior cancer surgery will be summarized by SOC and PT by treatment group in the treated set.

Prior cancer surgery, prior cancer radiotherapy and prior systemic cancer therapy will be listed for the treated set.

8.7.5 Prior and concomitant medication

All medications will be coded using the WHO-DD drug codes of the most recent version (WHO-DD Global B3 SEP 2021, or later version) resulting in ATC codes indicating therapeutic classification.

Prior and concomitant medications will be defined using start and stop dates recorded relative to the first and last IMP administration dates.

Prior medications will be defined as any therapy taken prior up to (but not including) the start date of first IMP administration.

Concomitant medications will be defined as any medication either ongoing at the start date of first IMP administration or with a start date on or after the first IMP administration date, including those taken during safety FU.

During survival FU, only medications and non-drug therapies related to “Suspected BNT151-related AEs” need to be documented.

The number and percentage of patients taking prior medication, concomitant medications and concomitant antineoplastic medications may be summarized by ATC therapeutic class (ATC level 2), ATC pharmacological class (ATC level 3), and standardized medication name for each treatment group based on the treated set.

Prior/concomitant medications, concomitant antineoplastic medications, and medications taken during survival FU will be separately listed for the treated set.

8.7.6 Other procedures

Concomitant cancer radiotherapy:

During the course of the trial, concomitant cancer radiotherapies after signing the informed consent, including those taken during safety FU, will be listed for the treated set.

Concomitant Procedures and Non-Drug Therapies:

During the course of the trial, concomitant procedures/non-drug therapies, e.g., invasive procedures/surgery, may take place and will be documented on the CRF page

"Procedure/Non-Drug Therapies". Data will be reported within 28 days prior to start of the first administration of IMP up to safety FU-D60 visit. Procedures and non-drug therapies will be coded using MedDRA (version 24.0 or later version).

All concomitant procedures/non-drug therapies will be listed for the treated set.

8.7.7 New anti-cancer therapies

New anti-cancer therapies will be summarized by the type of therapy and the reason for termination. A listing will also be provided for new anti-cancer therapies.

8.8 Efficacy analyses

8.8.1 Primary analysis

No primary efficacy endpoints are specified. The analysis of the primary safety endpoints is described in Section 8.9.1.

8.8.2 Supplementary analyses

Not applicable.

8.8.3 Secondary analyses

8.8.3.1 Secondary efficacy endpoints

Secondary endpoints include evaluation of anti-tumor activity in terms of objective response rate (ORR), disease control rate (DCR), and DOR. All endpoints will be evaluated using RECIST 1.1 criteria and will be assessed by on--treatment imaging at Week 6 (+7 days), every 6 weeks (±7 days) for 48 weeks, and every 12 weeks (±7 days) thereafter until disease progression is assessed by the investigator (unless the investigator elects to continue treatment), withdrawal of consent, trial termination by the sponsor, or death, whichever occurs first. All applicable tumor assessments (including from unscheduled visits) will be used in the analysis.

Objective response rate

ORR is defined as the proportion of patients in whom a CR or PR is observed as best overall response. Patients not meeting the criteria for CR or PR, including those without any post-baseline tumor assessments, will be considered as non-responders.

ORR will be summarized with absolute and relative frequencies along with two-sided 95% Clopper-Pearson CI by treatment group using the treated set.

Best overall response (all categories) will be summarized with absolute and relative frequencies.

Tumor response assessments as documented including all tumor assessments per RECIST 1.1 (target lesion, non-target lesion, new lesion) will be listed. Best overall response and confirmed CR/PR will be listed as well.

For best overall response, only best overall response before start of new anticancer therapy is analyzed (in line with the censoring rules as defined for PFS and DOR in Appendix 4).

Disease control rate

DCR is defined as the proportion of patients in whom a CR or PR or SD (SD assessed at least 6 weeks after the first IMP dose) is observed as best overall response.

DCR will be summarized with absolute and relative frequencies along with two-sided Clopper-Pearson 95% CI by treatment group.

Duration of response

DOR is defined as the time from the first objective response (CR or PR) to the first occurrence of objective tumor progression (progressive disease), or death from any cause, whichever occurs first. Only patients in whom a CR or PR has been confirmed will be analyzed for DOR.

DOR will be analyzed using Kaplan-Meier methodology by treatment group.

Censoring will be applied according to the rules defined in Appendix 4.

DOR expressed in months for patients whose objective response was confirmed CR or PR = (MIN [Date of first documented disease progression, date of death from any cause] – Date of first documented overall response (CR or PR) + 1) / 30.4375 for events

or (Date of last evaluable tumor assessment – Date of first documented overall response (CR or PR) + 1) / 30.4375 in case of censoring.

Time to response

For confirmed responders, time to response (TTR) is defined as the time from the date of the first dose of IMP to the first objective response (CR or PR) which has been confirmed with a subsequent assessment of objective response at least 4 weeks apart.

Time to response will be summarized using descriptive statistics.

TTR expressed in months for patients whose objective response was confirmed CR or PR = (Date of first documented overall response (CR or PR) – date of first IMP dose + 1) / 30.4375

8.8.4 Exploratory analyses

Exploratory endpoints include PK parameters (including but not to be limited to AUC_{0-t}, $AUC_{0-\infty}$, C_{max}, t_{max}, k_{el}, and $t_{1/2}$, see definitions in Appendix 3), anti-tumor and safety indicators. The analysis of these endpoints will be separately documented in a PK analysis plan. Analysis of exploratory efficacy endpoints, PFS and OS, are optional and if analyzed, their analysis is specified in this SAP.

8.8.4.1 Exploratory efficacy endpoints

Exploratory efficacy analyses will be performed using the treated set.

Progression-free survival

The exploratory endpoint PFS is optional and may not be analyzed. If analyzed, the following applies.

PFS, assessed using RECIST 1.1, is defined as the time from first dose of IMP to first objective tumor progression, or death from any cause, whichever occurs first. PFS will be analyzed using the Kaplan-Meier method by treatment group. Kaplan-Meier plots will be generated.

Patients alive and without disease progression at the data cut-off date or patients lost to FU will be censored at the day of their last tumor assessment before trial discontinuation. If no baseline or post-baseline tumor assessment is available, the patient will be censored at the date of the first dose of IMP. Detailed censoring rules are specified in Appendix 4.

PFS expressed in months = (MIN [Date of first disease progression, date of death from any cause] – date of first dose of IMP + 1) / 30.4375 if the patient has a PFS event

or

(Date of last evaluable tumor assessment – date of first dose of IMP + 1) / 30.4375 in case of censoring.

Overall survival

The exploratory endpoint OS is optional and may not be analyzed. If analyzed, the following applies.

OS is defined as the time from first dose of IMP to the date of death from any cause. OS will be analyzed using the Kaplan-Meier method. Kaplan-Meier plots will be generated.

Patients alive or patients lost to FU at the date of analysis cut-off will be censored at the day of their last date known to be alive.

OS expressed in months = (Date of death from any cause – date of first dose of IMP + 1) / 30.4375 for patients with available death date

or (Last date of known to be alive – date of first dose of IMP + 1) / 30.4375 in case of censoring.

Other exploratory endpoints

Analyses of the other exploratory endpoints are out of scope of this SAP and may be

described in a separate document, including:

- Changes in selected cytokines and other soluble innate and adaptive immune system activation markers compared to baseline.
- Changes in systemic and intra-tumoral immune response in blood and tumor tissue compared to baseline (e.g., immunophenotyping of immune cells and tumor microenvironment analysis).
- Correlate potential predictive biomarkers in tumor and periphery with antitumor response.

- Evaluate pre-existing (pre-treatment) and post-treatment ADAs and examine the immunogenicity incidence with treatment.

8.8.5 Further efficacy analyses

No further efficacy analyses are planned.

8.8.6 Interim analysis

No formal interim analysis is planned. However, data will be reviewed by SRC after each dose level; analyses as defined in this SAP may be generated at the end of dose escalation for safety evaluation as well.

8.9 Safety analyses

Safety data analyses include extent of exposure, AEs, occurrence of DLTs during dose escalation in Part 1, clinical laboratory assessments, vital signs and body weight, ECGs, physical examination and ECOG PS. All safety analyses will be based on the safety set and will be summarized by treatment group, except for the DLT(s) which will be based on the DLT evaluable set in Part 1 only.

All safety data will be listed for the safety set.

8.9.1 Primary safety endpoints dose-limiting toxicities and treatment-emergent adverse events

The primary safety endpoints are the occurrence of DLTs, treatment-emergent adverse events (TEAEs) including Grade ≥3, serious, fatal TEAE reported by causal relationship to trial treatment, grade, and seriousness according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 5.0 (CI CTCAE v5.0). The number and percentage of patients with any DLT, TEAEs and/or SAEs will be presented.

Moreover, a patient listing will be provided with all relevant dose exposure data of all patients enrolled into Part 1, and a listing of all recorded DLTs and/or adverse event (AE) will be presented including the reported term and SOC and PT terms coded using MedDRA®, its time of onset, relationship, NCI CTCAE grade, and seriousness including dose exposure data.

The respective analyses are specified in Section 8.9.3.

8.9.2 Extent of exposure

Exposure to BNT151 and each component of the combination treatments will be separately analyzed using the following dose exposure variables:

- Number of cycles

  The pre-conditioning cycle is not counted in the number of cycles.
- Number of infusions
- Actual treatment duration (week) defined as follows:

(date of last dose – date of first dose + planned duration)/7, whereas the planned duration (days) is defined as the planned time between two consecutive administrations. For BNT151, planned duration is 7 days for pre-conditioning cycle and 21 days for Cycle 1 and afterwards.

- Actual treatment duration (month) defined as follows:

  (date of last dose – date of first dose + planned duration)/30.4375, whereas the planned duration (days) is defined as the planned time between two consecutive administrations. For BNT151, planned duration is 7 days for pre-conditioning cycle and 21 days for Cycle 1 and afterwards.

- Planned treatment duration (week) is defined as follows:

  BNT151: Planned duration per cycle is 7 days for pre-conditioning cycle and 21 days for Cycle 1 and afterwards. Without pre-conditioning, planned treatment duration (week) = (number of cycles × 21) / 7, irrespectively if the patient was treated at each cycle or not. With pre-conditioning, planned treatment duration (week) = ((number of cycles +1) × 21 – 14) / 7, irrespectively if the patient was treated at each cycle or not. The pre-conditioning cycle is not counted in the 'number of cycles'.

- Actual cumulative dose (µg) is defined as sum of all administered doses.
- Planned cumulative dose is defined as sum of all planned doses, irrespectively if the patient was treated at each cycle or not.
- Dose intensity is defined as: Actual cumulative dose (µg) / Actual treatment duration (week)
- Relative dose intensity (RDI) is defined as follows:

$$\text{RDI}\ (\%) = \frac{\text{Dose intensity}\ \left(\frac{\mu g}{\text{week}}\right)}{\text{Planned dose intensity}\ \left(\frac{\mu g}{\text{week}}\right)}\ \text{x}\ 100,$$

whereas

$$\text{Dose Intensity}\ \left(\frac{\mu g}{\text{week}}\right) = \frac{\text{Actual cumulative dose}\ (\mu g)}{\text{Actual treatment duration}\ (\text{week})}$$

$$\text{Planned dose Intensity}\ \left(\frac{\mu g}{\text{week}}\right) = \frac{\text{Planned cumulative dose}\ (\mu g)}{\text{Planned treatment duration}\ (\text{week})}$$

Note: For BNT151, dose unit is "µg"; for the components of the combination treatments, a different unit may be used.

Additionally, the relative dose intensity will be presented categorically (i.e., number and percentage of patients with relative dose intensity of < 60%, 60 - < 80%, ≥ 80%). Moreover, the number and percentage of patients with any dose delay, with any dose modification, and any dose interruption will be presented.

All study drug administration data will be listed. All the above-mentioned variables related to actual exposure will be listed as well.

8.9.3 Adverse events

AEs will be coded using the MedDRA® coding system (version 24.0 or later version) to get a SOC and PT for each AE and graded for severity using NCI CTCAE v5.0.

A TEAE is defined as any AE with an onset date on or after the first administration of IMP (if the AE was absent before the first administration of IMP) or worsened after the first administration of IMP (if the AE was present before the first administration of IMP). AEs with an onset date more than 60 days after the last administration of IMP will be considered as TEAEs only if assessed as related to IMP by the investigator. TEAEs will be summarized by treatment group.

AEs started prior to first administration of IMP or AEs after last administration of IMP + 60 days and not assessed as related to IMP (FU AEs) will be included only in the AE listings.

AE leading to permanent study treatment discontinuation

AE leading to discontinuation are the AEs with the items "*Action taken*" ticked "Drug withdrawn".

AE leading to dose reduction

AE leading to dose reduction are the AEs with the items "*Action taken*" ticked "Dose reduced".

AE leading to dose interruption

AE leading to dose interruption are the AEs with the items "*Action taken*" ticked "Drug interrupted".

In case a patient has an AE with missing relationship status, the event will be assumed to be related and associated with the treatment received in the summaries (will be listed as collected in the listings). No imputation for missing NCI CTCAE grades and missing seriousness will be performed.

Dose limiting toxicities

In general, a DLT for a drug or other treatment is defined as an AE that prevents an increase of the dose level of that treatment.

For the purpose of dose escalation in Part 1, the DLT monitoring period will be the 21 days of Cycle 1.

The occurrence of any of the toxicities outlined in protocol Section 6.5 will be considered a DLT, excluding toxicities clearly related to disease progression or intercurrent illness. Serious adverse events (SAEs), non-serious Grade ≥3 AEs and clinically significant abnormal laboratory values Grade ≥3 will be collected and assessed for DLTs. NCI CTCAE v5.0 will be used to grade the intensity of AEs. Other clinically significant toxicities, including a single event or multiple occurrences of the same event, may be considered as DLTs.

Overall summary of adverse events

The number and percentage of patients reporting at least one TEAE will be summarized for each of the following AE types:

- TEAEs
- TEAEs related to treatment (BNT151)
- Grade ≥3 TEAEs
- Grade ≥3 TEAEs related to treatment (BNT151)
- TEAEs related to trial procedure
- Serious TEAEs
- Serious TEAEs related to treatment (BNT151)
- Serious TEAEs leading to death
- Serious TEAEs related to treatment leading to death (BNT151)
- TEAEs leading to dose reduction
- TEAEs leading to dose interruption
- TEAEs leading to permanent discontinuation of treatment
- DLTs

For Dose Escalation, only (serious) TEAEs related to BNT151 are applicable.

Analyses of adverse events

The number and percentage of patients for each category above (except for DLTs) will be summarized by SOC and PT as per MedDRA. If a SOC/PT is reported more than once for a patient, the patient will only be counted once for this SOC/PT. All AE summary tables will be sorted by descending frequency (%) by SOC and by PT within SOC in the total column. The number of events will be summarized in the same table.

The following AE tables by SOC and PT will be generated:

- TEAEs
- TEAEs related to treatment (BNT151)
- Grade ≥3 TEAEs
- Grade ≥3 TEAEs related to treatment (BNT151)
- TEAEs related to trial procedure
- Serious TEAEs
- Serious TEAEs related to treatment (BNT151)
- Serious TEAEs leading to death
- Serious TEAEs related to treatment leading to death (BNT151)
- TEAEs leading to dose reduction
- TEAEs leading to dose interruption
- TEAEs leading to permanent discontinuation of treatment

For Dose Escalation, only (serious) TEAEs related to BNT151 are applicable.

Analyses of adverse events by worst NCI CTCAE grade

Moreover, the number and percentage of patients with any TEAE will be summarized by worst NCI CTCAE grade by PT nested within SOC. If a TEAE is reported more than once by a patient for an SOC/PT, the patient will be counted for the worst grade for this SOC/PT. AEs with a missing grade will be presented in the summary table as a grade category of "missing". TESAEs related to treatment will be analyzed by the worst NCI CTCAE grade as well.

Treatment-emergent adverse events table per ClinicalTrials.gov

Moreover, the number and percentage of patients with any non-serious TEAE will be summarized by SOC and PT.

Deaths

All deaths will be listed. Deaths within 30 days after the last dose of trial treatment will be flagged.

Deaths will be summarized by primary cause for the following categories based on death time after last dose of trial treatment: within 30 days, within 31-60 days, post 60 days.

Serious TEAEs leading to death will be summarized by PT nested within SOC.

Serious TEAEs that are deemed as related to trial treatment and leading to death will be summarized by PT nested within SOC as well.

Adverse events listings

All AEs will be listed. Further, all deaths, SAE, all AEs leading to death, AEs leading to permanent discontinuation of treatment will be listed.

All DLTs will be presented with the reported term, PT and SOC, its time of onset, duration, and outcome, relationship, NCI CTCAE grade, and seriousness based on the DLT evaluable set.

8.9.4 Laboratory assessments

Clinical laboratory data to be summarized include hematology, coagulation factors, blood chemistry, urinalysis and endocrine test (thyroid function tests). The clinical laboratory parameters to be assessed are listed in Table 4 and the scheduled time points for assessment are presented in the SoAs (see protocol Section 1.3).

All laboratory parameters are reported in the clinical database using local safety laboratories. Measurements and corresponding Low and High ranges will be converted into Standard International units. Values outside the normal ranges are classified as "Not clinically significant" or "Clinically significant" by the investigator.

The last available assessment on treatment is defined as the assessment at End of Treatment Visit or the last safety assessment on treatment in case of treatment discontinuation due to death or withdrawal of consent.

The last available assessment on trial is defined as the assessment either at Safety FU-D30, Safety FU-D60 visit or the last safety assessment before discontinuation from the trial.

For the derivation of the minimum post-baseline assessment, maximum post-baseline assessment or worst post-baseline toxicity common toxicity criteria (CTC) grades, assessments collected at scheduled visits and those at unscheduled visits will be considered.

Continuous clinical laboratory parameters at each scheduled visit and change from baseline to each post-baseline time point will be summarized using descriptive summary statistics for each parameter by treatment group.

The number and percentage of patients reporting markedly abnormal clinical laboratory values as reported by the investigator at any point on the trial will be summarized for each parameter by visit and treatment group using the following categories as applicable: Clinically significant low, Low, Normal, High, Clinically significant high.

Shift tables from Baseline to Minimum post-baseline, Maximum post-baseline, Last available assessment on treatment and to Last available assessment on trial with respect to reference range values (as applicable: Clinically significant low, Low, Normal, High, Clinically significant high) will be generated for each parameter by treatment group.

Clinical laboratory results will be classified according to NCI CTCAE v5.0. Laboratory results not corresponding to a NCI CTCAE term will not be graded. CTCAE grades for the respective parameters will be reported using the following categories: Grade 0, Grade 1, Grade 2, Grade 3, Grade 4 by visit and treatment group. Shift tables from Baseline to Worst post-baseline toxicity grade, Last available assessment on treatment, and Last available assessment on trial will be provided for each laboratory parameter, if applicable, by treatment group.

Urinalysis, dipstick results will be reported only by number of patients and percentage for each parameter by visit and treatment group.

All clinical laboratory data will also be presented in the data listings. Abnormal clinical laboratory values and clinically significant values as reported by the investigator will be flagged in the listing. Microscopic urinalysis parameters (where available) and pregnancy assessment data (including childbearing potential) will be listed only.

Table 4: Local safety laboratory tests (blood and urine) – overview

Laboratory assessments	Parameters			
Hematology[1]	Platelet count	Red blood cell (RBC) Indices:MCV MCHC %Reticulocytes	White blood cell (WBC) count with Differential: Neutrophils Lymphocytes Monocytes Eosinophils Basophils	
	Red blood cell RBC count	Hemoglobin	Hematocrit	
Coagulation factors[1]	• Activated partial thromboplastin time • Prothrombin time • International normalized ratio			
Blood chemistry[1,2]	Blood urea nitrogen (BUN) or urea[3]	Potassium	Aspartate Aminotransferase (AST)	Total bilirubin
	Creatinine	Sodium	Alanine Aminotransferase (ALT)	Lactate dehydrogenase
	Glucose [Indicate if fasting, or non-fasting]	Calcium	Alkaline phosphatase	C-reactive protein
	Bicarbonate or carbon dioxide	Chloride	Total protein	Pre-albumin
	Phosphorus	Magnesium	Ferritin	Albumin
Routine urinalysis[1]	• Specific gravity • pH, glucose, protein, blood, ketones, bilirubin, urobilinogen, nitrite, leukocyte esterase by dipstick • Microscopic examination (if blood or protein is abnormal)			
Endocrine tests (thyroid function tests)[1]	• TSH, free-T3 and free-T4			
Pregnancy test (females of childbearing potential only)[1]	• Urine human chorionic gonadotropin (hCG) • Serum β-hCG (at screening only and if a urine hCG is equivocal or positive during the remainder of the trial)			
Serum autoantibody analysis[1]	• Anti-nuclear antibody • Anti-double-stranded-DNA • Circulating anti-neutrophil cytoplasmic antibody Perinuclear anti-neutrophil cytoplasmic antibody			

1 Local laboratory testing.
2 Time points and content of the reduced blood chemistry panel are detailed in Section 1.3.
3 Depending on the country.
MCHC = mean corpuscular hemoglobin concentration; MCV = mean corpuscular volume; T3 = triiodothyronine, T4 = thyroxine; TSH = thyroid stimulating hormone.

8.9.5 Vital signs

Vital sign parameters (blood pressure, pulse rate, respiratory rate, body temperature) and the scheduled visits and time points (pre-administration, during administration or post-administration) for assessments are presented in the SoA (see protocol Section 1.3).

Vital sign parameters at each time point and change from baseline to each post-baseline time point will be summarized using descriptive statistics for each parameter by treatment group.

The last available assessment on treatment is defined as the assessment at End of Treatment Visit or the last safety assessment on treatment in case of treatment discontinuation due to death or withdrawal of consent.

The last available assessment on trial is defined as the assessment either at Safety FU-D30, Safety FU-D60 visit or the last safety assessment before discontinuation from the trial.

Vital sign values for each parameter will be assigned an LNH classification according to whether the value is lower (L), within (N) or higher (H) the reference range for that parameter (Table 5).

Table 5: Normal ranges for vital signs

Parameter	Range
Systolic blood pressure	90-150 mmHg
Diastolic blood pressure	50-100 mmHg
Respiratory rate	12-20 Breaths/min
Pulse rate	50-100 bpm
Temperature (where applicable)	35-38 °C

All vital sign data will be presented in the data listings.

8.9.6 Electrocardiogram

ECGs parameters and the scheduled visits for assessment are presented in the SoA (see protocol Section 1.3).

ECGs will be judged by the investigator as normal, abnormal not clinically significant (NCS) or abnormal clinically significant (CS). The number and percentage of patients with clinically significant ECG findings per visit will be summarized by treatment group.

ECG parameters (HR, PR, QRS, QT, and corrected QT [QTc] intervals) at each visit and change from baseline to each post-baseline visit will be summarized using descriptive statistics treatment group.

Additionally, number and percentage of patients classified into each of the following categories will be summarized by treatment group:

- Maximum QTc interval at post-baseline
  - 450 – 480 ms

  - 481 – 500 ms
  - ≥501 ms
- Maximum increase in QTc interval from baseline
  - 30 – 60 ms
  - >60 ms
- HR ≤50 beats per minute (bpm) at post-baseline and/or decrease in HR from baseline ≥20 bpm
- HR ≥120 bpm at post-baseline and/or increase in HR from baseline ≥20 bpm
- PR ≥220 ms at post-baseline and/or increase in PR from baseline ≥20 ms
- QRS HR ≥120 ms at post-baseline

Results for all ECG parameters will be presented in listings; abnormal QTc values will be flagged.

8.9.7 Echocardiogram (ECHO)

Evaluation of left ventricular function (LVEF) by ECHO scan will be performed during screening, and if clinically indicated at other time points.

All LVEF assessment results (normal, abnormal not clinically significant or abnormal clinically significant) will be listed if available.

Left ventricular function is evaluated by an ECHO scan at screening and unscheduled timepoints. ECHO data will be listed only.

8.9.8 Physical examination

Comprehensive physical examinations will be performed at screening. During the course of the trial, the investigators should make an overall health judgement and make only in-depth examinations if clinically indicated.

Depending on timing (before or after signing the informed consent form for trial participation), the clinically significant findings of the physical examination have to be recorded as medical/surgical history or as AEs.

The complete physical examination at screening includes the following body systems: head, eyes, ears, nose, and throat, and the cardiovascular, dermatological, musculoskeletal, respiratory, gastrointestinal, genitourinary and neurological systems.

At all visits after screening, an abbreviated physical examination will be performed.

The abbreviated physical examination includes an overall health judgement (any new or worsened abnormalities since previous assessment [Y/N]) and brief interim history (change of symptoms). In-depth physical examinations are required if obvious pathological signs are visible or in the case the patient states any signs or symptoms.

The number and percentage of patients with assessments of normal, abnormal NCS, abnormal CS at screening visit as well as assessment about new or worsened abnormalities since previous assessment for each post-baseline visit will be displayed in a summary table.

All physical examination data will be listed.

8.9.9 ECOG performance status

ECOG PS (Table 6) will be assessed according to the SoA (see protocol Section 1.3).

The following ECOG PS grades will be reported:

Table 6: ECOG performance status

GRADE	ECOG PERFORMANCE STATUS
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all selfcare but unable to carry out any work activities; up and about more than 50% of waking hours
3	Capable of only limited selfcare; confined to bed or chair more than 50% of waking hours
4	Completely disabled; cannot carry on any selfcare; totally confined to bed or chair
5	Death

Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol.* 1982; 5: 649-655.
Available from: https://ecog-acrin.org/resources/ecog-performance-status

The number and percentage of patients in each ECOG PS grade will be summarized by visit and by treatment group.

Shift tables from baseline to scheduled post baseline visits with respect to ECOG PS grades by treatment group will be generated. All ECOG PS data will be listed. For each patient, the best and worst post-baseline assessment values will be flagged.

9 QUALITY CONTROL

SAS programs are developed to produce output such as analysis data sets, summary tables, data listings, figures or statistical analyses. An overview of the development of programs is detailed in Syneos Health Developing Statistical Programs SOP (3907).

Syneos Health Developing Statistical Programs SOP (3907), Conducting the Transfer of Biostatistical Deliverables SOP (3908) and the SAS Programming and Validation Plan (version 1.0, dated 25 MAR 2021) describes the quality control procedures that are performed for all SAS programs and output. Quality control is defined here as the operational techniques and activities undertaken to verify that the SAS programs produce the output by checking for their logic, efficiency and commenting and by review of the produced output.

10 REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics. 1982;38:29-41.

Clopper, CJ and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-413, 1934.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-247.

European Medicines Agency, Guideline on the evaluation of anti-cancer medicinal products in man (available at https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-evaluation-anticancer-medicinal-products-man-revision-6_en.pdf).

Kaplan E. L., Meier P. Nonparametric estimation of incomplete observations. Journal of the American Statistical Association. 1958;53:457-481.

Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5:649-655.

Simon R. Optimal two-stage designs for Phase II clinical trials. Control Clin. Trials. 1989;10(1):1-10.

U.S. Department of Health and Human Services, Food and Drug Administration, Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (available at https://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/).

U.S. National Comprehensive Cancer Network (NCCN 2020; available at https//www.nccn.org/professionals/physician_gls/pdf/prostate.pdf).

11 SUPPORTING DOCUMENTATION

Not applicable.

11.1 Appendix 1: Changes to protocol-planned analyses

To evaluate preliminary efficacy in this FIH study, there is no need to perform sensitivity analyses on the censoring rules since with the limited number of patients there is no difference expected. Thus, the censoring per the EMA guidelines stated in the protocol will not be performed.

The following are clarifications and not changes to protocol-planned analyses as they concern exploratory endpoints only.

Analyses of the following exploratory endpoints are out of scope of this SAP and may be described in a separate document:

- Changes in selected cytokines and other soluble innate and adaptive immune system activation markers compared to baseline

- Changes in in systemic and intra-tumoral immune response in blood and tumor tissue compared to baseline (e.g., immunophenotyping of immune cells and tumor microenvironment analysis).
- Correlate potential predictive biomarkers in tumor and periphery with antitumor response.
- Evaluate pre-existing (pre-treatment) and posttreatment ADAs and examine the immunogenicity incidence with treatment.

The following exploratory endpoints from the protocol are optional and may not be analyzed:

- PFS defined as the time from first dose of IMP to first objective tumor progression, or death from any cause, whichever occurs first.
- OS defined as the time from first dose of IMP to death from any cause.
- PK parameters (including but not limited to AUC, C_{max}, t_{max}, and $t_{1/2}$).

11.2 Appendix 2: List of abbreviations

Abbreviation	*Definition*
AE	*Adverse Event*
AJCC	*American Joint Committee on Cancer*
ATC	*Anatomical-Therapeutic-Chemical*
BMI	*Body Mass Index*
bpm	*beats per minute*
BSA	*Body surface area*
CI	*Confidence Interval*
CR	*Complete Response*
CRF	*Case Report Form*
CS	*Clinically Significant*
CTC	*Common Toxicity Criteria*
NCI CTCAE v5.0	*National Cancer Institute Common Terminology Criteria for Adverse Events, version 5.0*
CTMS	*Clinical Trail Management System*
DLT	*Dose-limiting toxicity*
DOR	*Duration of response*
DPP	*Data presentation plan*
DRM	*Data Review Meeting*
ECHO	*Echocardiogram*
ECG	*Electrocardiogram*
ECOG PS	*Eastern Cooperative Oncology Group Performance Status*
FU	*Follow-up*
HCC	*Hepatocellular carcinoma*
IMP	*Investigational Medicinal Product*
LNH	*Low, Normal, High*
LVEF	*Left ventricular function*
MedDRA™	*Medical Dictionary for Regulatory Activities*
NA	*Not Applicable*
NCS	*Not Clinically Significant*
NE	*Not Evaluable*
NCI-CTCAE	*National Cancer Institute - Common Terminology Criteria for Adverse Events*
NSCLC	*Non-small cell lung cancer*
ORR	*Objective Response Rate*
OS	*Overall Survival*
PD	*Progressive Disease*

PFS	*Progression-Free Survival*
PR	*Partial Response*
PT	*Preferred Term*
QC	*Quality Control*
RBC	*Red Blood Cell*
RDI	*Relative dose intensity*
RECIST	*Response Evaluation Criteria in Solid Tumors*
RP2D	*Recommended phase 2 dose*
SAE	*Serious Adverse Event*
SAP	*Statistical Analysis Plan*
SAS	*Statistical Analysis Software*
SCCHN	*Squamous cell carcinoma of head and neck*
SD	*Stable Disease*
SI	*International System of Units*
SOA	*Schedule of Activities*
SOC	*System Organ Class*
SOP	*Standard Operating Procedures*
SRC	*Safety Review Committee*
TEAE	*Treatment-Emergent Adverse Event*
TNBC	*Triple negative breast cancer*
TNM	*Tumor, nodes and metastasis*
TTR	*Time to response*
WHO DD	*World Health Organization Drug Dictionary*

11.3 Appendix 3: List of PK parameters

Parameter	***Definition***
AUC_{0-t}	*Area under the concentration versus time curve from time 0 to time t*
$AUC_{0-\infty}$	*Area under the concentration versus time curve from time 0 extrapolated to infinity*
C_{max}	*The maximum observed concentration*
k_{el}	*Elimination rate constant*
t_{max}	*The time to reach the maximum observed concentration*
$t_{1/2}$	*Terminal half-life*

11.4 Appendix 4: Censoring rules for PFS and DOR

Situation	Date Patient Has Event or is Censored	Situation Outcome
Inadequate baseline assessment (not applicable for DOR)	Start Date (date of first dose of IMP)	Censored
No on-trial assessments (and no death; not applicable for DOR)	Start Date (date of first dose of IMP)	Censored
Progression documented not after 2 or more missed consecutive scheduled/hypothetical assessment	Date of first objective tumor assessment showing objective progression	Progressed (Event)
Death not after 2 or more missed consecutive scheduled/hypothetical assessment, without objective progression	Date of death	Death (Event)
Death or progression after 2 or more consecutive missed scheduled/hypothetical assessments	Date of last adequate assessment with evidence of no progression prior to 2 or more consecutive missed scheduled/hypothetical assessments; if no adequate assessment exists (not applicable for DOR) then censored at date of first dose	Censored
New anticancer therapy started prior to progression or death	Date of last adequate assessment with evidence of no progression prior to new anticancer therapy; if no adequate assessment exists (not applicable for DOR) then censored at date of first IMP dose	Censored
No progression / lost to follow up	Date of last adequate assessment with evidence of no progression; if no adequate assessment exists (not applicable for DOR) then censored at date of first dose	Censored

Method for determining if 2 or more consecutive tumor assessments have been missed:

Subject has 2 consecutive missed response assessments = subject has more than 2*6*7+7+7 = 98 days between actual consecutive response assessments for the first 48 weeks or 6*7+12*7+7+7 = 140 if 1 assessment during first 48 weeks is missing and the next assessment is scheduled after week 48 and missing or 2*12*7+7+7 = 182 if only assessments after week 48 are considered. If the first planned two post-baseline assessments are missing, apply the following. Subject has 2 consecutive missed response assessments = first non-missing post-baseline assessment has happened after trial day 2*6*7+7 = 91.

Note: In 6*7+7+7 the 6*7 is number of days between response assessments during first 48 weeks as per schedule because the assessments are planned every 6 weeks and assessment window is ±7 days.