STATISTICAL ANALYSIS PLAN (SAP)
BNT111-01

Version: Final v7.0 **Date:** 05 Mar 2024
Sponsor: BioNTech SE

Protocol number:	BNT111-01
Protocol title:	Open-label, randomized Phase II trial with BNT111 and cemiplimab in combination or as single agents in patients with anti-PD-1/PD-L1-refractory/relapsed, unresectable Stage III or IV melanoma
Protocol version:	7.0
Protocol date:	09 Feb 2023
Compound:	BNT111, cemiplimab
CRO name:	ICON plc, South County Business Park, Leopardstown, Dublin 18, D18 X5R3, Ireland
SAP author:	PPD Principal Biostatistician

Prepared at ICON Clinical Research by:

PPD PPD
11 Mar 2024 04:57:10 UTC (Z)
REASON: I approve this document
PPD

PPD **Principal Biostatistician** **Date**

Reviewed at ICON Clinical Research by:

PPD
PPD
07 Mar 2024 15:26:26 UTC (Z)
REASON: I approve this document
PPD

PPD **Principal Biostatistician** **Date**

Approved at BioNTech SE by:

PPD PPD
07 Mar 2024 15:26:43 UTC (Z)
REASON: I approve this document
PPD

PPD **Date**

PPD
PPD
07 Mar 2024 15:21:42 UTC (Z)
REASON: I approve this document
PPD

PPD **Date**

TABLE OF CONTENTS

1 SAP APPROVAL

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

2 VERSION HISTORY

Table 1 SAP Version History Summary

SAP version	Approval date	Change	Rationale
Final v1.0	04Nov2021	First version	
Final v2.0	See signature page	Section 8.5.2: The text 'If an AE has a missing stop date, then the AE will be reported as treatment–emergent.'' Was updated to ''If an AE has a missing start date, then the AE will be reported as treatment–emergent Section 8.1: • Added clarification about the denominator that should be used for percentage calculations. • Clarified that for this study confirmation of CR and PR is not required. • Added text to clarify that A separate programming guideline will be created for programming best overall response using RECIST 1.1 and iRECIST criteria Section 8.4.2: Added clarity (from the iRECIST guidelines) about date of occurrence of objective tumor progression. Section 10.5 Appendix 5: Censoring Rules: • Amended 'before' to 'not after' • Added clarity that for iRECIST the date of iUPD is used if progression has been confirmed.	Mistake in previous text. All added for clarification Added for clarification To ensure meaning was understood.
Final v3.0	See signature page	Section 5: added masking of pre-final tables, listings and figures. Section 8.1: • Updated the analysis of the efficacy endpoints ORR, DCR, DOR, TTR, PFS	For information Change to analysis

		using the Add-on Treatment set to only use data from data of first add-on treatment • Added clarification about the data that will be used for analysis of the safety data for the Safety Set and the Add-on Treatment set)	Clarify selection of data
		Section 8.5.1: Added the category > 30 for the categorical variable Number of cycles, as a maximum of 34 cycles is possible.	Omission in previous text
		Section 5.2: Clarification that the following analyses are also within the scope of the interim analysis: – Analyses of ORR described in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. – Analyses of duration of response described in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. – Analyses of DCR described in Section 8.4.2.2 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. – Analysis of TTR described in Section 8.4.2.4 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. – Analysis of PFS described in Section 8.4.2.5 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. For PFS only Kaplan-Meier plots will be provided at the interim analysis. – Depth of response described in Section 8.4.3.2.	Clarification of IA scope due to initially overlooked details on planned analyses regarding secondary and exploratory objectives
		Appendix 5: Censoring rules • Removed ''Non-protocol treatment'' from the reason for censoring for PFS and DOR as this data is not collected in the CRF.	

		• Reworded “Treatment discontinuation due to toxicity or other” to “Treatment discontinuation due to toxicity(due to an adverse event that is related to IMP ie BNT111 or Cemiplimab) or for the ‘other’ category on the CRF” • Removed ‘prior to treatment discontinuation’ from censoring rules • Added calculation for determining if 2 or more consecutive assessments have been missed.”	
Final v4.0	See signature page	Section 3.2: corrected “Patients will be followed for survival every 3 months for at least 48 months after last trial assessment” to “Patients will be followed for survival every 3 months for at least 48 months after first randomization.” As per protocol. Section 8.1: added DEoR (Depth of response) to the list of tumor response endpoints for which general considerations defined in 8.1 apply Section 8.1: added clarification on how to impute subsequent start of new anti-cancer therapy partial dates Section 8.4.1: added clarification that best overall response for a patient is defined as the best response across all responses <u>prior to the start of new anti-cancer therapy</u> for that patient Section 8.4.3.2: • Added clarity that data collected in the CRF is Investigator data. • Added that only data prior to start of new anti-cancer therapy will be used for DeoR analyses • Added that for RECIST 1.1 analysis data collected up to and including first PD will be used for the DeoR analyses. • Added that for iRECIST analysis, <u>data up to and including first occurrence of objective tumor progression (i.e. the date of iUPD) will be used, provided that iCPD is confirmed at the next assessment</u> as defined by iRECIST guidelines will be used for DeoR analyses	Mistake in previous text. Added for clarification and to clarify selection of data Omission in previous text Omission in previous text Added for clarity Clarify of data selection for analyses

		• Added that separate summaries and waterfall plots will be presented for Randomized and Add-on treatment Section 8.5.1: For the definition of planned treatment duration, updated "scheduled date" to "scheduled study day" Appendix 5: clarified that "consecutive response assessments" is based on actual assessments.	Added for clarification
Final 5.0	See signature page	Section 3.1: - Table 3.1: Updated the objectives based on PA 7.0. - Added clarification in endpoints for Biomarker objectives. Section 3.2: Trial Design, Table 3.2: Updated the text "Combination group" as "Arm 1". Section 5.2: - Updated based on PA v7.0, included details of IA2. - Added the analysis separately for IA1 and IA2. - Clarification that the following analyses are also within the scope of the second interim analysis: o Subgroup analysis from Section 7.3. o Exploratory analysis of duration of disease control. Sections 8.4.3.2 & 8.4.3.3. o Analysis of overall survival, section 8.4.3.6. o Analysis of biomarker data. o Section 8.4.3.8. Section 6: Updated based on PA 7.0. Section 7.3: - Added 3 more sub-groups. - Added clarification for "Visceral metastasis" that the tumor assessment will be based on BICR. - Added clarification for "Presence of liver metastasis" that the tumor assessment will be based on BICR. - Added sub-group "Prior CTLA-4 inhibitor (Yes/No)". - Added clarification in subgroup definitions.	Added details based on PA 7.0 Added clarification Added clarification Added details based on PA 7.0 Added clarifications of IA1 and IA2 Added clarifications of IA2 Added details based on PA 7.0 Added details based on PA 7.0 Added clarification Added clarification Added additional subgroup Added clarification

		Section 8.1: - Modified 'stop' to 'start' date - Added text to separately deal with missing/partial dates for AEs for randomized and add-on treatment. - Imputation rules added for irAE, date of diagnosis and date of progression for prior cancer therapies. Rules clarified for medication dates imputation. - Reference to the programming guidelines for iRECIST data using RECIST data. Section 8.4.1: - Added clarity about tumor response of SD: "It is noted that a tumor response of SD can be confirmed from 36 days after first dose of study drug." - Updated Table 8.1 based on PA 7.0. - Updated details of sensitivity analysis for confirmed CR/PR based on RECIST 1.1 (BICR assessment) for primary analysis. - Added Swimmer plots. - Addition of Non-CR/Non-PD category in BICR analyses. Section 8.4.2.2: Addition of Non-CR/Non-PD category in BICR analyses. Section 8.4.2.6: - Added reference to section 8.4.2.3 - Added the following details of swimmer plot "a swimmer plot based on RECIST 1.1 for investigator assessment will be presented as described in section 8.4.1" Section 8.4.3.1: - Added clarification that sensitivity analysis (EMA censoring rules) is not required based on iRECIST for DOR and PFS. - Added the following details of swimmer plot "a swimmer plot based on iRECIST for BICR and investigator assessment will be presented as described in section 8.4.1". - Addition of Non-CR/Non-PD category in BICR analyses. Section 8.4.3.2: - Added new exploratory analysis: Duration of Disease Control for RECIST 1.1.	Added clarification Added clarification Added imputation rules Added guidelines reference Added clarification Added details based on PA 7.0 Added clarification Added plots Added clarification Added clarification Added clarification Added clarification Added clarification Added clarification Added clarification Added new analysis Added clarification

		- Addition of Non-CR/Non-PD category in BICR analyses. Section 8.4.3.3: Added new exploratory analysis: Duration of Disease Control for iRECIST. Section 8.4.3.4: - Added spider plots for percent change of target lesions in sum of diameter from baseline. - Added clarification in DEoR data used. Section 8.4.3.8.1.1: Added Biomarker analysis. Section 8.5.2: Added summaries by PT and SOC for related serious events for both randomized and add-on treatment for BNT111 and cemiplimab separately." Section 10.1: - Added details about censoring for duration of disease control. - Added additional information for missed tumor assessments. - Addition of Non-CR/Non-PD category in BICR analyses. Section 10.5: Appendix 5: Censoring rules – Added the following text for clarification: "If a subject met multiple censoring criteria, the earliest censoring date will be used". Section 10.7: Appendix 7 added. Section 10.8: Addition of Non-CR/Non-PD category in BICR analyses.	Added new analysis Added plots Added clarification Added new analysis Added new analysis Added clarification Added clarification Added clarification Added clarification Added new appendix Added clarification
Final 6.0	See signature page	Section 7.3: Relapsed/refractory classification rules modified. Section 8.1: Clarification to perform KM analysis only if at least 3 subjects in a treatment arm. Section 8.4: Removal of EMA analyses. Section 9.2: Removal of IPDs on Add-on treatment set. Section 9.3.3: List of variables analyzed adjusted.	Definition modified Added clarification Analysis removed Analysis removed Added clarification

		Section 8.4.1: Best overall response on confirmed CR/PR analysis removed for the Add-on treatment set.	Analysis removed
		Section 8.4.2.2: Clarification on patient results used in the analysis.	Added clarification
		Section 8.4.2.3: Clarification on patient results used in the analysis.	Added clarification
		Section 8.4.2.8: Clinically meaningful deterioration analysis removed.	Analysis removed
		Section 8.4.3.4: Clarification that RECIST and iRECIST DEoR summaries will be performed.	Added clarification
		Section 8.4.3.8: Removal of biomarkers analyses performed internally by the Sponsor	Analysis removed
		Section 8.5.1: Clarification on the BNT111 cycle count and clarification on AESI events.	Added clarification
		Section 8.5.2: Clarification on the death period definition.	Added clarification
		Section 8.5.3: Clarification on some data handling.	Added clarification
		Section 9: EMA censoring rules guidances removed.	Analysis removed
		Section 10: Clinically meaningful deterioration analysis removed.	Analysis removed
		Section 10.5: Table 5.1.2 removed and censoring cases and order added in Table 5.1 (Table 5.1.1 renamed to Table 5.1).	Analysis removed and modified
		Section 10.6: Clinically meaningful deterioration analysis removed.	Analysis removed
Final 7.0	See signature page	Section 8.5.1: Wording updated to clarify some derivations	Added clarification
		Section 8.5.3: Abnormal tables will be generated only on urinalysis parameters	Added modification
		Section 10.5: Censoring rules modified.	Added modification
		Section 10.8: No Evidence of Disease category added	Added modification

3 INTRODUCTION

This is an open-label, randomized Phase II trial in patients with anti-PD1-refractory/relapsed, unresectable Stage III or IV melanoma. This statistical analysis plan (SAP) describes the detailed procedures for the planned statistical analyses (except for analysis of Pharmacodynamic data which will be in a separate SAP) for protocol v1.0 dated 08JUL2020, protocol amendment v2.0, dated 03Nov2020, protocol amendment v4.0 dated 16Jun2021, protocol amendment v6.0 dated 12April2022 and protocol amendment 7.0 dated 09Feb2023 to support the deliveries of data presentations (tables, figures and listings (TFLs)) for the Clinical Trial Report (CTR) and the independent data and safety monitoring board (IDMC).

The statistical analyses will be conducted by BioNTech or ICON using SAS® software Version 9.4 or higher.

3.1 Objectives and endpoints

Table 3.1 Trial Objectives

Objectives	Endpoints
Primary objective	**Endpoint**
Demonstrate the anti-tumor activity of BNT111 + cemiplimab (Arm 1) in terms of objective response rate (ORR) according to RECIST 1.1	• ORR defined as the proportion of patients in whom a complete response (CR) or partial response (PR) is observed as best overall response by blinded independent central review (BICR).
Secondary objectives	**Endpoints**
Assess the anti-tumor activity of each single-agent (i.e., BNT111 and cemiplimab, respectively) in terms of ORR according to RECIST 1.1 (**key secondary endpoint**)	• ORR defined as the proportion of patients in whom a CR or PR is observed as best overall response by BICR.
Assess additional measures of anti-tumor activity of BNT111 + Cemiplimab (Arm 1), and each single agent (i.e., BNT111 and cemiplimab) according to RECIST 1.1	• Duration of response (DOR) defined as the time from first objective response (CR or PR) to first occurrence of objective tumor progression (progressive diseases, PD) by BICR or death from any cause (whichever occurs first). • Disease control rate (DCR) defined as the proportion of patients in whom a CR, PR or stable disease (SD; assessed at least 6 weeks [wks] +/- 1 wk after first dose) is observed as best overall response by BICR.

Objectives	Endpoints
Assess additional measures of anti-tumor activity of BNT111 + Cemiplimab (Arm 1), and each single agent (i.e., BNT111 and cemiplimab) according to RECIST 1.1	• Time to response (TTR) defined as the time from randomization to the first objective tumor response (CR or PR) by BICR. • Progression-free survival (PFS) defined as the time from randomization to first objective tumor progression (PD) or death from any cause, whichever occurs first by BICR. • ORR, DOR, DCR, TTR, PFS, as assessed by the investigator.
Assess overall survival of BNT111 + cemiplimab (Arm 1)	• Overall survival (OS) defined as the time from randomization to death from any cause.
Assess the safety and tolerability profile of BNT111 + Cemiplimab (Arm 1), as well as BNT111 and cemiplimab as single agents.	• Occurrence of treatment-emergent adverse events (TEAE) within a patient including Grade ≥3, serious, fatal TEAE by relationship. • Occurrence of immune-related adverse events (irAE). • Occurrence of dose reduction and discontinuation of trial treatment within a patient due to TEAE. • Changes in laboratory parameters compared to baseline. • Occurrence of abnormal laboratory parameters within a patient. • Changes in vital signs parameters compared to baseline. • Occurrence of abnormal vital signs parameters within a patient.
Assess health-related quality of life (HRQoL) of patients treated with BNT111 + Cemiplimab (Arm 1) and of patients receiving BNT111 and cemiplimab as single agents as measured by the European organization for research and treatment of cancer quality-of-Life questionnaire core 30 items (EORTC QLQ-C30).	• Mean changes from baseline in the global health status score of the EORTC QLQ-C30* • Mean changes from baseline in scores of the EORTC QLQ- C30 functional and symptoms scales* • Time to first clinically meaningful deterioration in global health status score as measured by EORTC QLQ-C30 • Time to first clinically meaningful deterioration in symptoms and functioning as measured by EORTC QLQ-C30*
Exploratory Objectives	**Endpoints**
Assess the anti-tumor activity of BNT111 + Cemiplimab (Arm 1) as well as each single agent (i.e., BNT111 and cemiplimab) according to iRECIST.	• ORR, DOR, DCR, TTR, PFS as assessed by BICR and the investigator.
Assess additional measures of anti-tumor activity of BNT111 + Cemiplimab (Arm 1), and each single agent (i.e., BNT111 and cemiplimab) according to RECIST 1.1/iRECIST.	• Depth of response as defined by maximum percentage reduction from screening/baseline* in the size of the tumor as assessed by BICR and the investigator.

Objectives	Endpoints
Assess the anti-tumor activity during add-on BNT111 + Cemiplimab (Arm 2 and 3) according to RECIST 1.1/ iRECIST.	• ORR, DOR, DCR, TTR, PFS as assessed by BICR and the investigator.
Assess overall survival of each single-agent (i.e., BNT111 and cemiplimab, respectively) with or without add-on treatment (Arm 2 and 3)	• OS
Assess systemic induction/expansion of BNT111 antigen-specific T cells.	• Patients in biomarker sub-trial (selected trial sites): Occurrence of de novo induction of increase in expansion of BNT111 antigen-specific T cells under treatment compared to screening/baseline (Arms 1 and 2 and add-on treatment). The endpoint is clarified as: • Occurrence of expansion of pre-existing or de novo induction of T cells specific for BNT111-encoded antigens at pre-specified time points compared to screening/baseline (categorical ex vivo ELISpot data) (Arm 1 and Arm 2 and add-on treatment)

Characterize the pharmacodynamic profile of BNT111 in combination with Cemiplimab (Arm 1 and add-on treatment) and identify potential predictive biomarkers for efficacy.	• BNT111 target antigen expression, potentially outcome-predictive biomarkers, immune signature molecules (e.g., PD-L1 expression), and potentially prognostic factors in archival tumor tissue at screening and pharmacodynamics markers (e.g., cytokine levels) under treatment. Patients in selected trial sites: tumor infiltrating lymphocytes (TIL) and T-cell receptor (TCR) repertoire characterization at screening/baseline and under treatment. Cytotoxicity assays with autologous tumor cell lines and T cells.
	These endpoints are split into the following endpoints.
	  • Occurrence of BNT111 target Tumor-associated antigen (TAA) expression in tumor tissue at baseline [Positive/Negative] (4 targets).
	  • Tumor mutational burden (TMB) at baseline [absolute values].
	  • Other potentially outcome predictive biomarkers. [1]
	  • Continuous Programmed Cell Death Ligand 1 (PD-L1) Tumor Proportion score (TPS) at baseline in percentage.
	  • Categorical PD-L1 expression (Positive/TPS≥1% vs Negative/TPS<1%) at baseline.
	  • Other immune signature molecules. [1]
	  • Peripheral cytokine profiles under treatment over time [absolute values].
	  • B-Raf proto-oncogene (BRAF) mutations at baseline.
	  • Other potentially prognostic factors in archival tumor tissue at screening and pharmacodynamics markers under treatment [1]

Objectives	Endpoints
	• Tumor infiltrating lymphocytes (TIL) and T-cell receptor (TCR) repertoire characterization at screening/baseline and under treatment. [1] • Cytotoxicity assays with autologous tumor cell lines and T cells [1]

* Baseline in the main trial (Arm 1) encompasses all data collected before start of trial therapy on Visit C1D1.
[1] endpoint will be done separately in a biomarker analysis report, which will not be a part of this SAP.

3.2 Trial design

This is an open-label, randomized, multi-site, Phase II, interventional trial designed to evaluate the efficacy, tolerability, and safety of BNT111 + cemiplimab in anti-PD1/PD-L1 refractory/relapsed patients with unresectable Stage III or IV melanoma. The contributions of BNT111 and cemiplimab will be delineated in single agent calibrator arms (Arms 2 and 3). Patients in Arms 2 and 3, who experience centrally verified (BICR) disease progression under single agent treatment, may be offered addition of the respective other compound to the ongoing treatment after re-consent.

Randomization will be stratified by metastatic status (M0, M1a, M1b vs. M1c, M1d) and number of prior lines of systemic treatment (1 vs 2 to 5).

Screening period:

Screening procedures will be performed within 21 days before the start of treatment on Day 1 of Cycle 1.

Treatment period:

Patients will be randomized in a 2:1:1 ratio to Arm 1 (BNT111 + cemiplimab) and calibrator Arm 2 (BNT111 monotherapy) and Arm 3 (cemiplimab monotherapy).

Treatments with BNT111 will be given once weekly for the first 6 weeks, followed by treatments once every 3 weeks (±3 d). Cemiplimab will be given once every 3 weeks (±3 d) throughout the trial.

The treatment period starts on Day 1 of Cycle 1 and continues for up to 24 months or until confirmed progression of disease, withdrawal of consent or unacceptable toxicity. Treatment with cemiplimab and/or BNT111 may be continued through initial radiological disease progression until symptomatic disease progression or unacceptable toxicity.

If one of the compounds needs to be discontinued, single agent treatment with the other can continue upon investigator's and sponsor's agreement.

Efficacy is assessed with regular efficacy assessments every 6 weeks for the first 3 months after treatment start, then every 9 weeks for the next 9 months and every 12 weeks thereafter until progressive disease, start of new anti-cancer therapy, withdrawal of consent, death or lost to follow-up.

Add-on therapy for patients in Arm 2 or 3

Patients who experience centrally verified (BICR) disease progression under monotherapy may continue with add-on therapy after re-consent.

For add-on of BNT111 to cemiplimab, BNT111 will be given once weekly for the first 6 weeks, followed by treatments once every 3 weeks (±3 d). For add-on of cemiplimab to BNT111, cemiplimab will be given once every 3 weeks (±3 d). Treatments will continue up to a total treatment duration (monotherapy and add-on therapy) of 24 months.

If one of the compounds needs to be discontinued, single agent treatment with the other can continue upon investigator's and sponsor's agreement.

Efficacy is assessed with regular efficacy assessments every 6 weeks for the first 3 months after treatment start, then every 9 weeks for the next 9 months and every 12 weeks thereafter until progressive disease, start of new anti-cancer therapy, withdrawal of consent, death or lost to follow-up.

Safety follow up

The treatment period is followed by a safety follow-up period with visits after 30 and 90 days after last treatment to assess adverse events (AEs). All adverse events will be reported until 90 days after the last dose of trial treatment unless a new anti-cancer therapy is started. If the patient is not able to visit the clinic/hospital, the AE assessment will be done by phone call.

Efficacy follow up

Patients whose last disease assessment on trial treatment was CR, PR or SD will continue to have efficacy assessments.

Survival follow up

Patients will be followed for survival every 3 months for approximately up to 48 months of OS follow-up from the randomization of the last patient.

3.3 Schema (graphical representation of the trial)

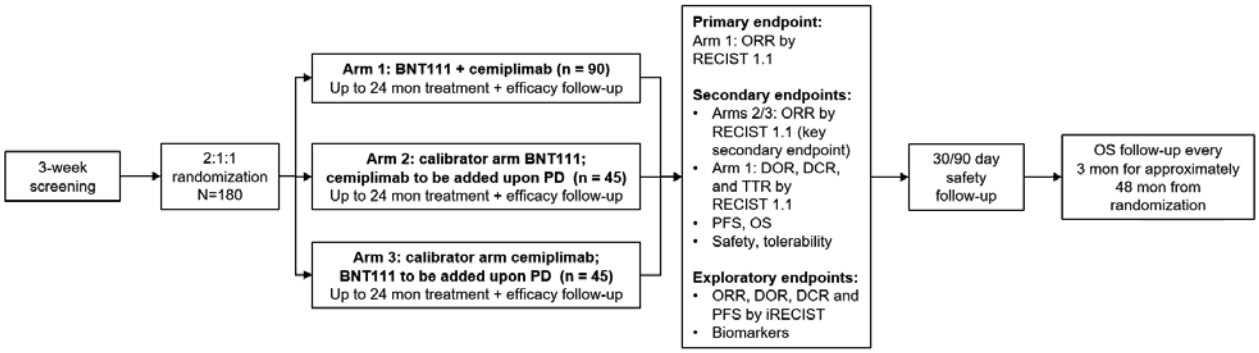

Figure 1: Trial schema

Abbreviations: DCR = disease control rate; DOR = duration of response; mon = months; N = number of patients; ORR = objective response rate; OS = overall survival; PD = progressive disease; PFS = progression-free survival; TTR = time to response.

Table 3.2 Trial design

Trial population	Patients with anti-PD1-refractory/relapsed, unresectable Stage III or IV melanoma
Geographic regions	The trial will take place in about 60 trial sites, in up to 15 countries.
Investigational medical product(s)	**Arm 1**: BNT111 + cemiplimab Name: BNT111 Dose: CCI µg total RNA for the first dose and subsequently at a dose of CCI µg total RNA. Schedule: BNT111 will be given once weekly for the first 6 weeks, followed by treatments once every 3 weeks (±3 d). Route of administration: intravenous (IV) injection Name: cemiplimab Dose: CCI mg Schedule: cemiplimab will be given once every 3 weeks (±3 d) throughout the trial. (cemiplimab will be administered after administration of BNT111 with at least 30 min interval). Route of administration: IV infusion **Calibrator Arm 2**: BNT111 mono Name: BNT111 mono Dose: CCI µg total RNA for the first dose and subsequently at a dose of CCI µg total RNA. Schedule: Same as for the Arm 1 Route of administration: Same as for the Arm 1 **Calibrator Arm 3**: cemiplimab mono Name: Cemiplimab Dose: CCI mg Schedule: Same as for the Arm 1 Route of administration: Same as for the Arm 1 **Add-on treatment:** BNT111 + add-on cemiplimab and cemiplimab + add-on BNT111 Name: Add-on Cemiplimab to BNT111 (Patients in Arm 2): Dose: Same dose of Monotherapy Schedule: Cemiplimab - same as for the Arm 1. BNT111 will be continued according to the planned schedule. A patient must have received a total of 6 weekly injections of BNT111 during the initial trial treatment and the add-on therapy before moving to a once every 3-week schedule of BNT111 injections. Route of administration: Same as for the Arm 1 Name: Add-on BNT111 to cemiplimab (Patients in Arm 3): Dose: Same dose of Monotherapy Schedule: BNT111 will be the same as for the Arm 1. Cemiplimab will be continued according to the planned schedule. Route of administration: Same as for the Arm 1

Treatment and trial duration	**Patient level:** Patients will receive treatment BNT111 + cemiplimab or BNT111 or cemiplimab, single agents up to 24 months and will continue until unacceptable toxicity, withdrawal of consent, discontinuation due to investigator's decision or (confirmed) disease progression. The duration of therapy for patients entering into add-on therapy is a total of 24 months for both the monotherapy and add-on therapy part. **Trial level:** Approximately 18 months of recruitment followed by approximately up to 48 months of OS follow-up from the randomization of last patient.
Planned number of patients	180 patients will be randomized overall (90 patients to Arm 1 (BNT111 + cemiplimab) and 45 patients each to calibrator Arms 2 (BNT111 mono), and 3 (cemiplimab mono).
Randomization and blinding	Automated centralized randomization. Stratification factors: metastatic status: M0, M1a, M1b versus M1c, M1d and by number of prior lines of systemic treatment (1 vs 2 to 5) Treatment Allocation Ratio: 2:1:1 Blinding: Open-label
Tumor assessment schedule:	Tumor response assessments will be performed at screening, every 6 (±1) weeks for the first 3 months after randomization, then every 9 (±1) weeks for the next 9 months and every 12 (±2) weeks thereafter. Until (confirmed) disease progression, death, withdrawal of consent, or until the start of next-line anti-cancer therapy, whichever occurs earlier. Tumor response will be assessed by investigator as well as by blinded independent central review (BICR) of radiographic images and will be measured using RECIST 1.1 and iRECIST. The same schedule will be followed if a patient receives Add-on treatment.
Other features	Per the IDMC charter, regular meetings by an IDMC will be scheduled to review and evaluate safety data. Details on the meeting frequency are described in Section 5.1. of this SAP. Review of Interim analysis will occur per Section 5.2 of this SAP.

3.4 Schedule of activities

See Appendix 4 for the schedule of activities.

4 STATISTICAL HYPOTHESES

The primary endpoint is ORR and the primary statistical objective is to demonstrate that the ORR of BNT111+cemiplimab is greater than CCI using an overall significance level of CCI

CCI

Regarding Arms 2 & 3, i.e., single agent BNT111 and cemiplimab, respectively, the primary endpoint ORR will be analyzed descriptively in a non-confirmatory sense in order to obtain an estimate for the single agent anti-cancer activity. No formal statistical comparisons between treatment groups will be performed and no statistical hypothesis is defined.

~~5~~ INTERIM ANALYSES, ANALYSIS SEQUENCE AND MASKING

A study specific procedure (SSP) has been created to describe the masking that will be used for the pre-final tables, figures and listings (TFLs) deliveries to BioNTech SE. For interim analyses, pre-specified BioNTech personnel will receive unmasked TFLs.

5.1 Independent Data Monitoring Committee (IDMC)

Safety data will be periodically reviewed and evaluated by an IDMC at regular review meetings. The frequency will be based on the number of enrolled patients and will occur at least once every 3 months for the first year and every 6 months thereafter.

In addition, an IDMC safety review meeting is scheduled to occur once 10 patients in Arm 1 (BNT111+cemiplimab combination therapy) have been treated for at least two cycles (six weeks) or have discontinued treatment for any reason during the first 2 cycles.

Specific stopping criteria will be based on the cumulative incidence of trigger events (TEs). TEs include all related serious adverse events (SAEs), related Grade 3 to 5 TEAEs/irAEs except for transient Grade 3 flu-like symptoms, clinically insignificant laboratory abnormalities and Grade 3 hypotension lasting for less than 6 h.

Furthermore, the occurrence of a single grade 5 TEAE (unless clearly unrelated to study drug e.g., due to underlying disease) or a cumulative incidence of ≥ 40% of Trigger Events during the first two cycles of the first 10 patients in Arm 1 will trigger an ad hoc meeting of the IDMC as well as an enrollment pause.

Data presentations described in Sections 8.2, 8.3.1, 8.3.2, 8.3.3, 8.5.1, 8.5.2, 8.5.3 will be provided at each regular review IDMC meeting. Presentations provided for the safety review meeting and ad hoc meeting are described in the IDMC charter.

5.2 Interim analysis

CCI

CCI The IA1 for efficacy has been performed after there were CCI of the planned number of CCI patients. IA1, which was based on CCI patients in the BNT111+cemiplimab group has been performed at a one-sided CCI significance level and requires an ORR of at least CCI to show statistical significance.

IA2 will be performed after approximately CCI of the planned number of CCI patients are response evaluable (i.e., having at least two post-screening tumor assessments or discontinued before).

The operational characteristics of interim analysis (IA1 and IA2) of ORR are summarized in Section 8.4.1. Analysis for the BNT111 monotherapy and cemiplimab monotherapy arms as described in Section 8.4.2.1 will also be presented.

The minimal statistically significant ORR at IA2 will be CCI i.e., the smallest ORR that will result in a statistically significant p value of CCI

The final analysis for efficacy (i.e., the primary analysis of the trial) will be performed once all the planned CCI patients in Arm 1 are response evaluable (i.e., having at least three post-screening tumor assessments or discontinued before). The minimal statistically significant ORR at final analysis would be CCI i.e., the smallest ORR that will result in a statistically significant p value of CCI

The following analyses have been provided at IA1.

- Analyses of ORR described in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST.
- Analyses of DOR described in Section 8.4.2.3 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST.
- Analyses of DCR described in Section 8.4.2.2 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1 using investigator assessment based on iRECIST.
- Analysis of TTR described in Section 8.4.2.4 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1 using investigator assessment based on iRECIST.
- Analysis of PFS described in Section 8.4.2.5 using BICR based on RECIST 1.1, in Section 8.4.2.6 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1 using investigator assessment based on iRECIST. For the first interim analysis only Kaplan-Meier plots were provided.
- Depth of response described in Section 8.4.3.2.

The following analyses will be provided at IA2:

- Subgroup analyses in Section 7.3 (the subgroup "Best response to prior PD-1 + CTLA-4 combination inhibitor therapy (CR, PR, SD, PD)" and "Best response to prior PD-1 inhibitor therapy (CR, PR, SD, PD)" are excluded.
- Analyses of ORR described in Section 8.4.3.5 using BICR based on RECIST 1.1, in Section 8.4.2.6/Section 8.4.3.5 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1/Section 8.4.3.5 using BICR/investigator assessment based on iRECIST.
- Analyses of DOR described in Section 8.4.2.3/Section 8.4.3.5 using BICR based on RECIST 1.1, in Section 8.4.2.6/Section 8.4.3.5 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1/Section 8.4.3.5 using BICR/investigator assessment based on iRECIST.
- Analyses of DCR described in Section 8.4.2.2/Section 8.4.3.5 using BICR based on RECIST 1.1, in Section 8.4.2.6/Section 8.4.3.5 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1/Section 8.4.3.5 using BICR/investigator assessment based on iRECIST.
- Exploratory analysis of duration of disease control (DDC) in Section 8.4.3.2 using BICR/Investigator assessment based on RECIST 1.1 and in Section 8.4.3.3 using BICR/Investigator assessment based on iRECIST.
- Analysis of TTR described in Section 8.4.2.4/Section 8.4.3.5 using BICR based on RECIST 1.1, in Section 8.4.2.6/Section 8.4.3.5 using investigator assessment based on RECIST 1.1, in Section 8.4.3.1/Section 8.4.3.5 using BICR/investigator assessment based on iRECIST.
- Analysis of PFS described in Section 8.4.2.5/Section 8.4.3.5 using BICR based on RECIST 1.1, in Section 8.4.2.6/Section 8.4.3.5 using investigator assessment based on RECIST 1.1 and in Section 8.4.3.1/Section 8.4.3.5 using BICR/investigator assessment based on iRECIST.
- Depth of response described in Section 8.4.3.2.
- Analysis of Overall survival in Section 8.4.2.7/8.4.3.6
- Analysis of Biomarker data Section 8.4.3.8.1

6 SAMPLE SIZE DETERMINATION

The original sample size calculation for the BNT111+cemiplimab group (Arm 1) is based on the following specifications and assumptions:

CCI [redacted]

CCI [redacted]

Based on these specifications at least CCI [redacted] CCI [redacted] CCI [redacted] patients are planned to be randomized into the BNT111 + cemiplimab group. The assumed true effect size for ORR is based on preliminary data of the ongoing Phase I Lipo-MERIT trial with a selected population of patients with melanoma.

After the originally planned interim analysis (first interim analysis [IA1]), where there were CCI [redacted] patients in the BNT111 + cemiplimab group, one additional interim analysis (second interim analysis [IA2]) will be performed. IA2 will assess efficacy after CCI [redacted] of the planned number of CCI [redacted] patients are response evaluable (i.e., having at least two post-screening tumor assessments or discontinued before) in the BNT111 + cemiplimab group. Given CCI [redacted] CCI [redacted]

The sample size for each calibrator arm (i.e., single-agent BNT111 and cemiplimab, respectively) CCI [redacted] CCI [redacted]

7 ANALYSIS SETS AND SUBGROUPS

7.1 Analysis sets

The following analyses sets are defined:

Table 7.1 Analysis Sets

Analysis Set	Description
Screened	The screened set is defined as all patients who signed informed consent.
Intent to treat	The Intent To Treat (ITT) set is defined as all patients who are randomized.
Modified intent to treat	The modified Intent To Treat (mITT) set is defined as all patients who are randomized, had at least one dose of trial treatment and have a baseline and at least one post-randomization tumor response assessment.
Add-on Treatment Set	The Add-on Treatment Set is defined as patients who received at least one dose of Add-on treatment
Safety	The safety set is defined as all patients who received trial treatment (i.e., at least one dose of BNT111 or cemiplimab).
Per protocol	The Per protocol Set (PPS) is defined as all patients who received trial treatment and fulfil the following criteria: • The absence of any important protocol deviations that could affect the primary efficacy analysis. • The completion of a minimal exposure to the treatment of 1 cycle i.e. the patient should receive all the treatment doses planned as per schedule of assessment for at least 1 cycle. • Availability of baseline and at least one on-treatment / post-randomization tumor assessment. Important deviations will lead to an exclusion of patients from the PPS and will be agreed at the data review meeting prior to database snapshot for the primary analysis.

7.2 Protocol deviations

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

- IPDs are those that are considered to have a significant effect on the treatment efficacy and hence would exclude the patient from the Per protocol set.
- Non-important protocol deviations are those that are not considered to significantly affect the efficacy evaluation and hence do not warrant patients' exclusion from the Per protocol set.
- IPDs will be identified by medical review prior to unblinding/database snapshot for main analysis.

The following criteria might be considered as important protocol deviations:

- Violation of important inclusion or exclusion criteria
- Assignment to incorrect treatment/dose (i.e., actual treatment/dose taken differs from the randomized/scheduled)
- Non-compliance (i.e., relative dose intensity less than 80% or greater than 120%)
- Intake of prohibited concomitant medication/occurrence of prohibited medical procedures

IPDs will be presented in a listing. The number and percentage of patients with each IPD type will be summarized by treatment group. The number of patients excluded from the per protocol set will be summarized by treatment group.

7.3 Subgroups

Subgroup analyses will be performed for the primary endpoint (ORR based on BICR assessment using RECIST 1.1) and secondary endpoints DCR and PFS based on BICR assessment using RECIST 1.1. The analysis will be performed if the number of patients is substantial within each subgroup category. Substantial number of patients is defined as at least 10 patients in each of the subgroup categories for the BNT111 + cemiplimab treatment group. If there is a substantial number of patients for the BNT111 + cemiplimab treatment group then the analyses for the BNT111 and cemiplimab monotherapy treatment groups will be performed regardless of the number of patients in each subgroup category. The category "missing" will not be considered for the subgroup analysis.

The analysis will be performed for the BNT111 + cemiplimab, BNT111 and cemiplimab treatment groups using the CCI set.

The following subgroups, based on data collected before start of study treatment will be defined:

- Age group (Age < 70 years, Age ≥ 70 years)
- Sex (Male, Female)
- ECOG PS (ECOG PS 0, ECOG PS 1)
- BRAF mutation status (Mutated and Wild Type)
- Metastatic status ((M0, M1a, M1b), (M1c, M1d))

  Number of prior therapy lines (1, 2 to 5). These subgroup categories will be identified from the prior cancer therapy CRF page.

- Anti-PD1 (programmed death protein 1) refractory/relapsed (Relapsed or refractory to prior PD-L1 treatment) (Yes, No)

  These subgroup categories will be identified from the prior cancer therapy data by medical review. The categorization will be based on the last therapy line where anti-PD-1 or anti-PD-L1 was given. If the best response during the therapy is CR, PR or SD/non-PD (or disease status at the end of the therapy if best response is

missing) then the patient is considered as relapsed, otherwise it is considered as refractory.

If best response is "No evidence of disease" (or disease status at the end of the therapy if best response is missing) and if disease progression is within 3 months from end date of last line then it is considered refractory, otherwise it is relapsed (i.e. date of progression > last anti-PD-1/anti-PD-L1 end date + 3*30.4375). If the date of progression is partial and cannot be imputed then the patient will be excluded from the subgroup analysis.

- Prior PD-1 inhibitor therapy (monotherapy, combination)

  These subgroup categories will be identified from the prior cancer therapy data by medical review.

- Prior BRAF/MEK inhibitor therapy (Yes, No)

  These subgroup categories will be identified from the prior cancer therapy data using ATC codes.

- Prior chemotherapy (Yes, No)

  Chemotherapy will be identified from the prior cancer therapy data using ATC codes.

- Brain metastasis (Yes, No)

  This will be selected using Target Lesion Assessment and Non-Target Lesion Assessment data where sub-location = Brain

- Visceral metastasis (Yes, No)

  This will be identified using Target Lesion Assessment and Non-Target Lesion Assessment data based on BICR assessment. If any of the site/any lesion is reported as having “visceral location” then visceral metastasis =yes.

- Time since the first diagnosis (>median time since the first diagnosis vs <=median time since the first diagnosis). The median time since the first diagnosis will be derived by treatment group. Time since the first diagnosis will be calculated as: Date of informed consent - Date of first diagnosis + 1.

- Sum of target lesion at baseline (>median sum of target lesions at baseline vs <=median sum of target lesions at baseline). The sum of target lesion at baseline will be derived by treatment group.

- Presence of liver metastasis (Yes, No) (target or non-target)

  This will be identified using Target Lesion Assessment and Non-Target Lesion Assessment data based on BICR assessment. If any of the site/any lesion is reported as having “Liver location” then liver metastasis = yes.

In addition, the below subgroups will also be considered:

- Use of systemic antibiotics (Yes, No)

Use of systemic antibiotics will be determined from the concomitant medication data. Use of systemic antibiotics during the treatment period will be checked.

- Use of systemic corticosteroids (Yes, No)

Use of systemic corticosteroids will be determined from the concomitant medication data. Use of systemic corticosteroids during the treatment period will be checked.

- Occurrence of irAE during the treatment and Safety follow-up period (Yes, No)
- Best response to prior PD-1 inhibitor therapy (CR, PR, SD, PD)

Note: In case of re-challenge (≥2 PD1i treatment received by patient prior to enrolment) "best response to PD1i" is defined as best response to last PD-1 inhibitor therapy.

These subgroup categories will be identified from the prior cancer therapy data using ATC codes and best response to the therapy. Only patients who received PD-1 + CTLA-4 in combination will be included in this subgroup analysis:

- Best response to prior PD-1 + CTLA-4 combination inhibitor therapy (CR, PR, SD, PD)

The best response of the last PD-1 + CTLA-4 combination treatment taken before the first study trial treatment will be used.

- Prior CTLA-4 inhibitor (yes, no)

These subgroup categories will be identified from the prior cancer therapy data and medical review.

Analysis detailed Sections 8.4.1, 8.4.2.1, 8.4.2.2 and 8.4.2.5 will be performed. All the subgroup analyses are exploratory, no adjustment for multiplicity will be performed. A p-value will not be presented.

8 STATISTICAL ANALYSES

8.1 General considerations

In general, the statistical analysis of efficacy will be performed by treatment group. Exploratory analysis will be performed for patients in the calibrator arms who went on to receive add-on treatment.

No formal statistical comparisons between treatment groups will be performed.

Continuous variables will be summarized by treatment group and total (as specified) using the following descriptive statistics: number of patients with non-missing data (n), mean, standard deviation (SD), median, minimum (min) and maximum (max).

Categorical variables will be summarized by treatment group and total (as specified), presenting absolute and relative frequencies (n and %) of patients in each category and the number of patients with missing data (missing' category will be presented if there is one or more missing value). The "(N=xx)" in the tables denotes the number of patients in the

analysis set within each treatment group. For event-driven occurrence data (e.g., adverse event, concomitant medication, etc.) the percentages will be based on the number of patients in the analysis set (N). For reported visit data (e.g., gender, ECOG, etc.) the percentages will be based on the number of patients with non-missing values (n).

The rates of binary endpoints will be summarized by treatment group with absolute and relative frequencies (n and %) along with two-sided 95%-confidence intervals (CI) using an exact method. Exact confidence intervals for binomial proportions will be derived using the Clopper-Pearson method (Clopper, CJ and Pearson, ES[5]). CCI [redacted] CCI [redacted]

Time-to-event-endpoints will be analyzed using Kaplan-Meier methodology by treatment group and censored in accordance with the FDA Guidance: "Clinical Trial Endpoints for the Approval of Non-Cell Lung Cancer Drugs and Biologics[2] (Table 5.1). The main analysis will be done for each treatment group separately. Comparison of the treatment groups may be done as an exploratory analysis, with stratified and non-stratified analysis considered.

The overall number and type of events and the overall number and reason for censoring will be summarized (n, %) by treatment group. The median event/survival time (including two-sided 95%-confidence limits according to Brookmeyer and Crowley) and the first and third quartile will be presented. Survival rates (including two-sided 95% confidence intervals based on Greenwood's formula) as well as the number of patients with events, censored and under risk will be displayed for the selected time points.

The time-to-event analysis will be illustrated using Kaplan-Meier plots by treatment group and by treatment and randomization strata. Kaplan-Meier analysis and plots will be generated only if there are at least 3 patients in one treatment group.

Tumor response will be assessed by investigator as well as by blinded independent central review (BICR) of radiographic images and will be measured using RECIST 1.1 and iRECIST. The primary analysis will be based on BICR using RECIST 1.1, while investigators assessment and assessments based on iRECIST (BICR and Investigator) will be considered as secondary/exploratory analysis.

The iRECIST data will be imputed as per the programming guidelines. A separate programming guideline has been created for imputing iRECIST data for BICR assessment and investigator assessment using the RECIST 1.1 assessment.

The best overall response (required for the analysis in Sections 8.4.1, 8.4.2.1, 8.4.2.4 and 8.4.2.6) for the BICR assessment and investigator assessment (using RECIST 1.1 and iRECIST) will be programmatically determined using the RECIST 1.1 and iRECIST guidelines. A separate programming guideline has been created for programming best overall response using RECIST 1.1 and iRECIST criteria. Confirmation of CR and PR is not required for the analysis, unless otherwise specified.

The primary and secondary efficacy analyses and analysis of efficacy data will be performed using the CCI [redacted] set. All data for patients in the CCI [redacted] set will be used for the analysis, even if the patient went on to receive add-on treatment, unless specified otherwise. CCI [redacted] CCI [redacted]

CCI The Safety set will be used for all safety analyses. The Add-on Treatment set will be used for analysis of both safety and efficacy data. Patients randomized to monotherapy (arm 2 and arm 3) who enter add-on treatment period are included in the analysis using both the mITT set, ITT set and the Add-on Treatment set, unless specified otherwise.

The following will occur for the analysis of ORR, DCR, DOR, TTR, PFS, OS, DEoR and EORTC endpoints and for the exploratory analysis of Duration of Disease Control (DDC):

- Tumor response endpoints ORR, DCR, DOR, DDC, TTR, DEoR and PFS: for the analysis using the CCI set or CCI set (sensitivity analysis for ORR (including DCR)), only the data until the day before first dose of add-on treatment will be used. For the analysis using the Add-on Treatment set, only the data from date of first dose of add-on treatment will be used.
- OS: for analysis using the CCI set and Add-on Treatment set, all data would be used. For the exploratory analysis using the Add-on Treatment set (BNT111 + add-on cemiplimab and cemiplimab + add-on BNT111 treatment groups) the data from first day of add-on treatment would be used.
- For EORTC endpoints (Section 8.4.2.8) the summaries/analyzes will be performed as follows:
  - (i) While using CCI set, data until the day before first dose of add-on treatment would be used.
  - (ii) In addition, a separate summary from first dose of add-on treatment to EOS, for the Add-on Treatment set.

All safety analyses will be based on the treatment actually received by the patient ('as treated'). All other analyses will be based on the treatment the patient was randomized to ('as randomized'), unless stated otherwise.

The following will occur for the analysis of safety data (apart from Adverse events)

- For the analysis of safety data using the Safety set, only data until the day before first dose of add-on treatment, if there is add-on treatment, or all data if there is no add-on treatment will be used.
- For the analysis of safety data using the Add-on Treatment set, only data from date of first dose of add-on treatment will be used.

The primary analysis of the trial will be performed based on all available data from a clinical data cut-off that will occur when all patients are randomized and response evaluable (i.e., having at least 3 post-baseline tumor assessments or discontinued before).

An analysis update will be performed when the last patient discontinued from the trial and once all patients have been randomized and have been followed-up for at least 24 months. An analysis update may be subsequently performed.

Data up to and including the clinical cut-off date will be taken into account for the statistical analysis.

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

Baseline is defined as last non-missing value prior to the first dose of trial treatment at C1D1. Assessments performed at C1D1 will be considered for baseline value, unless there is evidence the assessment was performed after trial treatment had been administered. For the add-on treatment if the baseline visit and AOC1D1 are on the same day, the assessment only needs to be taken once. In this case the AOC1D1 would be considered as the baseline value, unless there was evidence that the assessment was done after treatment had been administered.

To identify the baseline value, generally, if there are multiple assessments available for a particular day, the assessment that is closest to the day (and time if collected) of the first dose of the trial treatment (i.e., BNT111 or cemiplimab) will be used as the baseline value in the summary/analyses.

Multiple values at post-baseline visits: if there are multiple post-baseline assessments available within a visit/day, the first value and corresponding date/time will be considered.

Change from baseline will be calculated as follows:

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

If either the baseline or post-baseline assessment value is missing, the change from baseline is set to missing as well.

Duration (Other than "trial treatment duration" defined in Section 8.5.1) will be calculated as follows:

- Duration = 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

Trial Day and Treatment Day are defined as follows:

- Trial day:
    - If assessment date < randomization date then trial day = assessment date – randomization date
    - If assessment date >= randomization date then trial day = assessment date – randomization date + 1
- Treatment Day = treatment date – date of first dose +1

That is, trial day 1 indicates the date of treatment initiation.

All trial data will be presented at the visit as collected in the eCRF. No windowing of data for the purposes of assigning to a visit will occur.

Unscheduled assessments will be included in the analysis of efficacy endpoints for tumor response, time to event endpoints and shift from baseline to worst NCI-CTCAE Grade laboratory summary tables. All other unscheduled assessments will be listed only and not included in summaries/analysis.

If there is a difference between randomized and actual strata at randomization then the randomized strata will be used for analyses. A summary of randomized and actual strata at randomization will be presented. If there is a notable difference between the two summaries, then a sensitivity analysis for the primary endpoint will be considered.

Handling missing Data:

For the purposes of assigning treatment-emergent flag for AEs, partial or missing AE dates will be handled as follows:

- If the day of the month is missing, the onset day will be set to the first day of the month unless it is the same month and year as the first trial *treatment*. In this case, in order to conservatively report the event as treatment-emergent, the onset date will be assumed to be the date of the first trial treatment.
- If the onset day and month are both missing, the day and month will be assumed to be January 1, unless the event occurred in the same year as the first trial treatment. In this case, the event onset will be assumed to be the day and month of treatment in order to conservatively report the event as treatment-emergent.
- A completely missing onset date will be assumed to be the day of the first trial treatment.
- If an AE has a partial start date that is in the same month/year as the first dose of trial treatment, then the AE will be reported as treatment–emergent. If an AE has a missing start date, then the AE will be reported as treatment–emergent.

For the purposes of assigning treatment-emergent flag for irAEs, partial or missing irAE dates will be handled as follows:

- If the day of the month is missing, the onset day will be set to the first day of the month unless it is the same month and year as the first trial treatment of cemiplimab. In this case, in order to conservatively report the event as treatment-emergent, the onset date will be assumed to be the date of the first trial treatment of cemiplimab.
- If the onset day and month are both missing, the day and month will be assumed to be January 1, unless the event occurred in the same year as the first trial treatment of cemiplimab. In this case, the event onset will be assumed to be the day and month of treatment of cemiplimab in order to conservatively report the event as treatment-emergent.
- A completely missing onset date will be assumed to be the day of the first trial treatment of cemiplimab.

- If an irAE has a partial start date that is in the same month/year as the first dose of trial treatment of cemiplimab, then the irAE will be reported as treatment–emergent. If an irAE has a missing start date, then the irAE will be reported as treatment–emergent.

For the purposes of assigning prior or concomitant flag for medications, partial or missing medication dates will be handled as follows:

- If end day is missing and month/year are non-missing, then day is the earliest of treatment end date (if month/year of treatment end date is same as the medication month/year) and the last day of the month.
- If end day/month are missing and year is non-missing, then day/month is the minimum of treatment end date (if year of treatment end date is same as the medication year) and the end of the year (31DECYYYY).
- If imputed end date is less than the start date, use the start date as the imputed end date.
- If the start date year is missing, the start date is set to one day prior to treatment start date.
- If the start date year is less than the treatment start date year, then:
  - If the month is missing, the start date is assumed to be mid-year point (01JULYYYY).
  - Else if the month is not missing, the start date is assumed to be mid-month point (15MONYYYY).
- If the start date year value is greater than the treatment start date year, then:
  - If the month is missing, the start date is assumed to be the year start point (01JANYYYY).
  - Else if the month is not missing, the start date is assumed to be the month start point (01MONYYYY).
- If the start date year value is equal to the treatment start date year value:
  - If the month is missing or the month is equal to the treatment start date month, then the start date is assumed to be one day prior to the treatment start date.
  - Else if the month is less than the treatment start date month, the start date is assumed to the mid-month point (15MONYYYY).
  - Else if the month is greater than the treatment start date month, the start date is assumed to be the month start point (01MONYYYY).

If complete end date is available and the start date assumed from the steps above is greater than the end date, then the assumed start date should be set to the end date.

Partial anti-cancer therapy start dates will be handled as follows:

- If the start of subsequent anti-cancer therapy date is missing, this will be imputed to the date of the EOT visit.
- Missing day, month/year present: If the month/year is the same as the month/year of EOT Visit date, then impute missing start dates with EOT Visit date. Otherwise, for dates corresponding to a start date, impute with the first day of the month.
- Missing month/day, year present: If the year is the same as the year of EOT Visit date, then impute missing start dates with EOT Visit date. Otherwise, for dates corresponding to a start date, impute with the first day of the year.
- If the EOT Visit date is missing, the date of last treatment will be used for imputation.

Partial date of progression for prior cancer therapies will be handled as follows, for the derivation of the time since the first diagnosis subgroup:

- Assign the day to the 15th of the month.
- If the month/year is missing, the date of progression will not be imputed.

Partial diagnosis dates will be handled as follows, for the derivation of the time since the first diagnosis subgroup:

- Assign the day to the 15th of the month.
- If the month is missing, the initial diagnosis date is assumed to be mid-year point (01JULYYYY).
- Else if the month is not missing, the initial diagnosis date is assumed to be mid-month point (15MONYYYY).

For analysis of pharmacodynamic data:

- The actual date collected will be used. If the date is not available, the value will be considered as missing.
- The actual time of sample collection will be used. If the time is not available, the value will, at the discretion of study pharmacodynamics analysist (in consultation with BioNTech), be considered as missing or imputed to be the nominal sampling time.

8.2 Patient disposition

For the screened set, the number and percentage of patients having failed screening will be presented along with a summary of the primary reason for screening failure. A summary of reason for violation of inclusion/exclusion criteria will also be presented.

The number and percentage of patients in each analysis set will be summarized by treatment group and overall for randomized patients. For each analysis set (Section 7.1) the number and percentage of patients being excluded from the analysis set will be presented by treatment group and overall along with a summary of the reasons for exclusion based on the ITT set.

For the ITT set the number and percentage of patients who were randomized and treated will be presented. Also, the number of patients on treatment and who are off-treatment will be summarized. Patients who are "off-treatment" will be the patients who received at least one dose of trial treatment and then discontinued from trial treatment. Also, for the ITT set the number and percentage of patients having discontinued from trial treatment will be presented by treatment group and total along with a summary of the primary reason for premature treatment discontinuation (e.g., adverse events, investigator's decision, disease progression, death, withdrawal of consent, lost to follow-up). The percentages for on/off-treatment and treatment discontinuation will be based on treated patients. The summary will also be repeated for Add-on treatment set and mITT set.

For the ITT set, the number and percentage of randomized / treated patients will be presented by country and site including the number and percentage of IPDs (by type, subtype (as appropriate), treatment group and total).

The number of patients who entered the efficacy follow-up (on trial/off-trial – for IDMC TFLs) and the number of patients who discontinued from the trial together with the reasons for discontinuation will be summarized for the patients in ITT by treatment and total. For trial discontinuation, efficacy follow-up, survival follow-up the percentages will be based on randomized patients.

The patients who entered efficacy follow-up will be selected based on the below definition.

''Patients where EOT had occurred and reason for EOT is NOT progression and new cancer treatment has not started and EOS has not yet occurred.''

The summary of efficacy follow-up will be provided for Add-on Treatment set and mITT set as well.

In addition, survival follow-up status including entered (on/off survival follow-up) reason for discontinuation from survival follow-up and total number of deaths will be summarized by treatment group and total for the ITT set, Add-on Treatment set and mITT set.

The number of patients randomized by treatment group and randomization strata (from IWRS) will be presented based on ITT set.

8.3 Baseline characteristics

8.3.1 Demographics

Demographic and baseline variables will be summarized for patients in the ITT set. Age (years), weight (kg), height (cm), will be summarized as continuous data. Age (<70 years vs ≥70 years), Gender (male vs female) and race (Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, White and Not Reported)

and Ethnicity (Hispanic or Latino, Not Hispanic or Latino and Not Reported) will be summarized as categorical data.

Demographic and baseline variables will be summarized for the patients in Add-on treatment set and mITT set as well.

A patient data listing will be provided.

8.3.2 Disease characteristics

The following disease characteristics will be summarized for patients in the ITT set.

- Histologic subtype (cancer subcategory),
- BRAF V600 Mutation status,
- Tumor stage at diagnosis (American Joint Committee on Cancer Staging (AJCC)),
- Tumor stage at trial entry (baseline)
- Location at the initial diagnosis

The summary will be repeated for Add-on Treatment set and mITT as well.

A patient data listing will be provided.

8.3.3 Prior cancer therapies

Prior anti-cancer treatments (as collected on the CRF) will be summarized for the ITT set, Add-on Treatment set and mITT set.

Any prior surgery (curative, palliative and diagnostic), any prior radiotherapy (Adjuvant, Neoadjuvant, Metastatic, Maintenance and unknown), intent, number of therapy lines, prior cancer therapy types, number of prior anti-PD-1/anti-PD-L1 therapies (1 vs >1), duration of anti-PD-1/anti-PD-L1 therapy prior to disease progression (≤6 months, 6-12 months, ≥12 months) and time from anti-PD-1/anti-PD-L1 discontinuation to progression (on therapy, >1 month and ≤6 months, >6 months) will all be summarized as categorical data.

A listing of prior anti-cancer therapies will be provided.

8.3.4 Prior and concomitant medication/procedures/non-drug therapy

Prior and concomitant medications will be summarized for the ITT set and the Add-on Treatment set.

Prior and concomitant medications/non-drug therapies will be defined using start and stop dates recorded, relative to the first and last dose of trial medication (BNT111 or cemiplimab).

A prior medication/prior non-drug therapy will be defined as any medication/ non-drug therapy taken 28 days prior up to (but not including) the start date of trial treatment.

A concomitant medication/concomitant non-drug therapy will be defined as any medication/non-drug therapy either ongoing at the start of trial medication or with a start date on or after the first dose of trial medication up to the SFU2 visit. Section 8.1 contains details of how partial and missing dates will be handled.

Medications will be coded using the World Health Organization Drug Dictionary (WHODrug) version B3 March 2020. Resulting in Anatomical-Therapeutic-Chemical (ATC) codes indicating therapeutic classification.

The number and percentage of patients taking prior and concomitant medications will be summarized by ATC therapeutic class (ATC level 2), ATC pharmacological class (ATC level 3), and chemical substance (ATC level 5) for each treatment group and total. The summary will be presented alphabetically.

Procedures and Non-drug therapies will be coded using the Medical Dictionary for Regulatory Activities (MedDRA®) version 23.1 or later version if used.

Procedure and non-drug therapies will be listed only.

Patient data listings will be provided.

8.3.5 Medical history

Medical history will be summarized for the ITT set and the Add-on Treatment set.

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA®) coding system version 23.1 or later version. Medical history will be summarized by System SOC and PT of the MedDRA® dictionary by treatment group and overall.

A patient data listing will be provided.

8.4 Efficacy analyses

8.4.1 Primary analysis

The primary analyses will be performed using the CCI set.

Definition of primary endpoint

Overall Response Rate (ORR) defined as the proportion of patients in whom a CR or PR is observed as best overall response by blinded independent central review (BICR).

The BICR will provide a response for each patient assessment using the RECIST© 1.1 guidelines.

Main analytical approach

The primary statistical objective is to demonstrate that the ORR of BNT111+cemiplimab is greater than CCI using an overall significance level of CCI

The following null and alternative hypothesis are defined:

CCI

The best overall response for a patient is defined as the best response across all responses prior to the start of new anti-cancer therapy for that patient. It is noted that a tumor response of SD can be confirmed from 36 days after first dose of study drug as the tumor assessment at week 6 (day 42) can be performed in a +/- 1 week window.

Statistical analysis

Best overall response will be summarized with absolute and relative frequencies (n, %). The ORR will be summarized with absolute and relative frequencies (n, %) along with a two-sided 95% Clopper-Pearson confidence interval. CCI [redacted] Patients not meeting the criteria for CR or PR, including those without any post-baseline tumor assessments, will be considered as non-responders.

The minimal statistically significant ORR at final analysis would be CCI [redacted] The operational characteristics of IA1, IA2 and the final analysis of ORR are shown in Table 8.1.

CCI [redacted]

Additionally, a swimmer plot for duration on treatment and overall response assessments based on RECIST for BICR assessment will be presented based on CCI [redacted] set, add-on treatment set and combination of both (show both periods, including only for patients entering add-on treatment set). For IA1 the plot was generated for patients with BOR of CR/PR/SD and for IA2 the plot will be generated for patients with BOR of CR/PR/SD/PD/NE (for BICR data, the category Non-CR/Non-PD will be added). Tumor response and overall response will be listed.

Sensitivity analysis

Analysis described in this section (Section 8.4.1) will be repeated for the CCI [redacted] and CCI [redacted] as sensitivity analyses. CCI [redacted]

Additionally, sensitivity analysis based on confirmed CR/PR (RECIST 1.1) for BICR assessment for the CCI [redacted] set will be performed.

8.4.2 Secondary analyses

The secondary analyses will be performed for the BNT111 + cemiplimab, BNT111 (mono) and cemiplimab (mono) treatment groups, apart from overall survival which is secondary analysis for the BNT111 + cemiplimab treatment group only. All analyses will be done using the CCI set, unless specified otherwise.

8.4.2.1 Key Secondary Analysis

As described in Section 4 the primary endpoint (ORR) will be analyzed descriptively for the calibrator groups BNT111 (mono) and cemiplimab (mono), in a non-confirmatory sense in order to obtain an estimate for the single agent anti-cancer activity. Summaries as for the primary analysis will be presented.

8.4.2.2 Disease Control Rate (DCR) using BICR and RECIST 1.1

DCR is defined as the proportion of patients in whom a CR or PR or SD (for BICR data, the Non-CR/Non-PD response will be included in the DCR, similarly to SD response) (SD assessed at least 6 weeks [+/- 1 wk] after first dose) is observed as best overall response. Patients not meeting the criteria for CR or PR or SD (for BICR data, the Non-CR/Non-PD response will be included), including those without any post-randomization tumor assessments, will be considered as non-responders. Only patients with adequate (i.e. available) baseline and post-baseline assessments, in whom a CR, PR or SD is observed will be analyzed.

DCR will be summarized as a binary endpoint along with ORR (Section 8.1).

If the sensitivity analysis of ORR for the CCI and CCI is done then a summary of DCR will also be included.

8.4.2.3 Duration of Response (DOR) using BICR and RECIST 1.1

DOR is defined as the time in months from first objective response (CR or PR) to first occurrence of objective tumor progression (PD according to BICR and RECIST 1.1), or death from any cause, whichever occurs first. Only patients with adequate baseline and post-baseline assessments, in whom a CR or PR is observed will be analyzed for DOR. DOR will be analyzed as a time-to-event endpoint (Section 8.1). DOR will be calculated as: Date of first occurrence of PD - Date of first objective response + 1. See censoring rules in Table 5.1 for more details on the DOR derivation.

8.4.2.4 Time to Response (TTR) using BICR and RECIST 1.1

TTR is defined as the time in months from randomization to first objective response (CR or PR). TTR will only be summarized for patients in whom a CR or PR is assessed. TTR will be calculated as: Date of first objective response - Date of randomization + 1.

TTR will be summarized similar to other continuous variables mentioned in Section 8.1, summary statistics will be provided.

8.4.2.5 Progression Free Survival (PFS) using BICR and RECIST 1.1

PFS is defined as the time in months from randomization to first objective tumor progression (PD according to BICR and RECIST 1.1), or death from any cause, whichever occurs first. PFS will be calculated as: Date of first objective PD or death - Date of randomization + 1.

PFS will be analyzed as a time-to-event endpoint. Patients alive and without disease progression or patients lost to follow-up at data cut-off date will be censored per the censoring rules in Table 5.1.

PFS will be analyzed as a time-to-event endpoint (Section 8.1).

8.4.2.6 ORR, DCR, DOR, TTR, PFS using Investigator assessment and RECIST 1.1

Analysis described in Sections 8.4.1, 8.4.2.1, 8.4.2.2, 8.4.2.3, 8.4.2.4 and 8.4.2.5 will be repeated using Investigator assessment based on RECIST 1.1.

Additionally, a swimmer plot based on RECIST 1.1 for investigator assessment will be presented as described in Section 8.4.1.

8.4.2.7 Overall Survival (OS) for combination arm

OS is defined as the time in months from randomization to death from any cause. OS will be calculated as: Date of death - Date of randomization + 1.

OS will be analyzed using Kaplan-Meier methodology for BNT111 + cemiplimab treatment group. Patients alive or patients lost to follow-up at data cut-off will be censored per the censoring rules in Table 5.2.

8.4.2.8 Health-Related Quality of Life (HRQoL)

Health-Related Quality of Life (HRQoL) will be measured by the European organization for research and treatment of cancer quality-of-Life questionnaire core 30 items (EORTC QLQ-C30) using Patient Global Impression of Severity questionnaire (PGIS) and Patient Global Impression of Change questionnaire (PGIC) for calibration.

Observed and change from baseline values for the global health status score and the functional and symptoms scales for the EORTC QLQ-C30 will be summarized by visit using summary statistics. Score derivation is described in Section 10.6.

Patient Global Impression of Severity questionnaire (PGIS) and Patient Global Impression of Change questionnaire (PGIC) for calibration will be summarized by visit for each treatment group.

8.4.3 Exploratory analyses

Unless otherwise specified the exploratory analyses will be performed for the BNT111 + cemiplimab, BNT111 (mono) and cemiplimab (mono) treatment groups using the CCI set.

8.4.3.1 ORR, DOR, DCR, TTR and PFS using BICR/Investigator assessment and iRECIST

The analysis of endpoints for ORR, DCR, DOR, TTR and PFS described in Sections 8.4.1, 8.4.2.1, 8.4.2.2, 8.4.2.3, 8.4.2.4, 8.4.2.5 and 8.4.2.6 will be repeated for both the BICR and Investigator assessments using response based on the iRECIST guidelines. It is noted that, per the iRECIST guidelines, first occurrence of objective tumor progression is the date that progression criteria are met (i.e. the date of iUPD) provided that iCPD is confirmed at the next assessment and iSD (or Non-iCR/Non-iUPD for BICR data), iPR or iCR had not occurred before iCPD. Further details can be seen in the iRECIST guidelines.

Additionally, a swimmer plot based on iRECIST for BICR and investigator assessment will be presented as described in Sections 8.4.1.

8.4.3.2 Duration of Disease Control (DDC) using BICR/Investigator assessment and RECIST 1.1

DDC is defined as the time in months from first objective response (CR or PR or SD (for BICR data, the Non-CR/Non-PD response will be included in the DDC, similarly to SD response) (SD assessed at least 6 weeks [+/- 1 wk] after first dose)) to first occurrence of objective tumor progression (PD according to BICR and RECIST 1.1), or death from any cause, whichever occurs first. Only patients in whom a CR or PR or SD (or Non-CR/Non-PD for BICR data) (SD assessed at least 6 weeks [+/- 1 wk] after first dose) is observed will be analyzed for DDC. DDC will be analyzed as a time-to-event endpoint (Section 8.1) based on CCI set using patients who achieved disease control (i.e. patients with a CR or PR or SD observed as best overall response). DDC will be calculated as: Date of first occurrence of PD - Date of first objective response + 1. DOR will be calculated as: Date of first occurrence of PD - Date of first objective response + 1. See censoring rules in Table 5.1 for more details on the DDC derivation.

8.4.3.3 Duration of Disease Control (DDC) using BICR/Investigator assessment and iRECIST 1.1

The analysis of endpoint for DDC described in Section 8.4.3.2 will be repeated for both the BICR and Investigator assessments using response based on the iRECIST guidelines.

8.4.3.4 Depth of response (DEoR)

DEoR is defined as the maximum percent reduction from baseline in tumor size reported in the CRF by the investigator and as reported by the BICR. Tumor size is defined as sum of target lesion diameter at that assessment reported in the CRF. The sum of target lesion diameter includes nodal (short axis) and non-nodal (longest axis) lesions.

Percent change in tumor size from baseline will be calculated at each scheduled tumor assessment. The maximum percent reduction (DEoR) per patient will be selected from the data specified as follows.

- For CCI set, data up to and including first progression (PD)/first occurrence of objective tumor progression (i.e., the date of iUPD), provided that iCPD is confirmed at the next assessment until the day before first dose of add-on treatment, if there is

add-on treatment, will be used. If there is no progression/objective tumor progression (progression of iRECIST), data until the day before first dose of add-on treatment, if there is add-on treatment, will be used.

- For Add-on treatment set, data from the date of first dose of add-on treatment up to and including first progression (PD)/first occurrence of objective tumor progression (i.e., the date of iUPD), provided that iCPD is confirmed at the next assessment, will be used for deriving DEoR. If there is no progression/objective tumor progression, all data from the date of first dose of add-on treatment will be used.
- Only data collected before the start of new anti-cancer therapy will be used.
- If there is an assessment (not the one after the first occurrence of PD based on RECIST 1.1/objective tumor progression based on iRECIST) on the same date as the start of new anti-cancer therapy, this assessment data will be used.

Tumor size will be summarized by scheduled visit (the details of visit mapping are in Section 10.7) and DEoR will be summarized using summary statistics by treatment group based on RECIST 1.1 and iRECIST for the randomized treatment and Add-on treatment.

In addition, DEoR for RECIST 1.1 and iRECIST will be illustrated by treatment using waterfall plots, separately for the CCI set and Add-on treatment set, and listed in a patient data listing.

Additionally, spider plots for percent change of target lesions in sum of diameter from baseline will be presented by treatment based on RECIST 1.1 and iRECIST for BICR and investigator assessments based on CCI and add-on treatment set. Only percentage change used for deriving DEoR will be displayed in the spider plots.

8.4.3.5 Analysis for Add-on Treatment Set

The analysis of endpoints ORR, DCR, DOR, TTR, PFS described in Sections 8.4.2.1, 8.4.2.2, 8.4.2.3, 8.4.2.4, 8.4.2.5 and 8.4.2.6 will be repeated for the Add-on Treatment set using the response for BICR and Investigator assessment based on the RECIST 1.1 and iRECIST guidelines. Per Section 8.1, these analyses will be based on data collected from the date of first dose of add-on treatment and after. For TTR and PFS, the starting date is the date of first dose of add-on treatment.

The analysis of DEoR for the add-on treatment set is in Section 8.4.3.4.

8.4.3.6 Overall Survival for arm 2 and arm 3

Analysis of OS will be performed for the BNT111 mono and cemiplimab mono treatment groups using the same method as Section 8.4.2.7 for the CCI set.

The analysis will be repeated for each of the following:

- Add-on treatment Set using the BNT111 + add-on cemiplimab and cemiplimab + add-on BNT111 treatment groups (the definition of overall survival will be the same as that in Section 8.4.2.7 e.g. starts from the date of randomization)
- The CCI set excluding patients that are in the Add-on Treatment set

An additional analysis would also be performed for the BNT111 + add-on cemiplimab and cemiplimab + add-on BNT111 treatment groups using the Add-on Treatment set. OS, for this analysis, would be defined as the time in months from date of first dose of add-on treatment to death from any cause.

Censoring rules are in appendix Table 5.2.

8.4.3.7 Antigen-specific T cells

The presentation and analysis of antigen-specific T cell will be described in a separate Biomarker SAP that will be created by BioNTech.

8.4.3.8 Pharmacodynamic analysis

8.4.3.8.1 Biomarker Analysis

Biomarker investigations in this trial will include pharmacodynamic biomarkers to confirm biological activity of BNT111 as monotherapy and in combination with cemiplimab and predictive biomarkers that may identify patients that can benefit from therapy.

An overview of the biomarker samples that will be collected in this trial is provided in the SoA.

Biomarker assessment summaries for the CSR will focus on:

8.4.3.8.1.1 Occurrence of BNT111 target Tumor-associated Antigen (TAA) expression

Occurrence of BNT111 target TAA expression in tumor tissue at baseline [Positive/Negative] (4 targets: New York esophageal squamous cell carcinoma-1 (NY-ESO-1), tyrosinase, melanoma antigen A3 (MAGE-A3), transmembrane phosphatase with tensin homology (TPTE)).
This TAA expression and the cumulative TAA expression will be summarized using descriptive summary statistics by treatment group for the safety set based on patients with available data. This summary statistics will also be produced at IA2.

8.4.3.8.1.2 Continuous PD-L1 Tumor Proportion Score (TPS)

Continuous PD-L1 TPS at baseline in percentage (central PD-L1 test results). These biomarker data will be summarized using descriptive summary statistics by treatment group for the CCI set. This summary statistics will also be produced at IA2.

8.4.3.8.1.3 Categorical PD-L1 Tumor Proportion Score (TPS)

Categorical (Positive/TPS≥1% vs Negative/TPS<1%) PD-L1 expression at baseline (central PD-L1 test results). The biomarker values will be summarized using descriptive summary statistics by treatment group for the CCI set. This summary statistics will also be produced at IA2.

Additionally, summary statistics by response ((CR, PR) vs SD vs PD; based on BICR assessed RECIST 1.1) will be produced within treatment group for the CCI set. At IA2, the above summary statistics will also be produced.

BOR, ORR and DCR will be summarized with absolute and relative frequencies (n, %) by categorical PD-L1 expression at baseline within the treatment group. A box plot of PD-L1 expression by Response Status (BICR assessed RECIST 1.1) will be generated by treatment group for CCI set. At IA2, the plot will be produced. PFS and OS will be analyzed by Kaplan-Meier estimates and the corresponding Kaplan-Meier plots will be produced, according to BICR assessed RECIST 1.1.

8.4.3.8.1.4 Peripheral Cytokine Profiles

Peripheral cytokine profiles under treatment over time [absolute values]. The biomarker values at each time-point will be summarized using descriptive summary statistics and shown in boxplots by treatment group for the CCI set.

8.4.3.8.1.5 BRAF Mutations at Baseline

BOR, ORR and DCR will be summarized with absolute and relative frequencies (n, %) by BRAF mutation at baseline (Section 7.3). For BOR, ORR, DCR and PFS, the analyses will be performed according to BICR assessed RECIST 1.1 for the CCI set. At IA2, the above analyses will be performed.

8.5 Safety analyses

Safety data that will be summarized includes trial treatment exposure, adverse events (AEs), clinical laboratory assessments, vital signs and ECGs. All safety analyses will be based on the Safety set and Add-on Treatment set and will be summarized by treatment group unless otherwise stated.

All safety data will be listed.

See Section 8.1 for data selection using the Safety set and Add-on Treatment set.

8.5.1 Extent of exposure

The following dose exposure variables will be derived and analyzed for each trial treatment: The analyses will be based on patients in the Safety set and will be repeated for the Add-on Treatment set.

- Number of cycles received. For BNT111, a cycle is considered if at least 1 injection within the cycle is taken.
- Number of BNT111 injections and cemiplimab infusions
- (Actual) Trial Treatment Duration (weeks) for the summary using the Safety set, is defined as follows: (Date of last administration - Date of first administration + Planned Duration) / 7, where the Planned Duration (days) is defined as the planned time between two consecutive administrations. Planned duration for BNT111 is one week for the 2 first cycles, followed by once every 3 weeks; planned duration for cemiplimab is 3 weeks. Actual treatment duration is calculated for each treatment separately.
- (Actual) add-on Treatment Duration (weeks) for the summary using the add-on treatment Set is defined as follows: (Date of last administration of add-on - Date of first

administration of add-on + Planned Duration) / 7, where the Planned Duration (days) is defined as the planned time between two consecutive administrations. Planned duration for BNT111 is one week for the 2 first cycles, followed by once every 3 weeks; planned duration for cemiplimab is 3 weeks. Actual add-on treatment duration is calculated for each add-on treatment separately.

- Planned treatment duration (weeks) defined as (scheduled study day of last administration - scheduled study day of first administration + Planned Duration) / 7. Planned duration is as above.
  - The scheduled study day of last administration is the study day at which the last administration should be given according to the protocol schedule. For example, if the schedule is once every 3 weeks and the patients received 2 administrations (at C1D1 and C3D1), the scheduled study day of last administration would be day 43.
- (Actual) Treatment duration (weeks), of the randomized treatment, for the summary using the Add-on Treatment set, the duration of randomized treatment from first dose of Add-on treatment is defined as follows: (Date of last administration - Date of first administration received after enrollment into Add-on treatment + Planned Duration) / 7, where the Planned Duration (days) is defined as the planned time between two consecutive administrations. Planned duration for BNT111 is one week for the 2 first cycles, followed by once every 3 weeks; planned duration for cemiplimab is 3 weeks. Actual treatment duration is calculated for each treatment separately.
- (Actual) Cumulative Dose defined as sum of all administered doses, separately, for each trial treatment. BNT111 (µg) and cemiplimab (mg).
- Planned cumulative dose defined as sum of all planned doses, separately, for each trial treatment. BNT111 (µg) and cemiplimab (mg). This is calculated as planned cumulative dose across all the cycles.
- Dose Intensity (DI) is defined as Cumulative Dose for that trial treatment (unit for that trial treatment) / (Actual) Treatment Duration (weeks) for that trial treatment.
- Relative Dose Intensity (RDI) defined as follows:

$$\text{RDI (\%)} = \frac{\text{Actual Dose Intensity } \left(\frac{\mu g}{\text{week}}\right)}{\text{Planned Dose Intensity } \left(\frac{\mu g}{\text{week}}\right)} \text{ x } 100 = \text{DI x TI x } 100,$$

where,

$$\text{(Actual) Dose Intensity} \left(\frac{\mu g}{\text{week}}\right) = \frac{\text{(Actual) Cumulative Dose } (\mu g)}{\text{(Actual) Treatment Duration (weeks)}}$$

$$\text{Planned Dose Intensity } \left(\frac{\mu g}{\text{week}}\right) = \frac{\text{Planned Cumulative Dose } (\mu g)}{\text{Planned Treatment Duration (weeks)}}$$

$$\text{Dose Index (DI)} = \frac{\text{Total Administered Dose } (\mu g)}{\text{Total Planned Dose } (\mu g)}$$

$$\text{Time Index (TI)} = \frac{\text{Planned Treatment Duration (weeks)}}{\text{Actual Treatment Duration (weeks)}}$$

The units in the derivations above are for BNT111. For cemiplimab mg would be used.

The following variables will be presented with summary statistics by treatment group (as relevant for that treatment group) for the Safety set and the Add-on Treatment set:

- Number of cycles received (categorical, cycles 1,2,3,4,5,6,7,8,9,10-12,13-15,16-18,19-24,25-30, >30). For BNT111, a cycle is considered if at least 1 injection within the cycle is taken.
- Number of BNT111 injections and cemiplimab infusions (descriptive statistics)
- Number of BNT111 injections (categorical, ≤6 injection, 7-12 injections, 13-18 injections, >18 injections)
- Number of cemiplimab infusions (categorical, 1, 2, up to 9 then 10-12, 13-15 and so on)
- Duration of treatment
- Cumulative dose
- Dose intensity
- Relative dose intensity

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, dose modification, at least 1 dose modification and at least 2 dose modification any dose interruption and any dose prolonged will be presented by treatment group. This summary will also be repeated for mITT set.

8.5.2 Adverse events

Adverse events (AEs) will be coded using Medical Dictionary for Regulatory Activities (MedDRA®) coding system version 23.1 or later to get a System Organ Class (SOC) and Preferred Term (PT) for each AE and graded for severity using NCI CTCAE v5.0. 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-CTC grades will be performed.

TEAEs for randomized treatment

A treatment-emergent AE (TEAE) is defined as any AE with an onset date on or after the first administration of trial treatment (if the AE was absent before the first administration of

trial treatment) or worsened after the first administration of trial treatment (if the AE was present before the first administration of trial treatment). AEs with an onset date more than 90 days after the last administration of trial treatment will be considered as treatment emergent only if assessed as related to trial treatment by the investigator, prior to start of Add-on treatment, if applicable, whichever comes first.

An AE will be considered TE for the randomized treatment if it started within 90 days of last dose of randomized treatment or until start of add-on treatment whichever comes first.

Add-on TEAEs

Treatment-emergent for add-on treatment is defined as any AE with an onset date on or after the first administration of add-on treatment (if the AE was absent before the first administration of add-on treatment) or worsened after the first administration of the add-on treatment (if the AE was present before the first administration of add-on treatment). AEs with an onset date more than 90 days after the last administration of add-on treatment will be considered as TE only if assessed as related to add-on treatment by the investigator.

TEAEs for both randomized and add-on treatment

Treatment emergent for both randomized & add-on treatment is defined as any AE with an onset date on or after the first administration of randomized treatment / add-on treatment (if the AE was absent before the first administration of randomized treatment / add-on treatment) or worsened after the first administration of randomized treatment / add-on treatment (if the AE was present before the first administration of randomized treatment / add-on treatment). AEs with an onset date more than 90 days after the last administration of randomized/add-on treatment will be considered as TE only if assessed as related to randomized/add-on treatment by the investigator. If an AE is treatment emergent for the randomized treatment and worsens in intensity during the Add-on treatment, the AE will appear only once in the summary.

TEAEs for randomized, add-on TEAEs and TEAEs for randomized and add-on treatment will be summarized by treatment group and overall for patients in the Safety analysis set and the Add-on Treatment set.

Adverse Events of Special Interest (AESI)

- Grade ≥2 infusion-related reactions related to cemiplimab are considered AESIs.
- Grade ≥3 irAEs (immune-related AEs) related to cemiplimab are considered AESIs.

Overall summary of adverse events (AE)

The number and percentage of patients reporting at least one AE will be summarized for each of the following AE types (summary applies to both BNT111 and cemiplimab unless otherwise specified):

- Any AE
- TEAE
- Related TEAE (BNT111, cemiplimab and both)

- Grade ≥3 TEAE
- Related grade ≥3 TEAE (BNT111, cemiplimab and both)
- Any TESAE
- Related TESAE (BNT111, cemiplimab and both)
- TESAE leading to death
- Related TESAE leading to death (BNT111, cemiplimab and both)
- Any AESI
- Grade≥ 2 infusion-related reactions (IRRs) related to cemiplimab
- Grade≥ 3 irAEs (immune-related AEs) related to cemiplimab
- TEAEs leading to dose reduction (BNT111)
- TEAEs leading to dose interruption (BNT111, cemiplimab and both)
- TEAEs leading to permanent discontinuation of treatment
- Related TEAEs leading to dose reduction (BNT111)
- Related TEAEs leading to dose interruption (BNT111, cemiplimab and both)
- Related TEAEs leading to permanent discontinuation of treatment

Analyses of adverse events

The number and percentage of patients reporting TEAEs will be summarized by PT nested within SOC. If a SOC / PT is reported more than once for a patient, the patient will only be counted once for this SOC / PT. Strongest relationship or worst grade will be counted if a TEAE is reported more than once by the same patient for a SOC / PT.

All AE summary tables will be presented for BNT111 and cemiplimab separately and sorted alphabetically by SOC and PT within SOC.

Analyses of other adverse events (excluding serious AEs)

The number and percentage of patients reporting other TEAEs (excluding TESAEs) will be summarized by PT nested within SOC. 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 presented for BNT111 and cemiplimab separately and sorted alphabetically by SOC and PT within SOC.

Treatment-related AEs

The number and percentage of patients reporting treatment-related TEAEs will be summarized by PT nested within SOC for BNT111 and cemiplimab separately and related to both the treatments.

In addition, treatment-related for both randomized and add-on treatment TEAEs will be summarized by PT nested within SOC for BNT111 and cemiplimab separately.

AESIs (irAE and Infusion-related AEs)

The number and percentage of patients with AESIs, irAEs grade ≥3 related to cemiplimab and infusion-related reactions grade ≥2 related to cemiplimab will be summarized by PT nested within SOC.

TEAE by NCI-CTCAE grade

The number and percentage of patients with TEAEs will be summarized by worst NCI-CTCAE grade by PT nested within SOC. Worst grade will be counted if a TEAE is reported more than once by the same patient for a SOC / PT.

TEAE by NCI-CTCAE grade by relationship status

The number and percentage of patients with related and not related TEAEs will be summarized by worst NCI-CTCAE grade by PT nested within SOC for BNT111 and cemiplimab separately. Worst grade will be counted if a TEAE is reported more than once by the same patient for a SOC / PT.

Serious Adverse Events

The number and percentage of patients with TESAEs will be summarized by PT nested within SOC.

Related Serious Adverse Events

The number and percentage of patients with related TESAEs (for BNT111 and cemiplimab separately and related to both the treatments) will be summarized by PT nested within SOC.

In addition, treatment-related TESAEs for both randomized and add-on treatment will be summarized by PT nested within SOC for BNT111 and cemiplimab separately.

Related Serious Adverse Events by NCI-CTCAE grade

The number and percentage of patients with related TESAEs (for BNT111 and cemiplimab separately and related to both the treatments) will be summarized by worst NCI-CTCAE grade by PT nested within SOC. Worst grade will be counted if a TEAE is reported more than once by the same patient for a SOC / PT

Deaths

- All deaths including those that occurred on-treatment (deaths occurred between first and last trial treatment intake during the study + 30 days) and post-treatment (deaths occurred after the last trial treatment intake during the study + 30 days), will be listed. Post treatment deaths will be flagged. In addition to the summary of on-treatment deaths and post-treatment deaths, an additional summary of all deaths (on-treatment + post-treatment) will be provided along with primary reason for death.

- A separate summary of TESAEs leading to death will be provided by PT nested within SOC.
- A separate summary of related TESAEs leading to death (for BNT111 and cemiplimab separately and related to both the treatments) will be provided by PT nested within SOC.

AE listings

All TEAEs, fatal AEs, SAEs, TEAEs leading to dose reduction and interruption, TEAEs leading to permanent discontinuation of treatment, Infusion-related reactions, irAEs, and TEAEs related to trial procedures will be listed.

8.5.3 Laboratory assessments

Clinical laboratory data will be summarized for Safety analysis set and the Add-on Treatment set (as applicable).

Clinical laboratory data to be summarized includes hematology (for IAs, only Hemoglobin, White Blood Cells (WBC), Absolute Lymphocytes Count, Absolute Neutrophils Count, and Platelets will only be presented). For CSR, all parameters planned as per protocol will be presented), clinical chemistry (for IAs, only Total Bilirubin, Aspartate aminotransferase (AST), Alanine aminotransferase (ALT), Lipase, Lipase pancreatic, Amylase, Creatine kinase and Creatinine will be presented. For CSR, all parameters planned as per protocol will be presented), and urinalysis. The safety laboratory parameters to be assessed are listed in Section 10.2 of the protocol amendment. The scheduled time points for assessment are outlined in the SoAs (Section 10.4).

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

If a laboratory value is reported using a nonnumeric qualifier e.g., less than (<) a certain value, or greater than (>) a certain value, the given numeric value will be used in the summary statistics, ignoring the nonnumeric qualifier.

In case there are several results for the same parameter at the same visit, the record from the central lab will be used for the analyses. The local lab results will be only listed.

Safety laboratory results will be classified according to the most recent version of NCI-CTCAE. Shift tables from baseline to worst grade on treatment will be provided for each laboratory parameter by randomized treatment group.

Laboratory values that are below or above the reference ranges will be flagged.

Abnormal and clinically significant, abnormal and not clinically significant and normal laboratory values will be summarized for each urinalysis parameter by visit and randomized treatment group.

In addition, clinically laboratory values for each parameter will be plotted using trellis plots. Maximum post baseline values will be displayed vs baseline value by treatment group. The

plots will be repeated for minimum post baseline values vs baseline value by treatment group.

Notes for programming: For lab summary by visit tables: for baseline visit under value statistic, include all subjects having non-missing baseline data; for post-baseline visit under value statistic, include all subjects having non-missing post-baseline data; for post-baseline visit under CFB statistic, include subjects having both baseline and post-baseline data.

The data as collected will be presented in the listings along with the reference ranges and CTCAE grade. Laboratory values that are below or above the reference ranges will be flagged. Laboratory listings will be presented similarly to the laboratory tables, for IAs the same parameters as present in the tables will be presented. For the CSR all parameters planned as per protocol will be presented.

8.5.4 Vital signs

Vital sign parameters to be summarized by treatment group include systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate and body temperature.

Vital sign parameters at each visit will be summarized using descriptive summary statistics for each parameter as follows:

- change from baseline to each post-baseline visit, considering the value at 60-min prior treatment at Cycle 1, Day 1 as the baseline value (see Section 8.1 for definition of baseline). For post-baseline visits where treatment is administered, the 60-min prior treatment value at each post-baseline to be the visit value. Else the value collected at the visit will be used.
- change from pre-treatment value (60-min prior treatment) at each visit to each post-treatment time point value at that visit. So, considering the 60-min prior treatment value at a visit as the baseline for that visit for this calculation

Additionally, the occurrence of normal, abnormal clinically significant, abnormal not clinically significant results will be analyzed as categorical data for each parameter by visit and treatment group.

8.5.5 ECG

ECG parameters to be summarized include QT, QTcF, PR, and QRS intervals in msec, heart rate (bpm).

ECG parameters will be summarized by presenting summary statistics of the observed and change from baseline values by visit and treatment group.

Investigator interpretation of ECGs will be summarized as for categorical data by treatment group for each visit.

8.5.6 ECOG performance status

The ECOG performance status is a rating scale used to assess how a patient's disease is progressing as well as how the disease affects the daily living abilities of the patient. The

score values are 0 – 5, where 0 indicates “fully active, able to carry on all pre-disease performance without restriction” and 5 indicates the patient has died.

ECOG will be summarised as categorical data by treatment group and visit. Additionally, a shift table from baseline to each visit by treatment will be provided.

A listing of ECOG performance status will also be provided.

8.5.7 Physical examination

A listing of physical examination findings will be presented.

8.6 Other analyses

Subsequent anti-cancer therapies would be summarised for both the randomized and add-on treatment groups.

Type of therapy would be summarized at categorical data. Disease status at the end of the therapy would be summarised for both the for both the randomized and add-on treatment groups.

A patient data listing would be provided.

9 REFERENCES

CCI

3. Guidance for Industry: Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics (FDA Apr 2015)
4. Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics (FDA 2007)
5. Clopper, CJ and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934; 26: 404-13.
6. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1) Eur J Cancer. 2009; 45: 228–47.
7. Seymour, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet. Oncol. 2017; 18(3): e143-e152.
8. Musoro, JZ, Bottomley, A, Coens, C et al. Interpreting European Organisation for Research and Treatment for Cancer Quality of life Questionnaire core 30 scores as minimally importantly different for patients with malignant melanoma. Eur J Cancer. 2018: 104: 169-81.
9. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009; 45: 228–47.
10. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. The Lancet Oncology. 2017;18:e143 –52

10 SUPPORTING DOCUMENTATION

10.1 Appendix 1: Changes to protocol-planned analyses

- An additional analysis set have been defined, Add-on Treatment set to support planned analysis.
- Duration of disease control has been added as an additional exploratory analysis, not specified in the protocol.
- The Non-CR/Non-PD response has been added for BICR analyses of DDC and DCR.
- Per Sponsor's decision, the analyses of time to first clinically meaningful deterioration in global health status score and in symptoms and functioning as measured by EORTC QLQ-C30 are not performed and are removed from the SAP.

10.2 Appendix 2: List of abbreviations

AE	Adverse Event
AESI	Adverse Event of Special Interest
ATC	Anatomical Therapeutic Chemical
BICR	Blinded Independent Central Review
BRAF	B-Raf proto-oncogene
bpm	beats per minute
CI	Confidence Interval
CR	Complete Response
CTR	Clinical Trial Report
DCR	Disease Control Rate
DDC	Duration of Disease Control
DI	Dose Intensity
DOR	Duration of Response
DEoR	Depth of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
EOS	End of Trial
EOT	End of Treatment
FDA	Food and Drug Administration
FU	Follow-up
h	hour
HRQoL	Health-Related Quality of Life
ICH	International Conference on Harmonization
IDMC	Independent Data Monitoring Committee
IPD	Important Protocol Deviation
irAE	Occurrence of Immune-Related Adverse Events
ITT	Intent-To-Treat
kg	kilogram
MAGE-A3	Melonoma Antigen A3
max	maximum
MedDRA™	Medical Dictionary for Regulatory Activities
mITT	Modified Intent To Treat

min	minimum
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NY-ESO-1	New York Esophageal Squamous Cell Carcinoma-1
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PD-1	Programmed death protein
PD-L1	Programmed Cell Death ligand 1
PGIC	Patient Global Impression of Change Questionnaire
PGIS	Patient Global Impression of Severity Questionnaire
PFS	Progression-Free Survival
PP	Per Protocol
PPS	Per Protocol Set
PR	Partial Response
PRO	Patient Reported Outcome
PT	Preferred Term
QoL	Quality of Life
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis Software
SD	Stable Disease
SI	International System of Units
SOC	System Organ Class
SSP	Study Specific Procedure
TAA	Tumor-associated Antigen
TEAE	Treatment Emergent Adverse Event
TPS	Tumor Proportion Score
TLF	Tables, Listings, and Figures
TMB	Tumor Mutational Burden
TTR	Time to Response

10.3 Appendix 3: Reporting conventions

SAS version 9.4, or higher, will be used to produce all tables, listings, and figures.

For summary statistics, the mean and median will be displayed to one decimal place greater than the original value and the measure of variability (e.g., SD) will be displayed to two decimal places greater than the original value. Minimum and maximum will be reported to the same decimal places as the original value. Confidence interval will be displayed to the same number of places as the point estimate. P-values will be reported to three decimal places; p-values less than 0.001 will be reported as p<0.001.

The following treatment l will be used:

Treatment	Treatment Label	Treatment Code	Treatment Variables
BNT111 + cemiplimab	BNT111 + cemiplimab	1	TRT01PN
Calibrator arm BNT111	BNT111 mono	2	TRT01PN
Calibrator arm cemiplimab	cemiplimab mono	3	TRT01PN
BNT111 + add-on cemiplimab	BNT111 + add-on cemiplimab	4 for BNT111 and 5 for add-on cemiplimab	TRT02PN
Cemiplimab + add-on BNT111	Cemiplimab + add-on BNT111	6 for cemiplimb & 7 for add-on BNT111	TRT02PN

Note: The treatment code will be ued in the statistical analysis models.

10.4 Appendix 4: Schedule of activities

10.4.1 Schedule of activities for Arm 1, Arm 2, and Arm 3 patients (SOA1)

SOA1	Screening [a]		Treatment period							End of treatment [b]	Post-treatment follow-up period			
Cycle=21 d			Cycle 1			Cycle 2			Cycle 3 onward		Safety FU [c]		Efficacy FU [d]	Survival FU [e]
Visit			C1D1	C1D8	C1D15	C2D1	C2D8	C2D15	C3D1-CnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤21 d prior to randomization	≤10 d prior to randomization	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window	-21 d	-7 d		±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Assessments and procedures														
Informed consent	X [h]													
Eligibility		X [i]												
Demographic data	X													
Relevant medical history	X													
Underlying disease history	X [j]													
BRAF V600 mutation status [k]	X													
TNM staging at the time of trial entry		X [l]												

SOA1	Screening [a]		Treatment period							End of treatment [b]	Post-treatment follow-up period			
Cycle=21 d			Cycle 1			Cycle 2			Cycle 3 onward		Safety FU [c]		Efficacy FU [d]	Survival FU [e]
Visit			C1D1	C1D8	C1D15	C2D1	C2D8	C2D15	C3D1-CnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤21 d prior to randomization	≤10 d prior to randomization	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window	-21 d	-7 d		±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Concomitant medication and therapy [m]		X	X	X	X	X	X	X	X	X	X	X		
Vital signs [n]		X	X	X	X	X	X	X	X	X	X			
Weight		X	X	X	X	X	X	X	X	X	X			
Height		X												
Physical examination [o]		X	X	X	X	X	X	X	X	X	X			
ECOG PS		X	X	X	X	X	X	X	X	X	X			
12-lead ECG [p]		X	Perform, if clinically indicated							X				
CT and/or MRI [q]	X [r]		X [q]										X [q]	
Brain imaging	X		X [s]										X [s]	
Adverse events [t]	X	X	X	X	X	X	X	X	X	X	X	X		
PRO/QoL assessments [u]			X			X			X	X	X			
PGIS			X						C4D1, C8D1	X				
PGIC									C4D1, C8D1	X				

SOA1	Screening [a]		Treatment period							End of treatment [b]	Post-treatment follow-up period			
Cycle=21 d			Cycle 1			Cycle 2			Cycle 3 onward		Safety FU [c]		Efficacy FU [d]	Survival FU [e]
Visit			C1D1	C1D8	C1D15	C2D1	C2D8	C2D15	C3D1-CnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤21 d prior to randomization	≤10 d prior to randomization	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window	-21 d	-7 d		±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Survival status														X
Subsequent anti-cancer therapy										X	X	X	X	X
Treatment arms														
Randomization [v]			X											
ARM 1														
BNT111 [w]			X	X	X	X	X	X	X					
Cemiplimab (Q3W) [x]			X			X			X					
ARM 2														
BNT111 [w]			X	X	X	X	X	X	X					
ARM 3														
Cemiplimab (Q3W) [x]			X			X			X					
Laboratory assessments														
Hematology [y]		X	X	X	X	X	X	X	X	X	X			

SOA1	Screening [a]		Treatment period							End of treatment [b]	Post-treatment follow-up period			
Cycle=21 d			Cycle 1			Cycle 2			Cycle 3 onward		Safety FU [c]		Efficacy FU [d]	Survival FU [e]
Visit			C1D1	C1D8	C1D15	C2D1	C2D8	C2D15	C3D1-CnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤21 d prior to randomization	≤10 d prior to randomization	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window	-21 d	-7 d		±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Serology [z]		X												
Blood chemistry [aa]		X	X	X	X	X	X	X	X	X	X			
Coagulation (INR and aPTT) [bb]		X	Perform if clinically indicated							X	Perform if clinically indicated			
TSH and free T4 [cc]		X	X [cc]							X	X			
Pregnancy test [dd]		X	X [dd]							X	X			
Urinalysis [ee]		X	Perform if clinically indicated							X	X			
Cytokines [ff]			X			X			X [gg]					
Cellular immune responses [hh]		X [hh]	X [hh]							X				

HLA haplotyping[ll]		X[ll]												
Tumor biopsy [ii]	X								X [ii]					
Archival tumor sample [jj]	X													
Pharmacogenomic analyses[kk]	X													

Note: On treatment days, all assessments should be performed prior to dosing unless otherwise specified.

* Visits C1D8, C1D15, C2D8, and C2D15 are not to be conducted in patients in Arm 3 (cemiplimab only).

a Patients who do not meet the criteria for participation in this trial (screen failure) may be rescreened once. Patients must re-sign the informed consent form prior to re-screening, if the screening period exceeds 21 d. For both screening and re-screening, results of radiology scans performed prior to obtaining informed consent and within 28 d prior to randomization may be used. However, in case of signs of rapid progression, a new tumor evaluation may be necessary to obtain a reliable baseline assessment.

b Patients who discontinue trial treatment for whatever reason, will return to the clinic for End of Treatment (EoT) visit within 21 d after last dose. The visit at which response assessment shows (confirmed) progressive disease may be used as the EoT visit. The EoT visit and post-treatment follow-up period will only be performed after discontinuation of both trial drugs. The EoT visit and post-treatment follow-up period for patients in Arms 2 and 3 who will continue to receive add-on therapy will be performed as outlined in SOA2 (Section 10.4.2).

c All adverse events will be reported until 90 d after the last dose of trial treatment or until a new anti-cancer therapy is started. Additionally, in case a new anti-cancer therapy is started, the safety follow-up visit may be omitted.

d Patients who discontinue trial therapy prior to disease progression should continue to have regular efficacy assessments as scheduled per protocol, until disease progression, the start of next-line anti-cancer therapy, withdrawal of consent, or death, whichever occur first.

e After the last trial visit, information on survival status and new anti-cancer therapy will be collected via phone calls, patient medical records, and/or clinic visits approximately every 3 mon until death (unless the patient withdraws consent for survival follow-up or the sponsor terminates the trial). If a patient asks to be withdrawn from survival follow-up, this request must be documented in the source documents and signed by the investigator.

f "n" represents the number of the respective cycle (e.g., C4 means fourth cycle of treatment).

g Visit CnD1 should occur 21 (± 3) d after the first day (i.e., treatment visit) of the previous cycle.

h Informed consent must be documented before any trial-specific screening procedure is performed.

i Eligibility should be confirmed and documented in the eCRF. This should be performed when all screening results are available and within 21 d after initiation of screening.

j Underlying disease history should include histology, TNM stage at diagnosis (including which AJCC edition), prior anti-cancer therapies/procedures, date of last progression.

k BRAF V600 mutation status may be obtained from medical history data.

l TNM staging at the time of trial entry to be performed according to AJCC 8th edition.

m Concomitant medication and therapy includes any therapies (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated trial treatment from 28 d prior to initiation of trial treatment until the SFU2 visit.

n Vital signs (body temperature, pulse rate, respiratory rate, and blood pressure) should be measured within 60 min prior to the administration of the trial treatment(s) and at 30 (±10) min, 60 (±10) min, 4 h (±10 min), and 6 h (±10 min) after administration, or if clinically indicated (see Section 8.2.3 of the protocol for guidelines on reduction or expansion of these time points).

o At screening and at end of treatment assessments must include a full physical examination. Record abnormalities observed on the respective eCRF pages as appropriate. For all subsequent physical examinations perform a limited, symptom-directed examination at specified time points and as clinically indicated at other time points. Record new or worsened clinically significant abnormalities on the Adverse Event eCRF.

p Single 12-lead ECG recordings will be obtained during screening and as clinically indicated at other time points. Patients should be resting in a supine position for at least 10 min prior to ECG recording. The interpretation of the ECG recording will be done locally by the investigator.

q Tumor response will be measured by RECIST 1.1 / iRECIST. Patients will undergo tumor assessments at screening, every 6±1 wks for the 3 mon after trial treatment start, every 9±1 wks for the next 9 mon, and every 12±2 wks thereafter, regardless of dose delays, until (confirmed) disease progression (even if the patients discontinue trial treatment for other reasons) or until the start of next-line anti-cancer therapy (whichever occurs first). An extra scan should be scheduled if clinically indicated. Patients in Arms 2 and 3, who experience (confirmed) disease progression under single agent treatment, may be offered addition of the respective other compound to the ongoing treatment, in such cases tumor assessments are to continue (see Section 1.3.2).

r All measurable and evaluable lesions should be assessed at screening and target as well as non-target lesions identified and documented following RECIST 1.1 criteria. Tumor assessments performed as standard of care prior to obtaining informed consent and within 28 d prior to randomization do not have to be repeated at screening. Radiological progression to be confirmed by 2 scans 4-12 wks apart. If progression is accompanied by new symptoms or deterioration of performance status not attributed to toxicity one scan is sufficient.

s Any brain lesions present at baseline must be stable as per eligibility criteria and should be followed and assessed at all subsequent tumor assessment visits. If no brain lesions are present at baseline, brain imaging does not need to be repeated, only if clinically indicated as per investigator judgment. Brain imaging should be done by MRI, only in the case of a contra-indication for MRI, a contrast-enhanced CT brain scan will be acceptable.

t After informed consent has been obtained, all adverse events will be documented on the eCRF, and all serious adverse events should be reported to the sponsor within the appropriate time frame. All adverse events will be reported until 90 d after the last dose of trial treatment (i.e., SFU2 visit). In addition, the sponsor should be notified (even after SFU2 visit), if the investigator becomes aware of any serious adverse event that is believed to be related to prior exposure to trial treatment or a trial procedure.

u Patient reported outcomes/Quality of life assessment questionnaires must be completed before any assessments are done or any treatment is given on the dosing day. Patients will provide PRO/QoL assessment during trial visits at baseline, C2D1, C3D1, C4D1, every 9±1 wks for the next 9 mon, and every 12±2 wks thereafter until SFU1.

v Randomization must be done before administration of trial treatment.

w Treatment with BNT111 must be initiated no later than 7 d after confirmed eligibility. After administration of BNT111, an initial observation time of at least 6 h is required (see protocol Section 6.8.4.1). Treatment is to be continued up to 24 mon or until confirmed disease progression or patient withdrawal of consent or unacceptable toxicity.

x Treatment with Cemiplimab must be initiated no later than 7 d after confirmed eligibility. Cemiplimab will be given according to the valid Prescribing Information in the country. The first administration of cemiplimab will be on Cycle 1 Day 1, followed by every 3 wks (±3 d) thereafter. Cemiplimab will be administered after administration of BNT111 with at least a 30 min interval (if feasible, both trial drugs should be given on the same day, a maximal window of 1 wk between administrations would be allowed). There should be initial observation time of 6 h after first administration of cemiplimab. From second administration of cemiplimab no observation time is stipulated with reference to cemiplimab. Treatment to be continued up to 24 mon or until confirmed disease progression, unacceptable toxicity, or patient withdrawal of consent.

y Hematology assessment should be performed at screening, ≤24 h prior to each drug administration, on treatment discontinuation, and at 30 d SFU1 or if clinically indicated (see Section 10.2 for assessments to be done). Samples should be shipped to a central laboratory. Results need to be reviewed prior to drug administration based on local laboratory values. Unscheduled assessments can be performed locally.

z Serology assessments for patients will include HBsAG, anti-HBc, anti-HBs and anti-HCV. HIV (1/2) testing required if the patient has never been tested for HIV or if testing was prior to primary cancer diagnosis or performed ≥12 mon prior to trial start. Samples should be shipped to a central laboratory.

aa Chemistry panel should be performed at screening, ≤24 h prior to each drug administration, on treatment discontinuation, and at 30 d SFU1 or if clinically indicated (see Section 10.2 for assessments to be done). Samples should be assessed by both central and local laboratory. Results of local laboratory assessments need to be reviewed prior to drug administration . Unscheduled assessments can be performed locally or sent to the central laboratory.

bb Coagulation assessments to be performed at screening and if clinically indicated (see Section 10.2 for assessments to be done). Samples should be shipped to a central laboratory; unscheduled assessments can be performed locally.

cc TSH and free T4 will be assessed at screening, once every 2 cycles, on treatment discontinuation, at SFU1, and as clinically indicated. Samples should be shipped to a central laboratory. Results need to be reviewed prior to drug administration. Unscheduled assessments can be performed locally.

dd All women of childbearing potential will have a serum pregnancy test at screening, within 10 d prior to initiation of trial treatment. Urine or serum pregnancy test will be performed at every subsequent cycle or according to local requirements and at EoT and SFU1. If a urine pregnancy test is positive, hold dosing and confirm the result with a serum pregnancy test.

ee For urinalysis assessments dipsticks are to be used, microscopic evaluation should be performed per investigator's discretion. Tests will be done locally with material provided by the central laboratory.

ff Peripheral blood for assessment of cytokines will be drawn from all patients of Arms 1 and 2 prior to administration of BNT111 and 4-6 h post application. The tests will be performed centrally.

gg Only applicable for Cycle 3.

hh At selected trial sites, as part of the biomarker sub-study, blood for isolation of PBMCs for analysis of cellular immune responses should only be taken in patients randomized to Arm 1 or 2 at the following timepoints: prior to the first dosing with BNT111, on C3D1, C5D1, C8D1 and every six cycles afterwards, and at EoT. At C5D1, blood draws may be replaced by a leukapheresis.

ii At selected trial sites, paired biopsies at baseline and under treatment at approximately 6 to 8 wks after treatment start and optionally at progression will be performed. This serial tumor biopsy will be performed, if considered clinically feasible by the investigator. For other trial sites, if any biopsies or resections are to be performed as part of standard clinical care, tissue samples from such procedures should be provided (if feasible) for the sponsor for biomarker testing.

jj All patients must provide a tumor tissue sample (formalin fixed paraffin-embedded [FFPE] blocks/slides) from a fresh biopsy collected before Visit C1D1, or archival tissue., which should be shipped to the central laboratory only following confirmation of eligibility. The archival tissue can be an FFPE block (not older than 3 years) or freshly cut slides (special storage conditions and immediate shipment to a specialty lab are required), with tissue preferably derived from advanced disease stage. The fresh or archival tissue must have been taken using core needle biopsy (18-gauge needle or larger) or surgical excision. Note: fine needle aspirate biopsies will not be acceptable.

kk Optional whole blood sampling at screening if patient has signed ICF for pharmacogenomic analyses.

ll At selected trial sites, as part of the biomarker sub-study, a blood sample for HLA typing should be taken which should be stored at -80°C.

Abbreviations: AJCC = American Joint Committee on Cancer; d = day; CT = computer tomography; ECG = electrocardiogram; ECOG PS = Eastern Cooperative Oncology Group Performance Status; d = day(s); eCRF = electronic case report form; EoT = end of treatment; FU = follow-up; h = hour(s); HBc = hepatitis B core antigen; HBs = hepatitis B surface antigen; HBsAG = hepatitis surface antigen; HCV = hepatitis C virus; HIV = human immunodeficiency virus; HLA = human leukocyte antigen; INR = international normalized ratio; mon = month; MRI = magnetic resonance imaging; PBMC = peripheral blood mononuclear cells; PRO/QoL = patient reported outcomes/quality of life; SOA = schedule of activity; TNM = tumor, node, metastasis; TSH = thyroid stimulating hormone; wks = week(s).

10.4.2 Schedule of activities for patients receiving add-on treatment (SOA2)

SOA2	Baseline add-on therapy	Treatment period							End of treatment [a]	Post-treatment follow-up period			
Cycle=21 d		Add-on Cycle 1			Add-on Cycle 2			Add-on Cycle 3 onward		Safety FU [b]		Efficacy FU [c]	Survival FU [d]
Visit		AOC1D1 [e]	AOC1D8	AOC1D15	AOC2D1	AOC2D8	AOC2D15	AOCnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤10 d prior to AOC1D1	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window			±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Assessments and procedures													
Concomitant medication and therapy [h]	X	X [i]	X	X	X	X	X	X	X	X	X		
Informed consent	X												
Vital signs [j]	X	X [i]	X	X	X	X	X	X	X	X			
Weight	X	X	X	X	X	X	X	X	X	X			
Physical examination [k]	X	X [i]	X	X	X	X	X	X	X	X			
ECOG PS	X	X [i]	X	X	X	X	X	X	X	X			
12-lead ECG [l]	X	Perform, if clinically indicated							X				
CT and/or MRI [m]	X [m, n]	X [m]										X [m]	
Brain imaging	X [o]	X [o]										X [o]	
Adverse events [p]	X	X [i]	X	X	X	X	X	X	X	X	X		

SOA2	Baseline add-on therapy	Treatment period							End of treatment [a]	Post-treatment follow-up period			
Cycle=21 d		Add-on Cycle 1			Add-on Cycle 2			Add-on Cycle 3 onward		Safety FU [b]		Efficacy FU [c]	Survival FU [d]
Visit		AOC1D1 [e]	AOC1D8	AOC1D15	AOC2D1	AOC2D8	AOC2D15	AOCnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤10 d prior to AOC1D1	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window			±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
PRO/QoL assessments [q]		X			X			X	X	X			
PGIS		X						AOC4D1, AOC8D1	X				
PGIC								AOC4D1, AOC8D1	X				
Survival status													X
Subsequent anti-cancer therapy									X	X	X	X	X
BNT111 [r]		X	X [s]	X [s]	X	X [s]	X [s]	X					
Cemiplimab (Q3W) [t]		X			X			X					
Hematology [u]		X	X	X	X	X	X	X	X	X			
Blood chemistry [v]	X	X [i]	X	X	X	X	X	X	X	X			
Coagulation (INR and aPTT) [w]	X	Perform, if clinically indicated							X	Perform, if clinically indicated			
TSH and free T4 [x]	X	X [x]							X	X			

SOA2	Baseline add-on therapy	Treatment period							End of treatment [a]	Post-treatment follow-up period			
Cycle=21 d		Add-on Cycle 1			Add-on Cycle 2			Add-on Cycle 3 onward		Safety FU [b]		Efficacy FU [c]	Survival FU [d]
Visit		AOC1D1 [e]	AOC1D8	AOC1D15	AOC2D1	AOC2D8	AOC2D15	AOCnD1 [f]	EoT	SFU1	SFU2	EFU	LTFU
Day	≤10 d prior to AOC1D1	1	8	15	22	29	36	X [g]	0-21 d after last dose	30 d after last dose	90 d after last dose	6/9/12 wks	Every 3 mon after last trial assessment
Visit window			±1 d	±1 d	±1 d	±1 d	±1 d	±3 d		+5 d	±7 d	±7/14 d	±14 d
Pregnancy test [y]	X	X [y]							X				
Urinalysis [z]	X	Perform, if clinically indicated							X	X			
Cytokines [aa]		X			X			X [bb]					
Cellular immune responses [cc]	X [cc]				X [cc]			X [cc]	X [dd]				

HLA haplotyping[ee]	X [ee]												
Tumor biopsy [ff]	X [ff]							X [ff]					

Note: On treatment days, all assessments should be performed prior to dosing unless otherwise specified.

a Patients who discontinue trial treatment for whatever reason will return to the clinic for End of Treatment (EoT) visit within 21 d after last dose. The visit at which response assessment shows (confirmed) progressive disease may be used as the EoT visit. The EoT visit and post-treatment follow-up period will only be performed after discontinuation of both trial drugs.

b All AEs will be reported until 90 d after the last dose of trial treatment or until a new anti-cancer therapy is started. Additionally, in case a new anti-cancer therapy is started, the safety follow-up visit may be omitted.

c Patients who discontinue trial therapy prior to disease progression should continue to have regular efficacy assessments as scheduled per protocol, until disease progression, the start of next-line anti-cancer therapy, withdrawal of consent, or death, whichever occur first.

d After the last trial visit, information on survival status and new anti-cancer therapy will be collected via phone calls, patient medical records, and/or clinic visits approximately every 3 mon until death (unless the patient withdraws consent for survival follow-up or the sponsor terminates the trial). If a patient asks to be withdrawn from survival follow-up, this request must be documented in the source documents and signed by the investigator.

e Visit AOC1D1 should be no more than 6 wks after last trial treatment given.

f "n" represents the number of the respective cycle (e.g., AOC3 means the third cycle of add-on therapy).

g Visit X should occur 21 (± 3) d after the first day (i.e., treatment visit) of the previous cycle.

h Continued record of concomitant medication and therapy after add-on, includes any therapies (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated trial treatment until the SFU2 visit.

i If Baseline and AOC1D1 are on the same day, assessment only needs to be taken once.

j Vital signs (body temperature, pulse rate, respiratory rate, and blood pressure) should be measured prior to the administration of the trial treatments, and at 30 (±10) min, 60 (±10) min, 4 h (±10 min), and 6 h (±10 min) after administration, or if clinically indicated (see protocol for guidelines on reduction or expansion of these time points).

k At baseline and at end of treatment assessments must include a full physical examination. Record abnormalities observed on the respective eCRF pages as appropriate. For all subsequent physical examinations perform a limited, symptom-directed examination at specified time points and as clinically indicated at other time points. Record new or worsened clinically significant abnormalities on the Adverse Event eCRF.

l Single 12-lead ECG recordings will be obtained during baseline and as clinically indicated at other time points. Patients should be resting in a supine position for at least 10 min prior to ECG recording. The interpretation of the ECG recording will be done locally by the investigator.

m Tumor response will be measured by RECIST 1.1 / iRECIST. Patients will undergo tumor assessments every 6±1 wks for the 3 mon after add-on treatment start, every 9±1 wks for the next 9 mon, and every 12±2 wks thereafter, regardless of dose delays, until (confirmed) disease progression (even if the patients discontinue trial treatment for other reasons) or until the start of next-line anti-cancer therapy (whichever occurs first). An extra scan should be scheduled if clinically indicated.

n Baseline image may not be older than 21 d prior to AOC1D1. Target lesions and non-target lesions can be redefined at add-on baseline assessment.

o Patients with new brain lesions identified at baseline for add-on therapy cannot continue with add-on therapy, patients with known brain lesions during monotherapy should have stable brain lesions at baseline for add-on therapy, the lesions should be followed and assessed at all subsequent tumor assessments. If no brain lesions are present at baseline for add-on therapy, brain imaging does not need to be repeated, only if clinically indicated as per investigator judgment.

p All adverse events will be documented in the eCRF, and all serious adverse events should be reported to the sponsor within the appropriate time frame. All adverse events will be reported until 90 d after the last dose of trial treatment (i.e., SFU2 visit). In addition, the sponsor should be notified (even after SFU2 visit), if the investigator becomes aware of any serious adverse event that is believed to be related to prior exposure to trial treatment or a trial procedure.

q Patient reported outcomes/Quality of life assessment questionnaire must be completed before any assessments are done or any treatment is given on the dosing day. Patients will provide PRO/QoL assessments during trial visits at add-on baseline, AOC2D1, AOC3D1, AOC4D1, every 9±1 wks for the next 9 mon, and every 12±2 wks thereafter until EoT.

r After administration of BNT111, initially an observation time of at least 6 h is required (see protocol Section 6.8.4.1). Treatment to be continued up to 24 mon or until confirmed disease progression, unacceptable toxicity or patient withdrawal of consent.

s Weekly dosing applicable for patients cemiplimab monotherapy (Arm 3). Also applicable to patients who have received BNT111 monotherapy (Arm 2) if the patient has previously received ≤ 4 doses of BNT111.

t Cemiplimab will be given according to valid Prescribing Information in the country every 3 wks (±3 d). The first administration of cemiplimab will be on Cycle 1 Day 1, followed by every 3 wks (±3 d) thereafter. Cemiplimab will be administered subsequently after administration of BNT111 with at least 30 min interval. There should be an Observation time of 6 h after administration of cemiplimab should be performed as outlined in protocol Section 6.8.4.1. Treatment to be continued up to 24 mon or until confirmed disease progression, unacceptable toxicity or patient withdrawal of consent.

u Hematology assessment should be performed at baseline add-on therapy,≤24 h prior to each drug administration, on treatment discontinuation, and at 30 d SFU1 or if clinically indicated (see Section 10.2 for assessments to be done). Samples should be assessed by both a central and local laboratory. Results of local laboratory assessments need to be reviewed prior to drug administration based on local laboratory values. Unscheduled assessments can be performed locally or sent to the central laboratory.

v Chemistry panel should be performed at baseline add-on therapy,≤24 h prior to each drug administration, on treatment discontinuation, and at 30 d SFU1 or if clinically indicated (see Section 10.2 for assessments to be done). Samples should be shipped to a central laboratory. Results need to be reviewed prior to drug administration based on local laboratory values. Unscheduled assessments can be performed locally.

w Coagulation assessments are to be performed at baseline add-on therapy, at EOT, and if clinically indicated (see Section 10.2 for assessments to be done). Samples should be shipped to a central lab; unscheduled assessments can be performed locally.

x TSH and free T4 will be assessed at baseline add-on therapy, once every 2 cycles, on treatment discontinuation, and at SFU1. Samples should be shipped to a central laboratory and if required, to the local laboratory. Results need to be reviewed prior to drug administration. Unscheduled assessments can be performed locally or sent to the central laboratory.

y All women of childbearing potential will have a serum pregnancy test at baseline, within 10 d prior to initiation of trial treatment. Urine or serum pregnancy test will be performed at every third subsequent cycle or according to local requirements and at EoT and SFU1. If a urine pregnancy test is positive, hold dosing and confirm the result with a serum pregnancy test.

z For urinalysis assessments dipsticks are to be used, microscopic evaluation should be performed per investigator's discretion.

aa Peripheral blood for assessment of cytokines will be drawn from all patients prior to administration of BNT111 and 4-6 h post application. The tests will be performed centrally. This is only applicable for first 3 cycles with BNT111, if cytokines were collected in Arm 2 for 3 cycles, no further cytokines need to be collected.

bb Only applicable for Cycle 3.

cc At selected trial sites, as part of the biomarker sub-study, blood for isolation of PMBCsPBMCs for analysis of cellular immune responses should be taken during baseline (unless previous blood draw was < 7 wks, then the baseline sample can be omitted) and on AOC2D1AOC3D1, AOC5D1, AOC8D1, every six cycles afterwards, and at EoT. This applies only to patients in Arm 2 and Arm 3 add-on therapy who have consented to the biomarker sub-study.

dd Should be done within 3 wks after BNT111 administration.

ee Only for patients entering Arm 3 add-on therapy who have not yet provided a sample for HLA typing: a blood sample for HLA typing should be taken which should be stored at -80°C

ff At selected trial sites, as part of the biomarker sub-study, optional biopsies at baseline, under treatment and at progression will be performed. The tumor biopsy will be performed, if considered clinically feasible by the investigator. This applies only to patients in Arm 2 and Arm 3 add-on therapy who have consented to the biomarker sub-study. The tumor sample can be from a core biopsy or from resected tumor tissue. Fine needle aspirates and cytological specimens are not acceptable. In patients who have had biopsies taken during BNT111 monotherapy, the number of biopsies during add-on is to be adjusted as appropriate (up to three in total during the duration of the trial). It is recommended to perform the biopsy on progression only if 4 wks or more have elapsed since the last biopsy (on-treatment biopsy). If the time is shorter, no biopsy on progression should be taken. For other trial sites, if any biopsies or resections are to be performed as part of standard clinical care, tissue samples from such procedures should be provided (if feasible) for the sponsor for biomarker testing.

Abbreviations: AOCnDn = add-on, cycle "n", day "n"; d = day; CT = computer tomography; ECG = electrocardiogram; ECOG PS = Eastern Cooperative Oncology Group Performance Status; d =day(s); eCRF = electronic case report form; EoT = end of treatment; FU = follow-up; h = hour(s); INR = international normalized ratio; mon = month; MRI = magnetic resonance imaging; PBMC = peripheral blood mononuclear cells; PRO/QoL = patient reported outcomes/quality of life; SOA = schedule of activity; TSH = thyroid stimulating hormone; wks = week(s).

10.5 Appendix 5: Censoring rules

Table 5.1 Censoring Rules for PFS, DOR & DDC Based on Food & Drug Administration (FDA) Guidelines

For calculation of 2 or more missed consecutive scheduled assessments, see the text after the censoring rules tables.

If a subject met multiple censoring criteria, the earliest censoring date will be used.

Situation (Censoring Order)	Date Patient Has Event or is Censored	Situation Outcome
Inadequate baseline assessment* (1)	Randomization date	Censored
Inadequate post-baseline assessments* (2)	Randomization date	Censored
* Not applicable to DOR and DCR		
Progression documented on or between scheduled tumor assessments	Date of first objective tumor assessment showing objective progression	Progressed (Event)
Progression documented not after 2 or more missed consecutive scheduled assessments during the treatment period or not after 1 or more missed scheduled assessments during the Follow-up period	For RECIST 1.1, the Date of objective progression / for iRECIST, the date of iUPD if progression associated with that iUPD has been confirmed.	Progressed (Event)
Death (and progression has not occurred previously) not after 2 or more missed consecutive scheduled assessments during the treatment period or not after 1 or more missed scheduled assessments during the Follow-up period	Date of death	Death (Event)

Situation (Censoring Order)	Date Patient Has Event or is Censored	Situation Outcome
Death without objective progression	Date of death	Death (Event)
Death or progression after 2 or more consecutive missing scheduled assessments during the treatment period or after 1 or more missed scheduled assessments during the Follow-up period (3)	Date of last adequate assessment with evidence of no progression; if no adequate assessment exists then censored at the randomization date.	Censored
New anti-cancer therapy started prior to progression or death (4)	Date of last adequate assessment with evidence of no progression before anti-cancer therapy; if no adequate assessment exists then censored at the randomization date	Censored
No progression or death (6)	Date of last adequate assessment with evidence of no progression; if no adequate assessment exists then censored at the randomization date	Censored
Treatment discontinuation for undocumented progression (this reason would be assumed if for the end of treatment (EOT) CRF page "Other" has been ticked and details are "clinical progression") (5)	Date of last objective tumor assessment prior to treatment discontinuation *For DOR & DDC: Apply censoring on Day 1, if no post-baseline assessment available meeting objective response (for DOR) or disease control (for DDC) up to EOT with qualifying reasons of treatment discontinuation*	Censored

Situation (Censoring Order)	Date Patient Has Event or is Censored	Situation Outcome
Treatment discontinuation due to toxicity (due to an adverse event that is related to IMP ie BNT111 or Cemiplimab) or for the 'other' category on the CRF (7) with a documented disease progression	Date of first documented progression including protocol specified assessments continued through follow-up period	Event
Treatment discontinuation due to toxicity (due to an adverse event that is related to IMP ie BNT111 or Cemiplimab) or for the 'other' category on the CRF (8) without documented disease progression	Date of last objective tumor assessment including protocol specified assessments continued through follow-up period	Censored
Lost to follow-up (9)	Date of last adequate assessment with evidence of no progression; if no adequate assessment exists then censored at the randomization date.	Censored

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

1. Subject has 1 consecutive missed response assessments = subject has more than 6*7+7+7 days between actual consecutive response assessments for 3 months or 9*7+7+7 for next 9 months or 12*7+14+14 thereafter
2. Subject has 2 consecutive missed response assessments = subject has more than 2*6*7+7+7 days between actual consecutive response assessments for 3 months or 2*9*7+7+7 for next 9 months or 2*12*7+14 + 14 thereafter

If the missing assessments are in the first 3-month period and there are assessments present after the first 3 months then the following is done:

1. Subject has week 12 response assessment missing = subject has more than 9*7+7+7 days between actual consecutive response assessments
2. Subject has 2 consecutive missed response assessments in the first 3 months = subject has more than 6*7+9*7 +7 + 7 days between actual consecutive response assessments

Note: in 6*7+7+7 the 6*7 is number of days between response assessments as per schedule because the assessment window is ±7 days. Similarly, the window ±14 is applied per assessment schedule.

Table 5.2 Event and Censoring Rules for OS Based on Food & Drug Administration (FDA) Guidelines

Situation	Date Patient Has Event or is Censored	Situation Outcome
Death	Date of death	Event
Alive	Date last known to be alive (date of last contact collected in the CRF). *Check this date vs the latest date for a patient, if discrepant then raise for resolution.*	Censored

10.6 Appendix 6: EORTC-QLQ-C30

Table 8.2 Scoring of QLQ-C30

	Scale	*Number of items*	*Item range**	*Version 3.0 Item numbers*	*Function scales*
Global health status / QoL					
Global health status / QoL	*QL*	*2*	*6*	*29, 30*	
Functional scales					
Physical functioning	*PF*	*5*	*3*	*1 to 5*	*F*
Role functioning	*RF*	*2*	*3*	*6, 7*	*F*
Emotional functioning	*EF*	*4*	*3*	*21 to 24*	*F*
Cognitive functioning	*CF*	*2*	*3*	*20, 25*	*F*
Social functioning	*SF*	*2*	*3*	*26, 27*	*F*
Symptom scales / items					
Fatigue	*FA*	*3*	*3*	*10, 12, 18*	
Nausea and vomiting	*NV*	*2*	*3*	*14, 15*	
Pain	*PA*	*2*	*3*	*9, 19*	
Dyspnea	*DY*	*1*	*3*	*8*	
Insomnia	*SL*	*1*	*3*	*11*	
Appetite loss	*AP*	*1*	*3*	*13*	
Constipation	*CO*	*1*	*3*	*16*	
Diarrhea	*DI*	*1*	*3*	*17*	
Financial difficulties	*FI*	*1*	*3*	*28*	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

All of the scales and single-item measures range in score from 0 to 100. A high scale score represents a higher response level. Therefore, a high score for a functional scale represents a high/healthy level of functioning, a high score for the global health status / QoL represents a high QoL, but a high score for a symptom scale/item represents a high level of symptomatology / problems. The scoring procedure for each of the scales is the same and consists of computing the raw score (RS) and then computing the actual scale score (S) by making a linear transformation to standardize the score to values from 0 to 100 as shown below.

For all scales, the *RawScore*, *RS*, is the mean of the component items:

$$Raw\ Score = RS = \{I_1 + I_2 + \cdots + I_n\}/n$$

Then for **Functional scales**:

$$Score = \left\{1 - \frac{RS - 1}{range}\right\} \times 100$$

and for **Symptom scales / items** and **Global health status / QoL:**

$$Score = \left\{\frac{RS - 1}{range}\right\} \times 100$$

Where I_1, I_2, … I_n are the individual items and Range is the difference between the maximum possible value of RS and the minimum possible value. The range of RS equals the range of the item values. Most items are scored 1 to 4, giving Range = 3. The exceptions are the items contributing to the Global Health Status / QoL, which are 7-point questions with Range = 6.

Missing items

Score will be derived if at least half of the items from the scale are not missing, otherwise it will be set to missing.

10.7 Appendix 7: Algorithm for visit window mapping of tumor assessments for depth of response

All data points for tumor assessments will be used. This section explains how to assign visits to the assessments for the depth of response analysis.

Data is collected by evaluation number. This evaluation number will be used to group assessments/records for same targeted visit. The tumor assessment data, including unscheduled visit data, will be mapped per the details described below:

Derive AVISIT1/AVISIT1N based on tumor assessment date.

- Calculate trial day of tumor scan performed referenced to first dose date.
- Define target assessment day and analysis window range (lower and upper limits) based on protocol specified schedule.
- Check where the calculated trial day falls and assign AVISIT1.
- AVISIT1N will represent week number of the derived visit (numeric part of AVISIT1).
- For the summary analysis by visit, the assessment closest to the target day will be analyzed. In case, 2 assessments are equidistant, the earliest assessment will be used in the analysis.

Some illustrative examples:

AVISIT1	AVISIT1N	AWTARGET1 (Analysis Window Target)	AWLO1 (Analysis Window Beginning Timepoint)	AWHI1 (Analysis Window Ending Timepoint)
Screening	0	-28	-28	-1
Week 6	6	43	1	63
Week 12	12	85	64	116
Week 21	21	148	117	179
Week 30	30	211	180	242
Week 39	39	274	243	305
Week 48	48	337	306	378
Week 60	60	421	379	462
Week 72	72	505	463	546
Week 84	84	589	547	630
Week 96	96	673	631	714
Week 108	108	757	715	798
…				
Add-on Week 6	1006	43	1	63
Add-on Week 12	1012	85	64	116
Add-on Week 21	1021	148	117	179
Add-on Week 30	1030	211	180	242
Add-on Week 39	1039	274	243	305
Add-on Week 48	1048	337	306	378
Add-on Week 60	1060	421	379	462
Add-on Week 72	1072	505	463	546
Add-on Week 84	1084	589	547	630
Add-on Week 96	1096	673	631	714
Add-on Week 108	1108	757	715	798
…				

10.8 Rules for assigning date of overall response from tumor assessment page

		Each Evaluation #		
Overall Timepoint Response	*Overall Timepoint Response date*	Target Lesion Date	Non-Target Lesion Date	New Lesion Date
PD	min(dt1, dt2, dt3)	dt1*	dt2*	dt3*
CR	max(dt1, dt2)	dt1	dt2	-
PR	dt1	dt1	any	-
SD (Not a CR, PR or PD)	min(dt1, dt2)	dt1	dt2	-
NA, NE	min(dt1, dt2, dt3)	dt1	dt2	dt3

** indicates a PD response by individual Target or Non-Target or appearance of New Lesion, otherwise corresponding date can be excluded while picking the earliest scan date.*

Note: For BICR data, Non-CR/Non-PD and No Evidence of Disease responses can be present. The same rule as for SD will be followed.