Official Protocol Title:	A Phase 3, randomized, double-blind trial of pembrolizumab (MK-3475) with or without lenvatinib (E7080/MK-7902) in participants with treatment-naïve, metastatic nonsmall cell lung cancer (NSCLC) whose tumors have a tumor proportion score (TPS) greater than or equal to 1% (LEAP-007)
NCT number:	NCT03829332
Document Date:	15-Sep-2021

Title Page

THIS PROTOCOL AMENDMENT AND ALL OF THE INFORMATION RELATING TO IT ARE CONFIDENTIAL AND PROPRIETARY PROPERTY OF MERCK SHARP & DOHME CORP., A SUBSIDIARY OF MERCK & CO., INC., NJ, U.S.A. (MSD).

Protocol Title: A Phase 3, randomized, double-blind trial of pembrolizumab (MK-3475) with or without lenvatinib (E7080/MK-7902) in participants with treatment-naïve, metastatic nonsmall cell lung cancer (NSCLC) whose tumors have a tumor proportion score (TPS) greater than or equal to 1% (LEAP-007)

Protocol Number: 007-06 (E7080-G000-314)

Compound Number: MK-7902 (E7080/lenvatinib) and MK-3475 (pembrolizumab)

Sponsor Name:

Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc.
(hereafter referred to as the Sponsor or MSD)

The study is co-funded by MSD and Eisai.

Legal Registered Address:

One Merck Drive

P.O. Box 100

Whitehouse Station, New Jersey, 08889-0100, U.S.A.

Regulatory Agency Identifying Number(s):

IND	140221
EudraCT	2018-003794-98

Approval Date: 15 September 2021

Sponsor Signatory

Typed Name: Title:	Date

Protocol-specific Sponsor contact information can be found in the Investigator Study File Binder (or equivalent).

Investigator Signatory

I agree to conduct this clinical study in accordance with the design outlined in this protocol and to abide by all provisions of this protocol.

Typed Name: Title:	Date

DOCUMENT HISTORY

Document	Date of Issue	Overall Rationale
Amendment 06	15-SEP-2021	Data from an interim safety and futility eDMC for LEAP-007 (data cutoff: 19-MAY-2021) indicated the study met the prespecified nonbinding futility criteria for OS for the combination of lenvatinib plus pembrolizumab compared with placebo plus pembrolizumab and the benefit-risk assessment was not considered positive to continue the study. This futility analysis was requested at a previous eDMC meeting and conducted as described in the sSAP. Based upon these data, MSD implemented this amendment to unblind the study and remove lenvatinib and matching placebo from the study. In addition, the Second-course Treatment Phase was discontinued and collection of ePRO data stopped. The study will remain open to allow ongoing participants to continue treatment with open-label pembrolizumab monotherapy up to a maximum of 35 administrations.
Amendment 05	02-MAR-2021	To update the assumptions and timing of the analyses in the SAP, allow sufficient duration of follow-up, and to remove the initial alpha of ORR and reallocate to primary endpoints of PFS and OS
Amendment 04 – China-specific	11-FEB-2020	Extension of the enrollment period beyond the global study to achieve required exposure and number of events to investigate efficacy and safety in Chinese 1L NSCLC participants
Amendment 03	16-DEC-2019	Alignment with Lenvatinib and Pembrolizumab (LEAP) program protocol standards and Health Authority request
Amendment 02	27-JUN-2019	Alignment with Lenvatinib and Pembrolizumab (LEAP) program protocol standards

Document	Date of Issue	Overall Rationale
Amendment 01	17-JAN-2019	Alignment with Lenvatinib and Pembrolizumab (LEAP) program protocol standards
Original Protocol	09-NOV-2018	N/A

PROTOCOL AMENDMENT SUMMARY OF CHANGES

Amendment: 06

Overall Rationale for the Amendments:

Data from an interim safety and futility eDMC for LEAP-007 (data cutoff: 19-MAY-2021) indicated the study met the prespecified nonbinding futility criteria for OS for the combination of lenvatinib plus pembrolizumab compared with placebo plus pembrolizumab and the benefit-risk assessment was not considered positive to continue the study. This futility analysis was requested at a previous eDMC meeting and conducted as described in the sSAP. Based upon these data, MSD implemented this amendment to unblind the study and remove lenvatinib and matching placebo from the study. In addition, the Second-course Treatment Phase was discontinued and collection of ePRO data stopped. The study will remain open to allow ongoing participants to continue treatment with open-label pembrolizumab monotherapy up to a maximum of 35 administrations.

Summary of Changes Table:

Section # and Name	Description of Change	Brief Rationale
1.1 Synopsis 1.2 Schema	Revised text and added notes to clarify the removal of lenvatinib and placebo treatments, Second-course Phase, collection of ePRO, and collection of posttreatment imaging. Extension Study in China was also stopped	In accordance with the overall rationale for the amendment, administration of lenvatinib and matching placebo, the Second-course Treatment Phase, and the collection of ePRO data and posttreatment imaging have been removed from the study.

Section # and Name	Description of Change	Brief Rationale
1.3.1 Schedule of Activities – Initial Treatment Phase	Added a clarifying note regarding the removal of lenvatinib and matching placebo. Revised table to remove lenvatinib and matching placebo administration, collection of ePRO data and posttreatment imaging, collection of most of the blood biomarkers, and all procedures/assessments for cycles ≥36. Updated vital signs, ECG, and urine testing requirements.	In accordance with the overall rationale for the amendment, lenvatinib and matching placebo administration have been removed and assessments/procedures simplified.
1.3.2 Schedule of Activities – Second-course Treatment Phase	Removed table and added a clarifying note.	In accordance with the overall rationale for the amendment, the Second-course Treatment Phase has been removed from the study.
2 Introduction 2.2 Background	Added notes to clarify the removal of lenvatinib and matching placebo. Sections have been left unchanged for reference.	In accordance with the overall rationale for the amendment, lenvatinib and matching placebo administration have been removed, but background information is maintained for reference.
2.3 Benefit/Risk Assessment	Added a note to clarify that the benefit/risk assessment of pembrolizumab plus lenvatinib is no longer applicable. Section updated accordingly.	Lenvatinib and matching placebo have been removed from the study; thus, benefit/risk assessment of pembrolizumab plus lenvatinib is no longer relevant.
3. Hypotheses, Objectives, and Endpoints	Added a note to clarify that efficacy endpoints will no longer be collected, and exploratory objectives may not be further pursued.	As the study did not meet the prespecified criteria for OS, the scope of the study is reduced and further collection of efficacy data is not required.

Section # and Name	Description of Change	Brief Rationale
4.1 Overall Design	Revised text and added notes to clarify the removal of lenvatinib and placebo treatments from the study. Extension Study in China was also stopped.	In accordance with the overall rationale for the amendment, administration of lenvatinib and matching placebo and the Second-course Treatment Phase have been removed from the study.
4.2.1.3 Patient-reported Outcomes 8.2.2 Patient-reported Outcomes	Added note to clarify that PRO data will no longer be collected.	In accordance with the overall rationale for the amendment, PRO data are no longer being collected.
4.3.1 Maximum Dose/Exposure for This Study	Revised text and added note to clarify the removal of lenvatinib, placebo, and the Second Course from the study.	Administration of lenvatinib and matching placebo and the Second-course Treatment Phase have been removed from the study.
4.3.2.1 Lenvatinib Dosing	Added note stating that this section is no longer applicable.	Lenvatinib and matching placebo administration have been removed, but rationale for lenvatinib dosing is maintained for reference.
6 Study Intervention	Added notes to clarify the removal of lenvatinib and matching placebo. Section updated accordingly.	Lenvatinib and matching placebo have been removed from the study.
6.3.3 Blinding 8.1.10 Participant Blinding/Unblinding	Added note to clarify that blinding is no longer applicable and removed the content in these sections.	Lenvatinib and matching placebo have been removed from the study. All participants will receive open-label pembrolizumab monotherapy and the study has been unblinded.

Section # and Name	Description of Change	Brief Rationale
6.6. Dose Modification	Added notes to clarify the removal of lenvatinib and matching placebo. Subsections related to dose modifications due to lenvatinib toxicity or the toxicity of lenvatinib in combination with pembrolizumab removed or updated as appropriate.	Lenvatinib has been removed from the study.
6.8 Clinical Supplies Disclosure	Added note stating that this section is no longer applicable and removed the content.	Lenvatinib and matching placebo have been removed from the study and the study has been unblinded.
7.1 Discontinuation of Study Intervention	Added note to clarify the removal of lenvatinib and matching placebo. Removed criteria related to ALT/AST elevation specific to toxicity of the combination of lenvatinib + pembrolizumab. Replaced with a general criterion noting that discontinuation criteria specific to pembrolizumab-related toxicity as defined in Section 6.6 should be followed.	Lenvatinib has been removed from the study.

Section # and Name	Description of Change	Brief Rationale
7.1.1 Second-course Treatment Phase 8.2.1.4 Second-course Treatment Tumor Scans	Deleted content and added a note to clarify that the Second Course has been removed from the study.	Second-course Treatment Phase has been removed from the study.
8.1.5.3 Concomitant Medications	Added a note to clarify that the Second Course has been removed from the study. Section updated accordingly.	Second-course Treatment Phase has been removed from the study. Recording of concomitant medications for the Second Course is no longer required.
8.1.8.1.1 Lenvatinib/Matching Placebo	Added clarifying note due to removal of lenvatinib and matching placebo from the study.	Lenvatinib and matching placebo have been removed from the study.
8.2.1 Tumor Scans and Assessments of Disease	Added a note to clarify that imaging obtained will not be assessed, verification of PD is no longer needed, and posttreatment imaging is not required. Subsections and Table 5 updated accordingly. Figure 3 removed.	In accordance with the overall rationale for the amendment, assessments/procedures have been simplified.
8.11.5.2 Efficacy Follow-up Visits	Added a note to clarify that imaging obtained will not be assessed, verification of PD is no longer needed, and posttreatment imaging is not required.	

Section # and Name	Description of Change	Brief Rationale
8.3 Safety Assessments	Added notes to clarify the removal of lenvatinib and matching placebo and updated guidelines for BP and ECG measurements to remove assessments/procedures that were specific to lenvatinib safety monitoring.	Lenvatinib and matching placebo have been removed from the study.
8.5 Treatment of Overdose	Added note to clarify the removal of lenvatinib and matching placebo. Section left unchanged for reference.	Lenvatinib and matching placebo have been removed from the study.
8.9 Biomarkers	Removed the collection of the following biomarkers: • Blood for serum biomarkers • Blood for plasma biomarkers • Blood for RNA analysis • Blood for circulating tumor nucleic acids	In accordance with the overall rationale for the amendment, assessments/procedures have been simplified.
8.11 Visit Requirements	Added notes to clarify that the Second-course Treatment Phase has been removed from the study. Subsections updated accordingly to remove text specific to the Second-course Treatment Phase.	Administration of lenvatinib, matching placebo, and Second-course Treatment Phase have been removed from the study. Therefore, study visit requirements were updated.

Section # and Name	Description of Change	Brief Rationale
9 Statistical Analysis Plan 9.1 Statistical Analysis Plan Summary 9.5 Analysis Populations	Added notes to clarify that the study has been unblinded, enrollment in the Extension Study in China has been stopped, and no further prespecified analysis of primary, secondary, and exploratory endpoints will be performed. Subsections updated accordingly.	In accordance with the reduced scope of the study and discontinuation of efficacy data collection.
9.2 Responsibility for Analyses/In-house Blinding	Added note to clarify that the study has been unblinded. Updated content to reflect the current status of the trial with regards to blinding.	
9.6 Statistical Methods	Added note to clarify that the selected analyses of safety and ePRO endpoints will be performed at end of study, and there will be no further analyses of efficacy endpoints. Sections have been retained for reference.	
9.7 Interim Analyses 9.8 Multiplicity 9.9 Sample Size and Power Calculations	Removed content and added notes to clarify that no interim or final analyses of the study will be performed and sections are no longer applicable.	
10.7.3 Canada-specific Requirements	Section updated to remove lenvatinib references.	Lenvatinib has been removed from the study.

Section # and Name	Description of Change	Brief Rationale
Throughout	All references to lenvatinib and placebo were removed. Where deletion of text could cause confusion due to the original design of the study or where text is useful to understand the overall context of the study, the text has been left unchanged and a note was added. Added clarifying notes regarding removal of lenvatinib and matching placebo from the study.	In accordance with the overall rationale for the amendment.
	Made typographical corrections and minor administrative edits.	To correct typographical errors and clarify intended meaning.

ADDITIONAL CHANGE(S) FOR THIS AMENDMENT:

No additional changes.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 PROTOCOL SUMMARY

1.1 Synopsis

Protocol Title: A Phase 3, randomized, double-blind trial of pembrolizumab (MK-3475) with or without lenvatinib (E7080/MK-7902) in participants with treatment-naïve, metastatic nonsmall cell lung cancer (NSCLC) whose tumors have a tumor proportion score (TPS) greater than or equal to 1% (LEAP-007)

Short Title: Pembrolizumab with or without lenvatinib in TPS ≥1% NSCLC

Acronym: Protocol 007 (LEAP-007)

Hypotheses, Objectives, and Endpoints:

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. All participants remaining on study will continue on open-label pembrolizumab monotherapy. No further analyses of efficacy endpoints will be performed. Exploratory objectives may not be pursued.

Participants were screened for up to 28 days, then randomized, and continued study treatment until discontinuation (defined in Section 7.1) OR stop/unblinding of study intervention due to the futility analysis criterion for OS being met.

In adult participants with treatment-naïve, metastatic NSCLC expressing PD-L1 (TPS ≥1%):

Primary Objectives	Primary Endpoints
- Objective: To compare progression-free survival (PFS) as assessed by blinded independent central review (BICR) according to Response Evaluation Criteria in Solid Tumors, version 1.1 (RECIST 1.1), adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo - Hypothesis (H1): The combination of pembrolizumab + lenvatinib has superior PFS per RECIST 1.1, as defined in Section 4.2.1.1.1, based on BICR	- PFS, defined as the time from randomization to the first documented progressive disease (PD) or death due to any cause, whichever occurs first*

- Objective: To compare overall survival (OS) for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo. - Hypothesis (H2): The combination of pembrolizumab + lenvatinib has superior OS	- OS, defined as the time from randomization to death due to any cause*
Secondary Objectives	Secondary Endpoints
- Objective: To compare objective response rate (ORR) as assessed by BICR according to RECIST 1.1, as defined in Section 4.2.1.1.1, for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo - Hypothesis (H3): The combination of pembrolizumab + lenvatinib has superior ORR per RECIST 1.1, as defined in Section 4.2.1.1.1, based on BICR	- Objective response (OR), defined as a confirmed complete response (CR) or partial response (PR)
- Objective: To evaluate the safety and tolerability for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	- Adverse events (AEs) and study intervention discontinuation due to AEs
- Objective: To compare the mean change from baseline in the global health status/quality of life (QoL), cough, chest pain, dyspnea, and physical functioning for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	- Change from baseline for the following patient-reported outcomes (PROs) scales/items: global health status/QoL (EORTC QLQ-C30 items 29 and 30), cough (EORTC QLQ-LC13 item 31), chest pain (EORTC QLQ-LC13 item 40), dyspnea (EORTC QLQ-C30 item 8), and physical functioning (EORTC QLQ-C30 items 1-5)

- Objective: To compare the time to true deterioration (TTD) in global health status/QoL, cough, chest pain, dyspnea, and physical functioning for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	- TTD, defined as the time from baseline to the first onset of a ≥10-point deterioration from baseline with confirmation by the subsequent visit of a ≥10-point deterioration from baseline in global health status/QoL (EORTC QLQ-C30 items 29 and 30), cough (EORTC QLQ-LC13 item 31), chest pain (EORTC QLQ-LC13 item 40), dyspnea (EORTC QLQ-C30 item 8), and physical functioning (EORTC QLQ-C30 items 1-5) - TTD in the composite endpoint (combination of cough [QLQ-LC13 item 31], chest pain [QLQ-LC13 item 40], or dyspnea [QLQ-C30 item 8]) defined as the time to first onset of a ≥10-point deterioration from baseline in any one of 3 scale items with confirmation by the subsequent visit of a ≥10-point deterioration from baseline in the same scale as the first onset

*This study will be considered to have met its success criteria if the combination of pembrolizumab + lenvatinib is superior to pembrolizumab + matching placebo in PFS or OS.

Overall Design:

Study Phase	Phase 3
Primary Purpose	Treatment
Indication	First-line treatment of metastatic NSCLC with PD-L1 expression (TPS ≥1%)
Population	Adult participants with treatment-naïve, metastatic NSCLC (TPS ≥1%)
Study Type	Interventional
Intervention Model	Parallel This is a multi-site study.

Type of Control	Active Control with Placebo **NOTE: As of Amendment 007-06, the placebo treatment is removed from the study.**
Study Blinding	Double-blind **NOTE: As of Amendment 007-06, the study is unblinded.**
Masking	Participant or Subject Investigator Sponsor
Estimated Duration of Study	The Sponsor estimates that the study will require approximately 5 years from the time the first participant (or their legally acceptable representative) provides documented informed consent until the last participant's last study-related contact. Extension Study in China: The Sponsor estimates that the study will require approximately an additional 1 year (beyond the global study's last participant last study-related contact) from the time the first participant (or their legally acceptable representative) provides documented informed consent until the last participant's last study-related contact. **NOTE: As of Amendment 007-06, enrollment in the Extension Study in China has been stopped.**

Number of Participants:

Global Study: Approximately 620 participants will be randomized.

Extension Study in China: Approximately 120 Chinese participants overall will be enrolled in the global study and the extension study.

NOTE: As of Amendment 007-06, enrollment in the Extension Study in China has been stopped.

Intervention Groups and Duration:

Intervention Groups

Intervention Group Name	Drug	Dose Strength	Dose Level	Route of Admin.	Regimen/ Treatment Period	Use
Arm A	Pembrolizumab	100 mg/vial	200 mg	IV Infusion	Q3W	Standard of care/Experimental
	Lenvatinib	10 mg/4 mg	20 mg	Oral	QD	Experimental
Arm B	Pembrolizumab	100 mg/vial	200 mg	IV Infusion	Q3W	Standard of care/Experimental
	Placebo	NA	NA	Oral	QD	Placebo

Abbreviations: NA = not applicable; Q3W = every 3 weeks; QD = daily.

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the treatment groups. Participants remaining on treatment will receive open-label pembrolizumab monotherapy as per protocol.

Total Number of Intervention Groups/Arms

2 arms

Duration of Participation

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. This section has been updated accordingly.

After a screening phase of up to 28 days, each participant received assigned study intervention until one of the conditions for discontinuation of study intervention was met (defined in Section 7.1).

Participants will be permitted to continue study intervention beyond RECIST 1.1-defined PD as long as the treating investigator considers that the participant may experience clinical benefit with continued treatment as per iRECIST, and the participant is tolerating study intervention, until PD is confirmed by iRECIST. All decisions to continue treatment beyond confirmed PD by iRECIST must be approved by the Sponsor.

Participants who have been on pembrolizumab for at least 24 weeks and attain a CR may consider stopping pembrolizumab treatment.

After the end of treatment, each participant will be followed for the occurrence of AEs and spontaneously reported pregnancy, as described in Section 8.4.5. Participants will not be required to continue with posttreatment follow-up imaging for disease status.

All participants will be followed for OS until death, withdrawal of consent, lost to follow-up, or the end of the study. The end of the study will be when the last participant completes the last study-related telephone call or visit, withdraws from the study, or is lost to follow-up (ie, the participant is unable to be contacted by the investigator).

Study Governance Committees:

Steering Committee	Yes
Executive Oversight Committee	Yes
Data Monitoring Committee	Yes
Clinical Adjudication Committee	No
Study governance considerations are outlined in Appendix 1.	

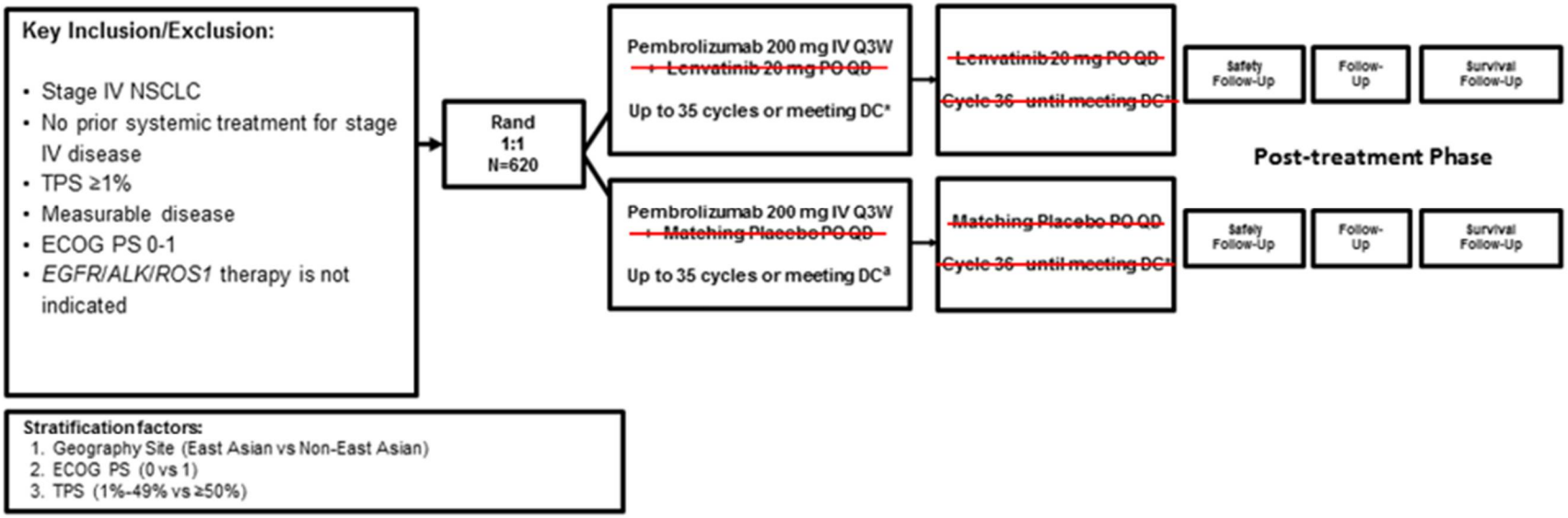
Study Accepts Healthy Volunteers: No

A list of abbreviations used in this document can be found in Appendix 8.

1.2 Schema

NOTE: As of Amendment 007-06, lenvatinib and matching placebo and the Second-course Treatment Phase have been removed from the study. The original study design and updates to this design as of Amendment 007-06 (indicated by strikethrough) are depicted in Figure 1.

Figure 1 Original and Updated Study Design: Initial Treatment Phase



Abbreviations: *ALK* = ALK receptor tyrosine kinase; DC = discontinuation; ECOG PS = Eastern Cooperative Oncology Group Performance Status; *EGFR* = epidermal growth factor receptor; IV = intravenous; NSCLC = non-small cell lung cancer; PD = progressive disease; PO = orally; Q3W = every 3 weeks; QD = once daily; *ROS1* = ROS proto-oncogene 1; TPS = tumor proportion score.
* DC criteria are listed in Section 7.1.

1.3 Schedule of Activities (SoA)

1.3.1 Initial Treatment Phase

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. Participants who remain on treatment will receive open-label pembrolizumab monotherapy as per protocol. Administration of lenvatinib and placebo, collection of ePRO data, and collection of most of the blood biomarkers have been removed from the SoA. Other study procedures specific to lenvatinib/matching have also been removed.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Administrative and General Procedures															
Informed consent	X														ICF must be signed before any protocol-specific screening procedures are performed. Additional consent is required at PD.
Inclusion/exclusion criteria	X														
Participant identification card	X	X													Update at C1D1.
Demographic and medical history	X														

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		+ 7	**± 7**	**± 14**	
Prior/concomitant medications	X	←								→					Prior concomitant medications received within 30 days before the first dose of study intervention through 30 days after the last dose of study intervention (or 90 days if used to treat an SAE) will be recorded.
Treatment randomization		X													Dose within 3 days of randomization.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Study Intervention Administration															
Pembrolizumab administration		X			X		X	X	X	X					Pembrolizumab 200 mg IV Q3W.
Tumor Scans															
Tumor scans and response assessment (thorax, abdomen, pelvis)	X*	←								→	X				*Perform Q9W (63±7 days) from randomization through Week 54, then Q12W (84 ±7 days) subsequently (ie, after Weeks 9, 18, 27, 36, 45, 54, 66, 78, 90, etc.). -Schedule should be followed regardless of treatment delays. -If imaging was obtained within 4 weeks before DC, scan at DC is not mandatory.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Brain MRI	X														-Required at screening. -Follow-up MRI is required if clinically indicated and at time of CR in participants with brain metastasis at Screening. -If MRI is contraindicated or cannot be performed, CT of the head with contrast is acceptable.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Efficacy Procedures															
Subsequent antineoplastic therapy status											X	X	X	X	-All anticancer therapy will be recorded until time of death or termination of survival follow-up. -If a clinic visit is not feasible, follow-up information may be obtained via telephone or e-mail.
Survival status		←								→				X	Participants may be contacted for survival status at any time during the course of the study.
Tumor Tissue Collection															
Newly obtained/archival tissue sample for PD-L1 analysis	X														-May use archival tissue sample obtained before screening period as part of the participant's SOC. -Central PD-L1 result will be blinded to the site.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
EGFR, *ALK*, and *ROS1* molecular status	X														-Not required for participants with squamous histology or K-*ras* mutation. -May send tumor tissue to central laboratory for molecular testing if status is unknown and cannot be determined locally.
Clinical Procedures/Assessments															**All screening procedures should be performed within 28 days of randomization, unless otherwise noted.**
Full physical examination	X*										X				*To be performed within 7 days before start of study intervention.
Directed physical examination		X		X	X	X	X	X	X	X		X			

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Phone contact			X												Participants will be contacted by telephone on C1D8 to assess for development of early toxicity. If early toxicity is suspected, an unscheduled visit can occur before C1D15 if deemed necessary by the investigator.
Height	X														
Vital signs, weight	X	X		X	X	X	X	X	X	X	X	X			Vital signs must be taken in the clinic.
12-lead ECG with QTc measurement	X	X			X						X				ECG at screening, C1D1, C2D1, D1 of every fourth cycle (12 weeks) thereafter (eg, C6, C10, C14, etc.) up until discontinuation. ECGs are only required at the EOT Visit for those participants that had received lenvatinib.
ECHO/MUGA	X										X				Additional assessments may be performed as clinically indicated.
ECOG performance status	X	X			X		X	X	X	X	X	X			At screening, perform within 7 days before C1 but before randomization.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
AE monitoring	X	←								→			X		Report AEs occurring within 30 days after the last dose of study intervention. Report SAEs occurring within 90 days after the last dose of study intervention, or 30 days after the last dose of study intervention if a new anticancer therapy is initiated, whichever is earlier.
Laboratory Procedures/Assessments: Analysis by Local Laboratory															**All screening procedures should be performed within 28 days of randomization, unless otherwise noted.**

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Serum β-HCG or urine pregnancy test (WOCBP only)	X				X		X	X	X	X	X	X			WOCBP require negative test prior to randomization. If more than 24 hours have elapsed prior to first dose of study intervention, another pregnancy test is required prior to starting study intervention. A serum or urine pregnancy test will be performed per Appendix 2. Refer to Sections 5.1, 5.2, 8.3.5 and 10.5.3.
HIV, hepatitis B and C screen	X														Required at baseline if mandated by local health authority.
Hematology and clinical chemistry laboratory assessment	X			X	X		X	X	X	X	X	X			-Performed locally within 10 days before first dose. -After C1, collect within 3 days before dosing.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Urinalysis/urine dipstick testing	X	X							X	X	X	X			-Performed locally within 7 days before first dose. After C1, collect within 3 days before dosing. -Urinalysis is required at screening and every 4 cycles for participants who received lenvatinib/matching placebo. After C1, collect within 3 days before dosing.
PT/INR and aPTT/PTT	X														-Screening samples collected within 10 days of treatment initiation. -Additional testing to be conducted as clinically indicated for participants taking anticoagulant therapy.

Study Period	**Screening**	**Treatment Cycle = 21 Days**									**EOT**	**Posttreatment**			**Notes**
												Safety Follow-up	**Follow-up**	**Survival Follow-up**	**All procedures are to be performed before study intervention administration unless otherwise indicated. Refer to Section 8.11 for visit details.**
Visit Timing/ Cycle Number	**-28 to -1**	**1**			**2**		**3**	**4**	**5**	**6 to 35**	**At DC**	**30 Days post last dose**	**Every 12 weeks**	**Every 12 weeks**	
Cycle Day (±3 days unless otherwise specified)		**1**	**8**	**15**	**1**	**15**	**1**	**1**	**1**	**1**	**–**				
Scheduling Window (Days)		**+ 3**	± 3	±3	± 3	± 3	± 3	± 3	± 3	± 3		**+ 7**	**± 7**	**± 14**	
Thyroid function (T3, T4, and TSH)	X				X			X		X	X	X			At Screening, C2D1, then every 2 cycles thereafter. -Screening samples to be collected within 10 days before first dose. -Participants may be dosed in subsequent cycles after C1D1 while thyroid function tests are pending. -After C1, collect within 3 days before dosing. -Free T3 and Free T4 are acceptable. **-May use central laboratory only if local laboratory is not capable.**
Biomarkers: Analysis by Central Laboratory															
Blood for genetic analysis		X													-Collect before study intervention.

Abbreviations: AE = adverse event; *ALK* = anaplastic lymphoma kinase gene; ANC = absolute neutrophil count; aPTT = activated partial thromboplastin time; BP = blood pressure; CT = computed tomography; CXDY= Cycle X Day Y; DC = discontinuation; ECG = electrocardiogram; ECHO = echocardiogram; ECOG = Eastern Cooperative Oncology Group; *EGFR* = epidermal growth factor receptor; EOT = end of treatment; ePRO = electronic patient-reported outcome; FSH = follicle-stimulating hormone; FT3 = free triiodothyronine; FT4 = free thyroxine; β-HCG = β human chorionic gonadotropin; HIV = human immunodeficiency virus; ICF = informed consent form; INR = international normalized ratio; IV = intravenous; K-*ras* = Kirsten rat sarcoma oncogene; MRI = magnetic resonance imaging; MUGA = multigated acquisition; PD = progressive disease; PD-L1 = programmed cell death ligand 1; PK = pharmacokinetic(s); PO = oral(ly); PT = prothrombin time; PTT = partial thromboplastin time; QXW = every X weeks; QD = once daily; *ROS1* = ROS proto-oncogene 1; SAE = serious adverse event; ; T3 = triiodothyronine; T4 = thyroxine; WOCBP = women of childbearing potential; SoA = schedule of activities; SOC = standard of care.

1.3.2 Second-course Treatment Phase

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from the study.

2 INTRODUCTION

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. Subsections 2.1 Study Rationale and 2.2 Background are based on the original study design and have been left unchanged for reference.

2.1 Study Rationale

The global incidence of lung cancer was 1.8 million in 2012, resulting in an estimated 1.6 million deaths [World Health Organization 2012]. In the United States, the 2018 estimated incidence of new diagnoses was 234,030, and the estimated number of deaths was 154,050 [National Cancer Institute 2018]. NSCLC represents approximately 80% to 85% of all lung cancers. Of patients with NSCLC, tumor histology is approximately 40% to 60% adenocarcinoma, 10% to 15% squamous, 5% neuroendocrine, and the rest, not otherwise specified [Sulpher, J. A., et al 2013].

Approximately 80% of patients with NSCLC have advanced disease not amenable to surgical resection at the time of diagnosis. The 5-year relative survival for patients with any lung cancer overall and metastatic lung cancer specifically has been reported to be 18.6% and 4.7%, respectively [National Cancer Institute 2018].

Pembrolizumab monotherapy is the current standard of care (SOC) for the treatment of patients with an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 and previously untreated advanced or metastatic NSCLC with programmed cell death ligand 1 (PD-L1) TPS ≥50% and with no epidermal growth factor receptor (*EGFR*) or anaplastic lymphoma kinase (*ALK*) genomic tumor aberrations. Approximately 30% of patients with newly diagnosed, advanced NSCLC highly express PD-L1 with a TPS ≥50% [Reck, M., et al 2016].

KEYNOTE-024, a Phase 3, randomized, open-label study, compared pembrolizumab monotherapy to standard first-line (1L) platinum-doublet therapy in 305 previously untreated participants with advanced NSCLC and PD-L1 TPS ≥50%. Results from KEYNOTE-024 indicated a significant PFS benefit for pembrolizumab over platinum-doublet therapy, with a hazard ratio (HR) of 0.50 (95% confidence interval [CI], 0.37-0.68, $p<0.001$). Median PFS was 10.3 months (95% CI, 6.7 months to not reached) with pembrolizumab and 6.0 months (95% CI, 4.2-6.2 months) with platinum-doublet. Additionally, there was an OS benefit, with an HR of 0.60 (95% CI, 0.41-0.89; $p=0.005$); median OS was not reached in either treatment arm. The Kaplan-Meier estimates showed 80.2% (95% CI, 72.9%-85.7%) of participants treated with pembrolizumab and 72.4% (95% CI, 64.5%-78.9%) of participants treated with platinum-doublet therapy were alive at 6 months. Furthermore, the ORR was higher in the pembrolizumab arm than in the platinum-doublet arm (44.8% vs 27.8%, respectively) [Reck, M., et al 2016].

The results from KEYNOTE-024 established pembrolizumab as 1L therapy for patients whose tumors have a TPS ≥50% with no *EGFR* or *ALK* genomic tumor aberrations.

Additionally, KEYNOTE-042, a randomized, open-label, Phase 3 study, compared pembrolizumab monotherapy to SOC platinum-based chemotherapy in previously untreated participants with advanced or metastatic NSCLC whose tumors express PD-L1 TPS ≥1%, confirmed the OS treatment effect of pembrolizumab monotherapy observed in participants with TPS ≥50% NSCLC in KEYNOTE-024, and extended these benefits to a broader population with TPS ≥1% NSCLC. Results from KEYNOTE-042 indicated a significant OS benefit for pembrolizumab over platinum-doublet therapy, with an HR of 0.81 (95% CI, 0.71-0.93, *p*=0.0018). Median OS was 16.7 months (95% CI, 13.9-19.7 months) with pembrolizumab and 12.1 months (95% CI, 11.3-13.3 months) with platinum-doublet [Lopes, G., et al 2018]. The Kaplan-Meier estimates showed 39.3% of participants treated with pembrolizumab and 28.0% of participants treated with platinum-doublet therapy were alive at 6 months (data on file).

While the benefits observed in KEYNOTE-024 (in participants with advanced or metastatic NSCLC whose tumors expressed PD-L1 TPS ≥50 %) and KEYNOTE-042 (in participants with advanced or metastatic NSCLC whose tumors expressed PD-L1 TPS ≥1%) were substantial for these participants with NSCLC, the benefits may be improved further in metastatic NSCLC with PD-L1 TPS ≥1% population, possibly through the combination of pembrolizumab with other therapies such as lenvatinib, a multiple RTKi with a potentially differentiated profile.

Lenvatinib has been studied as both monotherapy and in combination with chemotherapy for the treatment of advanced NSCLC. As a single agent, a randomized Phase 2 study of lenvatinib versus placebo was conducted in 135 participants with locally advanced or metastatic nonsquamous NSCLC who had failed to respond to at least 2 systemic anticancer regimens. Eighty-nine participants received 24 mg of lenvatinib QD and 46 received placebo; all participants received best supportive care in addition to study intervention. Median OS was 38.4 weeks for lenvatinib and 24.1 weeks for placebo, median PFS was 20.9 weeks for lenvatinib and 7.9 weeks for placebo, ORR was 10.1% for lenvatinib and 2.2% for placebo, and the disease control rate (DCR) was 42.7% for lenvatinib and 19.6% for placebo. Grade 3 and/or 4 AEs occurred in 69% of lenvatinib recipients and 50% of placebo recipients. Grade 3 and/or 4 AEs were dyspnea and pneumonia in both treatment groups, as well as hypertension in the lenvatinib group [Havel, L., et al 2014].

Lenvatinib was studied in combination with chemotherapy in a multicenter, open-label, Phase 1, dose-finding study (Study E7080-J081-110). This study, conducted in Japanese participants with chemotherapy-naïve NSCLC, investigated lenvatinib in combination with carboplatin/paclitaxel to establish the maximum tolerated dose (MTD). The MTD of lenvatinib was determined to be 4 mg BID after dose-limiting toxicities (DLTs) were experienced at 6 mg BID (febrile neutropenia, gingival infection [n=2]). A dose-expansion cohort was added and 28 participants were dosed in total: 12 in the dose-finding cohort and 16 participants in the dose-expansion cohort. In total, 22 participants were dosed at 4 mg BID with 6 at 6 mg BID. The ORR was 61% (95% CI: 41%, 79%) and comprised 1 CR and 16 partial responses (PRs). The median PFS was reported as 9 months. Of the 22 participants in the 4-mg BID dose level, the most frequently reported AEs were thrombocytopenia, neutropenia, leukopenia, peripheral sensory neuropathy, arthralgia, and alopecia. The most

frequently reported Grade 3 or 4 AEs were neutropenia, leukopenia, hypertension, thrombocytopenia, and febrile neutropenia. No AEs resulting in death were reported [Nishio, M., et al 2013].

The present study is designed to further evaluate the safety and efficacy of combination therapy of lenvatinib and pembrolizumab in adult participants with treatment-naïve, metastatic NSCLC with a TPS ≥1%. The outcomes for participants evaluated in this study would be further improved if the safety profile of a pembrolizumab plus lenvatinib combination remains acceptable and is shown to improve outcomes compared with pembrolizumab monotherapy; thus, this study could support the regulatory approval of this combination in this patient population.

2.2 Background

Pembrolizumab is a potent humanized IgG4 mAb with high specificity of binding to the PD-1 receptor, thus inhibiting its interaction with PD-L1 and PD-L2. Based on preclinical in vitro data, pembrolizumab has high affinity and potent receptor blocking activity for PD-1. Pembrolizumab has an acceptable preclinical safety profile and is in clinical development as an IV immunotherapy for advanced malignancies. Keytruda® (pembrolizumab) is indicated for the treatment of patients across a number of indications. For more details on specific indications, refer to the pembrolizumab Investigator's Brochure (IB).

Lenvatinib (also known as E7080 or MK-7902) inhibits the kinase activities of vascular endothelial growth factor (VEGF) receptors VEGFR1 (FLT1), VEGFR2 (KDR), and VEGFR3 (FLT4). Lenvatinib inhibits other kinases that have been implicated in pathogenic angiogenesis, tumor growth, and cancer progression in addition to their normal cellular functions, including fibroblast growth factor (FGF) receptors FGFR1, 2, 3, and 4; platelet-derived growth factor receptor alpha (PDGFRα), KIT, and RET. Lenvatinib also exhibited antiproliferative activity in cell lines dependent on activated FGFR signaling with a concurrent inhibition of FGF-receptor substrate 2α phosphorylation. Once daily (QD) dosing of lenvatinib combined with pembrolizumab is currently being developed for the treatment of metastatic NSCLC. Refer to the respective IBs/approved labeling for detailed background information on pembrolizumab and lenvatinib.

2.2.1 Pharmaceutical and Therapeutic Background

2.2.1.1 Lenvatinib

Angiogenesis, the formation of new blood vessels from a pre-existing vascular network, is essential for tumor growth and metastasis. VEGF and its family of receptors (VEGFRs 1-3) play a major role in tumor angiogenesis [Ferrara, N., et al 2003] [Ellis, L. M. 2008] [Tammela, T. 2010]. Accumulated evidence suggests that FGF and its receptor tyrosine kinase, FGFR, also play important roles in tumor angiogenesis [Cross, M. J. 2001] [Lieu, C., et al 2011] [Limaverde-Sousa, G., et al 2014].

Lenvatinib is a potent multiple RTKi that selectively inhibits VEGF receptors, VEGFR1 (FLT1), VEGFR2 (KDR), VEGFR3 (FLT4), FGFR1-4, PDGFRα, c-kit, and RET. Among

known kinase inhibitors in clinical use, lenvatinib is one of the only inhibitors currently labeled with a mechanism of action as an inhibitor of not only VEGFRs but also FGFRs, both of which are currently believed to be very important for tumor angiogenesis.

Lenvatinib inhibited cell free kinase activities for VEGFR1-3 and FGFR1-3 with Ki values around 1 nmol/L, and 8-22 nmol/L, respectively. In cell-based assays, lenvatinib inhibited VEGF-derived and FGF-derived tube formation of HUVEC with IC_{50} values of 2.1 and 7.3 nmol/L, respectively. Analysis of the signal transduction molecules revealed that lenvatinib inhibited both the MAPK pathway and the mTOR-S6K-S6 pathway in HUVECs triggered by activated VEGFR and FGFR. Furthermore, lenvatinib (10, 30 mg/kg) significantly inhibited both VEGF- and FGF-driven angiogenesis in a murine in vivo model [Yamamoto, Y., et al 2014]. In vivo, lenvatinib exhibited antitumor activity against various human tumor xenografts in athymic mice including 5 types of thyroid carcinomas (differentiated [papillary and follicular], anaplastic, squamous, and medullary thyroid carcinomas), RCC, HCC, melanoma, gastric cancer, NSCLC, ovarian cancer, Ewing's sarcoma, and osteosarcoma. In addition, the antitumor activity of lenvatinib in combination with other anticancer agents in several xenograft models was greater than that of lenvatinib or the other agents alone.

In summary, lenvatinib inhibited VEGF-driven VEGFR2 phosphorylation and suppressed proliferation and tube formation in human umbilical vein endothelial cell (HUVEC) models. Antitumor activity of lenvatinib in vivo has been shown in numerous xenograft animals. These results suggest that lenvatinib may be a novel anticancer therapy through inhibition of angiogenesis and may be useful as either monotherapy or in combination with other anticancer drugs.

2.2.1.2 Pembrolizumab

Keytruda® (pembrolizumab) is indicated for the treatment of patients across a number of indications. For more details on specific indications, refer to the pembrolizumab IB.

2.2.1.3 Pembrolizumab With Lenvatinib

The importance of intact immune surveillance function in controlling outgrowth of neoplastic transformations has been known for decades [Disis, M. L. 2010]. Accumulating evidence shows a correlation between tumor-infiltrating lymphocytes in cancer tissue and favorable prognosis in various malignancies. In particular, the presence of CD8+ T-cells and the ratio of CD8+ effector T-cells/FoxP3+ regulatory T-cells (T-regs) correlates with improved prognosis and long-term survival in solid malignancies, such as ovarian, colorectal, and pancreatic cancer; hepatocellular carcinoma; malignant melanoma; and renal cell carcinoma. Tumor-infiltrating lymphocytes can be expanded ex vivo and reinfused, inducing durable objective tumor responses in cancers such as melanoma [Dudley, M. E., et al 2005] [Hunder, N. N., et al 2008]. In preclinical models, lenvatinib decreased the tumor-associated macrophage (TAM) population, which is known as an immune-regulator in the tumor microenvironment. The decrease in TAM population was accompanied by increases in activated cytotoxic T-cell populations through stimulation of interferon-gamma signaling, resulting in increased immune activation [Kimura, T., et al 2018]. The immune-modulating effect of lenvatinib may result in a potent combination effect with PD-1/L1 signal inhibitors.

The effect of combining lenvatinib with an antihuman PD-1 humanized mAb was investigated in 4 murine tumor isograft models, which showed significant tumor growth inhibition compared with control. In the RAG murine tumor isograft tumor model, survival in the group treated with the combination was significantly longer than that of the respective monotherapy groups. In the CT26 murine tumor isograft model, treatment with the combination significantly increased the population of activated cytotoxic T-cells compared with that of the respective monotherapy groups [Kato, Y., et al 2019]. All treatments were well tolerated and severe body weight loss was not observed.

2.2.2 Preclinical and Clinical Studies

2.2.2.1 Completed Studies With Pembrolizumab and Lenvatinib

Refer to the respective IBs for further information on completed studies with lenvatinib and pembrolizumab.

2.2.3 Ongoing Clinical Studies of Pembrolizumab and Lenvatinib

Pembrolizumab is under evaluation in patients with NSCLC as monotherapy and in combination with chemotherapy, immunotherapy, and targeted therapies. Lenvatinib is being studied in patients with different types of solid tumors, including NSCLC, and in combination with other therapies, including PD-1-targeted therapies. A full list of ongoing studies can be found in the respective IBs of pembrolizumab [IB Edition 16 2018] and lenvatinib [IB Edition 15-Eisai 2018]. Details of the ongoing study 111/KEYNOTE-146 are outlined below.

Study 111/KEYNOTE-146

Study 111/KEYNOTE-146, is a multicenter, open-label, Phase 1b/2 clinical study to evaluate the efficacy and safety of lenvatinib in combination with pembrolizumab. The primary objective of the Phase 1b portion of the study is to determine the MTD in participants with unresectable solid tumors (endometrial cancer, melanoma, NSCLC, RCC, squamous cell carcinoma of the head and neck, and urothelial cancer) who have progressed after treatment with approved therapies or for which there are no available standard effective therapies. The primary endpoint of the initial part of Phase 2 is ORR after 24 weeks of treatment, with select secondary endpoints, including OR, DCR, PFS, and duration of response (DOR).

As of the data cutoff of 01-DEC-2017, 20 participants with NSCLC were enrolled (data on file). Of the enrolled participants, 9 (45%) were PD-L1(+) (TPS ≥1%), 5 (25%) were PD-L1(−), and 6 (30%) were not tested; 3 (15%) were treatment-naïve; and 6 (30%), 9 (45%), and 2 (10%) had 1, 2, and ≥3 prior lines of systemic therapy, respectively. The primary endpoint of ORR at Week 24 was 30.0% (95% CI, 11.9%–54.3%). Grade 3 and 4 treatment-related AEs occurred in 11 participants (55%) and 1 participant (5%), respectively (ie, increased aspartate aminotransferase [AST]). There was 1 fatal treatment-related AE (exsanguination that was deemed possibly related to study intervention). The most common Grade 3 treatment-related AEs were hypertension (30% [6 of 20 subjects]) and fatigue, diarrhea, and hyponatremia (15% [3 of 20 subjects]). These findings indicate that the

combination of lenvatinib with pembrolizumab shows promising clinical activity with a manageable safety profile in previously treated patients with metastatic NSCLC who were not preselected for PD-L1 status.

2.3 Benefit/Risk Assessment

NOTE: Data from an interim safety and futility eDMC for LEAP-007 (data cutoff: 19-MAY-2021) indicated that the study met the prespecified nonbinding futility criteria for OS for the combination of lenvatinib plus pembrolizumab compared with placebo plus pembrolizumab and the benefit-risk assessment was not considered positive to continue the study as designed. Therefore, the benefit-risk assessment of pembrolizumab plus lenvatinib is no longer applicable and has been removed from this section.

It cannot be guaranteed that participants in clinical studies will directly benefit from treatment during participation, as clinical studies are designed to provide information about the safety and effectiveness of an investigational medicine.

Additional details regarding specific benefits and risks for participants participating in this clinical study may be found in the accompanying IB and informed consent documents.

3 HYPOTHESES, OBJECTIVES, AND ENDPOINTS

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. All participants remaining on study will continue on open-label pembrolizumab monotherapy. N**o further analyses of efficacy endpoints will be performed. Exploratory objectives may not be pursued.**

Participants were screened for up to 28 days, then randomized, and continued study treatment until discontinuation (defined in Section 7.1) OR stop/unblinding of study intervention due to the futility analysis criterion for OS being met.

In adult participants with treatment-naïve, metastatic NSCLC expressing PD-L1 (TPS ≥1%):

Objectives	Endpoints
Primary	
• Objective: To compare progression-free survival (PFS) as assessed by blinded independent central review (BICR) according to Response Evaluation Criteria in Solid Tumors, version 1.1 (RECIST 1.1), adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo – Hypothesis (H1): The combination of pembrolizumab + lenvatinib has superior PFS per RECIST 1.1, as defined in Section 4.2.1.1.1, based on BICR	• PFS, defined as the time from randomization to the first documented progressive disease (PD) or death due to any cause, whichever occurs first*
• Objective: To compare overall survival (OS) for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo – Hypothesis (H2): The combination of pembrolizumab + lenvatinib has superior OS	• OS, defined as the time from randomization to death due to any cause*
Secondary	
• Objective: To compare objective response rate (ORR) as assessed by BICR according to RECIST 1.1, as defined in Section 4.2.1.1.1, for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo – Hypothesis (H3): The combination of pembrolizumab + lenvatinib has superior ORR per RECIST 1.1, as defined in Section 4.2.1.1.1, based on BICR	• Objective response (OR), defined as a confirmed complete response (CR) or partial response (PR)

Objectives	Endpoints
• Objective: To evaluate the safety and tolerability for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	• Adverse events (AEs) and study intervention discontinuation due to AEs
• Objective: To compare the mean change from baseline in the global health status/quality of life (QoL), cough, chest pain, dyspnea, and physical functioning for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	• Change from baseline for the following patient-reported outcomes (PROs) scales/items: global health status/QoL (EORTC QLQ-C30 items 29 and 30), cough (EORTC QLQ-LC13 item 31), chest pain (EORTC QLQ-LC13 item 40), dyspnea (EORTC QLQ-C30 item 8), and physical functioning (EORTC QLQ-C30 items 1-5)
• Objective: To compare the time to true deterioration (TTD) in global health status/QoL, cough, chest pain, dyspnea, and physical functioning for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	• TTD, defined as the time from baseline to the first onset of a ≥10-point deterioration from baseline with confirmation by the subsequent visit of a ≥10-point deterioration from baseline in global health status/QoL (EORTC QLQ-C30 items 29 and 30), cough (EORTC QLQ-LC13 item 31), chest pain (EORTC QLQ-LC13 item 40), dyspnea (EORTC QLQ-C30 item 8), and physical functioning (EORTC QLQ-C30 items 1-5) • TTD in the composite endpoint (combination of cough [QLQ-LC13 item 31], chest pain [QLQ-LC13 item 40], or dyspnea [QLQ-C30 item 8]) defined as the time to first onset of a ≥10-point deterioration from baseline in any one of 3 scale items with confirmation by the subsequent visit of a ≥10-point deterioration from baseline in the same scale as the first onset

Objectives	Endpoints
Tertiary/Exploratory	
• Objective: To compare the PFS and ORR per adjusted RECIST 1.1 for Immune-based Therapeutics (iRECIST) as assessed by the investigator for the combinations of pembrolizumab + lenvatinib versus pembrolizumab + matching placebo	• PFS, defined as the time from randomization to the first documented PD or death due to any cause, whichever occurs first • OR, defined as a confirmed CR or PR
• Objective: To evaluate DOR per RECIST 1.1, as defined in Section 4.2.1.1.1, for the combinations of pembrolizumab + lenvatinib and pembrolizumab + matching placebo	• DOR, defined as the time from the earliest date of qualifying response until earliest date of PD or death from any cause, whichever comes first
• Objective: To evaluate and compare participant's health status as assessed by the EuroQoL 5 Dimensions 5-Level (EQ-5D-5L) questionnaire to generate utility scores for use in economic models	• Health utilities assessed using the EuroQoL EQ-5D-5L
• Objective: To evaluate the pharmacokinetics (PK) of lenvatinib when coadministered with pembrolizumab	• Plasma concentration of lenvatinib versus time
• To identify molecular (genomic, metabolic, and/or proteomic) biomarkers that may be indicative of clinical response/resistance, safety, and/or the mechanism of action of pembrolizumab and lenvatinib in all participants	• Molecular (genomic, metabolic, and/or proteomic) determinants of response or resistance to treatments, using blood and/or tumor tissue

*This study will be considered to have met its success criteria if the combination of pembrolizumab + lenvatinib is superior to pembrolizumab + matching placebo in PFS or OS.

4 STUDY DESIGN

4.1 Overall Design

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. Participants who remain on treatment will receive open-label pembrolizumab monotherapy, as per protocol. The Second-course Treatment Phase is removed from the study. This section has been updated accordingly.

The original study design was a randomized, double-blind, active-control with placebo, parallel-group, multisite study of IV pembrolizumab and oral lenvatinib in participants with Stage IV NSCLC and PD-L1 TPS ≥1% who have not had previously received systemic therapy for metastatic disease, whose tumors express PD-L1 with a TPS ≥1%, and in whom *EGFR*-, *ALK*-, or ROS proto-oncogene 1 (*ROS1)*-directed therapy is not indicated. Participants were randomized 1:1 to Arm A (pembrolizumab + lenvatinib) or Arm B (pembrolizumab + matching placebo). Randomization was stratified by site geographic region (East Asia vs non-East Asia), ECOG PS (0 vs 1), and TPS (1%-49% vs ≥50%).The original study design is depicted in Figure 1 (Initial Treatment Phase).

As of Amendment 007-06, lenvatinib and matching placebo have been removed from the treatment arms (Figure 1). All participants remaining on treatment in the study will receive open-label pembrolizumab monotherapy. Treatment with pembrolizumab will continue up to Cycle 35 or until reaching a discontinuation criterion (defined in Section 7.1). Imaging will continue until progression by local investigator has been determined.

Participants will be evaluated with radiographic imaging to assess response to treatment every 9 weeks from randomization through 54 weeks, then every 12 weeks until PD, initiation of a new anticancer regimen, or completion of 35 cycles. All imaging obtained on study will be submitted to the imaging vendor for BICR, which will assess the images using RECIST 1.1 (see Section 8.2.1.5) for determination of OR, PFS, and DOR. After amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed, similarly, verification of PD is no longer needed before treatment discontinuation. Treatment-based decisions may use site-assessed iRECIST, as described in Section 8.2.1.6, which allows for participants with initial site-assessed PD to continue treatment until PD is confirmed by the site 4 to 8 weeks later.

Adverse event monitoring will be ongoing throughout the study, and AEs will be graded by severity according to the guidelines outlined in the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), version 4.0.

The study will be conducted in conformance with Good Clinical Practice (GCP).

Extension Study in China

NOTE: As of Amendment 007-06, enrollment in the Extension Study in China has been stopped; the planned enrollment goal of 120 participants was not reached.

Lenvatinib and matching placebo are removed from the extension study and it will proceed identical to the global study as per this protocol amendment (eg, study endpoints, primary and secondary objectives, study procedures), with the exception of an additional supplemental statistical analysis plan (sSAP), which will provide details of analyses specific to Chinese participants.

Specific procedures to be performed during the study, as well as their prescribed times and associated visit windows, are outlined in the SoA in Section 1.3. Details of each procedure are provided in Section 8.

4.2 Scientific Rationale for Study Design

This Phase 3 study is being conducted to evaluate the efficacy and safety of pembrolizumab in metastatic NSCLC when administered in combination with lenvatinib. Lenvatinib (20 mg QD) will be evaluated in combination with the standard dose of pembrolizumab (200 mg Q3W).

4.2.1 Rationale for Endpoints

4.2.1.1 Efficacy Endpoints

This study will use PFS based on RECIST 1.1 criteria, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, as assessed by BICR as a primary endpoint and OR as a secondary efficacy endpoint. PFS and OR are acceptable measures of clinical benefit for a late-stage study that demonstrates superiority of a new antineoplastic therapy, especially if the magnitude of the effect is large and the therapy has an acceptable risk/benefit profile. The use of BICR and RECIST 1.1 to assess PFS and OR are typically considered acceptable by regulatory authorities. Images will be read by a central imaging vendor blinded to treatment assignment to minimize bias in the response assessments. In addition, the final determination of radiologic progression will be based on the central assessment of progression, rather than a local site investigator/radiology assessment. Real time determination of radiologic progression as determined by central review will be communicated to the site.

The primary efficacy endpoint OS has been recognized as the gold standard for the demonstration of superiority of a new antineoplastic therapy in randomized clinical studies.

For additional details about assessing efficacy endpoints using RECIST 1.1 and iRECIST, see Appendix 6.

4.2.1.1.1 RECIST 1.1

RECIST 1.1 will be used by the BICR when assessing images for efficacy measures and by the local site when determining eligibility (Section 8.2.1.5). Although traditional RECIST 1.1 references a maximum of 5 target lesions in total and 2 per organ, this protocol has implemented an adjustment to RECIST 1.1 to follow a maximum of 10 target lesions in total and 5 per organ.

4.2.1.1.2 Adjusted RECIST 1.1 for Immune-based Therapeutics iRECIST

RECIST 1.1 will be adapted to account for the unique tumor response characteristics seen following treatment with pembrolizumab (Section 8.2.1.6). Immunotherapeutic agents such as pembrolizumab may produce antitumor effects by potentiating endogenous cancer-specific immune responses. The response patterns seen with such an approach may extend beyond the typical time course of responses seen with cytotoxic agents, and patients treated with pembrolizumab may manifest a clinical response after an initial increase in tumor burden or even the appearance of new lesions. Thus, standard RECIST 1.1 may not fully capture the treatment benefits from immunotherapeutic agents such as pembrolizumab. Based on an analysis of participants with melanoma enrolled in KEYNOTE-001 (KN001), 7% of evaluable participants experienced delayed or early tumor pseudoprogression. Of note, participants who had PD by RECIST 1.1 but not by the immune-related response criteria [Wolchok, J. D., et al 2009] had longer OS than participants with PD by both criteria [Hodi, F. S., et al 2014]. Additionally, the data suggest that RECIST 1.1 may underestimate the benefit of pembrolizumab in approximately 15% of participants. These findings support the need to apply an adjustment to RECIST 1.1 that takes into account the unique patterns of atypical responses in immunotherapy and enables treatment beyond initial radiographic progression, if the participant is clinically stable.

iRECIST assessment has been developed and published by the RECIST Working Group, with input from leading experts from industry and academia, along with participation from the United States Food and Drug Administration (FDA) and the European Medicines Agency (EMA) [Seymour, L., et al 2017]. The unidimensional measurement of target lesions, qualitative assessment of nontarget lesions, and response categories are identical to RECIST 1.1, until progression is seen by RECIST 1.1. However, if a participant is clinically stable, additional imaging may be performed to confirm radiographic progression. iRECIST will be used by investigators to assess tumor response and progression and make treatment decisions as well as for exploratory efficacy analyses where specified.

4.2.1.2 Safety Endpoints

Safety parameters commonly used for evaluating investigational systemic anticancer treatments are included as safety endpoints including, but not limited to, the incidence of, causality, and outcome of AEs/serious AEs (SAEs); and changes in vital signs and laboratory values. Adverse events will be assessed as defined by the NCI Common Terminology Criteria for Adverse Events (CTCAE), Version 4.0.

4.2.1.3 Patient-reported Outcomes

NOTE: As of Amendment 007-06, ePRO data are no longer being collected.

Symptomatic improvement is considered a clinical benefit and accepted by health authorities. As part of the analyses for this study, participants will provide information regarding their health-related quality of life (HRQoL) via the following assessment tools: European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30, EORTC QLQ-LC13, and EuroQoL EQ-5D-5L questionnaires. The ePROs are not pure efficacy or safety endpoints because they are affected by both PD and treatment tolerability.

4.2.1.3.1 EORTC QLQ-C30

EORTC QLQ-C30 is the most widely used cancer-specific, HRQoL instrument, which contains 30 items and measures 5 functional dimensions (physical, role, emotional, cognitive, and social), 3 symptom items (fatigue, nausea/vomiting, and pain), 6 single items (dyspnea, sleep disturbance, appetite loss, constipation, diarrhea, and financial impact), and a global health and QoL scale [Aaronson, N. K., et al 1993]. The EORTC QLQ-C30 is a psychometrically and clinically validated instrument appropriate for assessing QoL in oncology studies [Aaronson, N. K., et al 1993].

4.2.1.3.2 EORTC QLQ-LC13

The EORTC Quality of Life Questionnaire and Lung Cancer Module 13 (QLQ-LC13), a supplemental lung cancer-specific module used in combination with QLQ-C30, comprises multi-item and single-item measures of lung cancer-associated symptoms (cough, hemoptysis, dyspnea, and site-specific pain) and treatment-related symptoms (sore mouth, dysphagia, peripheral neuropathy, and alopecia) [Bergman, B., et al 1994]. It is scored on a 4-point scale (1 = not at all, 2 = a little, 3 = quite a bit, 4 = very much) and has been translated and validated into more than 60 languages.

The EORTC QLQ-C30 and QLQ-LC13 are the most frequently used PRO measures in lung cancer clinical studies. The reliability, validity, and practicality of these instruments have been reported [Bergman, B., et al 1994] [Aaronson, N. K., et al 1993].

4.2.1.3.3 EuroQoL EQ-5D-5L

The EQ-5D-5L is a standardized instrument for use as a measure of health outcome and will provide data to develop health utilities for use in health economic analyses [Rabin, R. 2001]. The 5 health state dimensions in the EQ-5D-5L include the following: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension is rated on a 5 point scale from 1 (no problem) to 5 (unable to/extreme problems). The EQ-5D-5L also includes a graded (0 to 100) vertical visual analog scale on which the participant rates his or her general state of health at the time of the assessment. This instrument has been used extensively in cancer studies and published results from these studies support its validity and reliability [Pickard, A. S., et al 2007].

4.2.1.4 Pharmacokinetic Endpoints

Prior to Amendment 05, blood samples were obtained to measure the PK of pembrolizumab and lenvatinib. The Sponsor has determined sufficient pembrolizumab/lenvatinib PK data in NSCLC participants have been collected. As such, no further PK samples will be collected.

Standard PK parameters of clearance (CL) and volume of distribution (V) at steady state are planned to be calculated for lenvatinib when coadministered with pembrolizumab, using the accepted mixed-effects modeling approach. PK data from this study may be combined with data from other studies and analyzed using standard population PK techniques to further characterize basic PK parameters, evaluate the effect of extrinsic and intrinsic factors in support of the proposed dosing regimen, and evaluate safety in the proposed participant population.

4.2.1.5 Planned Exploratory Biomarker Research

Cancer immunotherapies represent an important and novel class of antitumor agents. However, the mechanism of action of these exciting new therapies, including novel combinations with antiangiogenesis therapy, is not completely understood and much remains to be learned regarding how best to leverage these new drugs in treating patients. Thus, to aid future patients, it is important to investigate the determinants of response or resistance to cancer immunotherapy and other treatments administered, as well as determinants of AEs in the course of our clinical studies. These efforts may identify novel predictive/PD biomarkers and generate information that may better guide single agent and combination therapy with immuno-oncology drugs. To identify novel biomarkers, biospecimens (ie, blood components, tumor material) will be collected to support analyses of cellular components (eg, protein, deoxyribonucleic acid [DNA], ribonucleic acid [RNA], metabolites) and other circulating molecules. Investigations may include but are not limited to:

Germline (blood) genetic analyses (eg, SNP analyses, whole exome sequencing, whole genome sequencing)

This research may evaluate whether genetic variation within a clinical study population correlates with response to the treatment(s) under evaluation. If genetic variation is found to predict efficacy or AEs, the data might inform optimal use of therapies in the patient population. Furthermore, it is important to evaluate germline DNA variation across the genome in order to interpret tumor-specific DNA mutations.

Genetic (DNA) analyses from tumor

The application of new technologies, such as next generation sequencing, has provided scientists the opportunity to identify tumor-specific DNA changes (ie, mutations, methylation status, microsatellite instability) contributing towards the development/progression of cancer and/or driving response to therapy. Key molecular changes of interest to immuno-oncology drug development include the mutational burden of tumors and the clonality of T-cells in the tumor microenvironment. Increased mutational burden (sometimes referred to as a 'hyper-mutated' state) may generate neo-antigen presentation in the tumor microenvironment. To

conduct this type of research, it is important to identify tumor-specific mutations that occur across all genes in the tumor genome. Evaluation of molecular targets and signaling pathways including angiogenesis- or and growth factor related signaling pathways related to pembrolizumab and lenvatinib may also be explored. Thus, genome-wide approaches may be used for this effort. Note that in order to understand tumor-specific mutations, it is necessary to compare the tumor genome with the germline genome. Circulating tumor DNA and/or RNA may also be evaluated from blood samples.

Tumor and blood RNA analyses

Both genome-wide and targeted messenger RNA (mRNA) expression profiling and sequencing in tumor tissue and in blood may be performed to define gene signatures that correlate to clinical response to treatment with pembrolizumab or other immunotherapies. Pembrolizumab induces a response in tumors that likely reflects an inflamed/immune phenotype. Specific immune-related gene sets (ie, those capturing interferon-gamma transcriptional pathways) may be evaluated and new signatures may be identified. Individual genes related to the immune system and growth factor signaling pathways (eg, VEGF and FGF) may also be evaluated. MicroRNA profiling may also be pursued as well as exosomal profiling.

Proteomics and immunohistochemistry (IHC) using blood or tumor

Tumor and blood samples from this study may undergo proteomic analyses (eg, PD-L1 IHC). PD-L1 protein level in tumor sections, assessed by IHC, has been shown to correlate with response to pembrolizumab in patients with NSCLC, and an in vitro diagnostic (IVD) device has been developed for use with pembrolizumab in NSCLC. Preliminary data indicates that this association may also be true in additional cancer types (ie, triple negative breast cancer, head and neck, and gastric). Additional tumor or blood-derived proteins may also correlate with response to pembrolizumab and lenvatinib combination therapy. Therefore, tumor tissue may be subjected to proteomic analyses using a variety of platforms that could include but are not limited to immunoassays and liquid chromatography/mass spectrometry. This approach could identify novel protein biomarkers that could aid in patient selection for pembrolizumab (MK-3475) and lenvatinib combination therapy.

Other blood-derived biomarkers

In addition to expression on the tumor tissue, PD-L1, circulating cytokines and angiogenic factors, and other tumor derived proteins can be shed from tumor and released into the blood. Assays such as enzyme-linked immunoassay (ELISA) measure such proteins in serum. Correlation of expression with response to pembrolizumab and lenvatinib combination therapy may identify new approaches for predictive biomarkers in blood, representing a major advance from today's reliance on assessing tumor biomarkers. This research would serve to develop such assays for future clinical use.

Other molecular changes of interest include the subtype of T-cells in the tumor microenvironment. The T-cell repertoire from tumor tissue and blood components may be evaluated.

4.2.2 Rationale for the Use of Comparator/Placebo

Based on the results of KEYNOTE-024, pembrolizumab monotherapy has become the SOC for treating Stage IV NSCLC in patients with PD-L1 TPS ≥50% with no *EGFR* or *ALK* genomic tumor aberrations. The study resulted in a significant PFS benefit for pembrolizumab over the SOC at the time (platinum-doublet therapy), with an HR of 0.50 (95% CI, 0.37-0.68; $p<0.001$; median PFS, 10.3 months [95% CI, 6.7 to not reached] with pembrolizumab and 6.0 months [95% CI, 4.2-6.2 months] with platinum-doublet), as well as an OS benefit with an HR of 0.60 (95% CI 0.41-0.89; $p=0.005$; median OS was not reached in either treatment arm) [Reck, M., et al 2016]. The results from KEYNOTE-024 established pembrolizumab as 1L therapy for patients whose tumors have a TPS ≥50% and in whom *EGFR*- or *ALK*-directed therapies are not indicated, and the regimen has received regulatory approval for this use by the FDA and EMA. Additionally, based on results from KEYNOTE-042, pembrolizumab monotherapy is considered one of the treatment options in treating Stage IV NSCLC in patients with PD-L1 TPS ≥1% with no *EGFR* or *ALK* genomic tumor aberrations. The study resulted in a significant OS benefit for pembrolizumab over the SOC at the time (platinum-doublet therapy), with an HR of 0.81 (95% CI, 0.71-0.93; $p=0.0018$). Median OS was 16.7 months (95% CI, 13.9-19.7 months) with pembrolizumab and 12.1 months (95% CI, 11.3-13.3 months) with platinum-doublet therapy [Lopes, G., et al 2018].

The use of a lenvatinib matching placebo in combination with pembrolizumab will ensure the objectivity of the local investigators' treatment decision and AE causality assessments, while still providing participants the SOC treatment.

4.3 Justification for Dose

4.3.1 Maximum Dose/Exposure for This Study

NOTE: As of Amendment 007-06, lenvatinib, matching placebo, and Second-course treatment have been removed from the study. This section has been updated accordingly.

The maximum dose/exposure of pembrolizumab allowed in this study is 200 mg Q3W up to 35 cycles (approximately 2 years) for the Initial Treatment Phase.

4.3.2 Rationale for Dose Interval and Study Design

4.3.2.1 Lenvatinib Dosing

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. This section is no longer applicable, but has been left unchanged for reference.

The dosing regimen of lenvatinib was selected based on the results of the Phase 1b portion of Phase 1b/2 Study 111/KEYNOTE-146, of which the primary endpoint was to determine the MTD and RP2D for lenvatinib in combination with pembrolizumab 200 mg Q3W. Thirteen participants (lenvatinib 24 mg/day + pembrolizumab 200 mg IV Q3W: n=3; lenvatinib 20 mg/day + pembrolizumab 200 mg: n=10) were enrolled into the Phase 1b portion of the

study. Eight of the participants had RCC, 2 had NSCLC, 2 had endometrial cancer, and 1 had melanoma. There were 2 DLTs at the dose of lenvatinib 24 mg/day + pembrolizumab 200 mg IV Q3W (1 participant had Grade 3 arthralgia and another had Grade 3 fatigue); hence, this was defined as the toxic dose. No DLTs were reported in the next 10 participants (expansion part), all of whom received the lenvatinib 20 mg/day + pembrolizumab 200 mg Q3W dose.

Based on review of all of the clinical data from these 13 participants, the MTD and RP2D were determined to be 20 mg of lenvatinib daily in combination with a fixed dose of 200 mg of pembrolizumab given Q3W. Based on the promising antitumor efficacy and tolerable safety profile seen in both the endometrial carcinoma and RCC expansion cohorts from Study 111/KEYNOTE-146 [Makker, V., et al 2018], 2 Phase 3 studies have been initiated for both of these tumor types, Study E7080-G000-309/KEYNOTE-775 and Study E7080-G000-307/KEYNOTE-581.

4.3.2.2 Pembrolizumab Dosing

The rationale for the use of a fixed dose of pembrolizumab in participants with solid tumors was based on the following:

- Similar efficacy and safety of pembrolizumab when dosed at 2 mg/kg Q3W, 10 mg/kg Q3W, or 10 mg/kg every 2 weeks in participants with melanoma and NSCLC

- Flat exposure-response relationships of pembrolizumab for both efficacy and safety in the dose range of 2 mg/kg Q3W to 10 mg/kg Q3W

- The lack of clinically relevant effect of tumor burden or indication on distribution behavior of pembrolizumab (as assessed by the population pharmacokinetics [PopPK] model)

- Pharmacology data showing full target saturation in both systemic circulation (inferred from PK data) and tumor (inferred from physiologically based PK analysis) at 200 mg Q3W

The choice of pembrolizumab 200 mg Q3W as an appropriate fixed dose was based on simulations performed using the PopPK model of pembrolizumab, which showed the following for a fixed dose of 200 mg Q3W:

- Provide similar control of PK variability as weight-based dosing, with considerable overlap in the distribution of exposures from the 200 mg Q3W fixed dose and 2 mg/kg Q3W dose

- Maintain individual participant exposures in the exposure range associated with maximal efficacy response that was established in participants with melanoma and NSCLC

- Result in individual participant exposures within a range that was well tolerated and safe in participants with melanoma and NSCLC

Clinical data have shown meaningful improvement in benefit-risk, including OS, at 200 mg Q3W across multiple indications. Additionally, a fixed-dose regimen simplified the dosing regimen to be more convenient for physicians and to reduce the potential for dosing errors. A fixed-dosing scheme also reduced complexity in the logistical chain at treatment facilities and reduced wastage.

4.4 Beginning and End of Study Definition

The overall study begins when the first participant (or their legally acceptable representative) provides documented informed consent. The overall study ends when the last participant completes the last study-related contact, withdraws consent, or is lost to follow-up (ie, the participant is unable to be contacted by the investigator).

4.4.1 Clinical Criteria for Early Study Termination

The clinical study may be terminated early if the extent (incidence and/or severity) of emerging effects/clinical endpoints is such that the risk/benefit ratio to the study population as a whole is unacceptable. In addition, further recruitment in the study or at (a) particular study site(s) may be stopped due to insufficient compliance with the protocol, Good Clinical Practice (GCP), and/or other applicable regulatory requirements, procedure-related problems or an unacceptably high number of discontinuations or withdrawals due to administrative reasons.

5 STUDY POPULATION

Male and female participants with Stage IV NSCLC, who express PD-L1 (TPS ≥1%), have received no systemic anticancer therapy for their Stage IV NSCLC, and are at least 18 years of age will be enrolled in this study.

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

5.1 Inclusion Criteria

To be eligible for inclusion in this study, the participant must:

Type of Participant and Disease Characteristics

1. Have a histologically or cytologically confirmed diagnosis of NSCLC.

   Note: Mixed tumors will be categorized by the predominant cell type; if small-cell elements are present, the participant is ineligible.

2. Have Stage IV NSCLC (American Joint Committee on Cancer [AJCC], version 8).

3. Have confirmation that EGFR-, *ALK*-, or *ROS1*-directed therapy is not indicated as primary therapy (documentation of the absence of tumor-activating *EGFR* mutations [eg, DEL19 or L858R], AND absence of *ALK* and *ROS1* gene rearrangements OR presence of a Kirsten rat sarcoma [K-*ras*] mutation).

   Note: If participant's tumor is known to have a predominantly squamous histology, molecular testing for *EGFR* mutation and *ALK* and *ROS1* translocations will not be required, as this is not part of current diagnostic guidelines.

4. Have measurable disease based on RECIST 1.1, as determined by the local site.

   Note: Lesions that appear measurable but are situated in a previously irradiated area can be considered measurable (eligible for selection as target lesions) if they have shown documented growth since the completion of radiation.

5. Tumor tissue that demonstrates PD-L1 expression in ≥1% of tumor cells (TPS ≥1%) as assessed by IHC 22C3 pharmDx at a central laboratory.

   Note: Assessment of PD-L1 expression must be made from provided archival tumor tissue sample or newly obtained core or excisional biopsy of a tumor lesion not previously irradiated. (A fine-needle aspirate, frozen sample, plastic-embedded sample, cell block, clot, bone, bone marrow, cytologic specimen, or decalcified or formalin-fixed sample that was frozen at any point will not be acceptable for analysis). Formalin-fixed, paraffin-embedded tissue blocks are preferred to slides. Newly obtained biopsies are preferred to archived tissue.

Demographics

6. Be ≥18 years of age, inclusive, at the time of signing the ICF.

7. Have a life expectancy of at least 3 months.

8. Have an ECOG performance status of 0 or 1 within 7 days before the first dose of study intervention but before randomization.

Contraceptive use

Contraceptive use by men and women should be consistent with local regulations regarding the methods of contraception for those participating in clinical studies. If the contraception requirements in the local label for any of the study interventions is more stringent than the requirements above, the local label requirements are to be followed.

Male Participants

9. Male participants are eligible to participate if they agree to the following during the intervention period and for at least 7 days after the last dose of lenvatinib/matching placebo:

   - Be abstinent from heterosexual intercourse as their preferred and usual lifestyle (abstinent on a long-term and persistent basis) and agree to remain abstinent

     OR

   - Must agree to use contraception unless confirmed to be azoospermic (vasectomized or secondary to medical cause [Appendix 5]) as detailed below:

     – Agree to use a male condom plus partner use of an additional contraceptive method when having penile-vaginal intercourse with a woman of childbearing potential (WOCBP) who is not currently pregnant.

       Note: Men with a pregnant or breastfeeding partner must agree to remain abstinent from penile-vaginal intercourse or use a male condom during each episode of penile-vaginal penetration.

     – Please note that 7 days after lenvatinib/matching placebo is stopped, if the participant is on pembrolizumab only, no male contraception measures are needed.

Female Participants

10. A female participant is eligible to participate if she is not pregnant or breastfeeding, and at least one of the following conditions applies:

    - Is not a WOCBP.

    OR

    - Is a WOCBP and using a contraceptive method that is highly effective (with a failure rate of <1% per year), with low user dependency, or be abstinent from heterosexual intercourse as their preferred and usual lifestyle (abstinent on a long term and persistent basis), as described in Appendix 5 during the intervention period and for at least 120 days post pembrolizumab or 30 days post lenvatinib/matching placebo, whichever occurs last. The investigator should evaluate the potential for contraceptive method failure (ie, noncompliance, recently initiated) in relationship to the first dose of study intervention.

    - A WOCBP must have a negative highly sensitive pregnancy test (urine or serum as required by local regulations) within 24 hours before the first dose of study intervention.

- If a urine test cannot be confirmed as negative (eg, an ambiguous result), a serum pregnancy test is required. In such cases, the participant must be excluded from participation if the serum pregnancy result is positive.

- Additional requirements for pregnancy testing during and after study intervention are located in Appendix 2.

- The investigator is responsible for review of medical history, menstrual history, and recent sexual activity to decrease the risk for inclusion of a woman with an early undetected pregnancy.

Informed Consent

11. The participant (or legally acceptable representative if applicable) has provided documented informed consent/assent for the study.

Additional Categories

12. Have adequately controlled blood pressure (BP) with or without antihypertensive medications, defined as BP ≤150/90 mm Hg and no change in antihypertensive medications within 1 week prior to randomization.

13. Have adequate organ function as defined in the following table (Table 1). Specimens must be collected within 10 days before the start of study intervention.

Table 1 Adequate Organ Function Laboratory Values

System	Laboratory Value
Hematological	
ANC	≥1500/µL
Platelets	≥100,000/µL
Hemoglobin	≥9.0 g/dL or ≥5.6 mmol/L[a]
Renal	
Creatinine or Measured or calculated[b] CrCl (GFR can also be used in place of creatinine or CrCl)	≤1.5 × ULN or ≥30 mL/min for participant with creatinine levels >1.5 × institutional ULN
Hepatic	
Total bilirubin	≤1.5 × ULN or direct bilirubin ≤ULN for participants with total bilirubin levels >1.5 × ULN
AST (SGOT) and ALT (SGPT)	≤2.5 × ULN (≤5 × ULN for participants with liver metastases)
Coagulation	
INR or PT aPTT/PTT	≤1.5 × ULN unless participant is receiving anticoagulant therapy as long as INR/PT or aPTT/PTT is within therapeutic range of intended use of anticoagulants

Abbreviations: ALT (SGPT) = alanine aminotransferase (serum glutamic pyruvic transaminase); ANC = absolute neutrophil count; aPTT = activated partial thromboplastin time; AST (SGOT) = aspartate aminotransferase (serum glutamic oxaloacetic transaminase); CrCl = creatinine clearance; GFR = glomerular filtration rate; INR = international normalized ratio; PT = prothrombin time; ULN = upper limit of normal.

Note: This table includes eligibility-defining laboratory value requirements for treatment; laboratory value requirements should be adapted according to local regulations and guidelines for the administration of specific chemotherapies.

[a] Criteria must be met without erythropoietin dependency and without packed red blood cell transfusion within last 2 weeks.

[b] CrCl should be calculated per institutional standard.

5.2 Exclusion Criteria

The participant must be excluded from the study if the participant:

Medical Conditions

1. Has known untreated central nervous system metastases and/or carcinomatous meningitis. Participants with previously treated brain metastases may participate provided they are radiologically stable ie, without evidence of progression for at least 4 weeks by repeat imaging (note: repeat imaging should be performed during study screening), clinically stable, and without requirement of steroid treatment for at least 14 days before first dose of study intervention.

2. Has clinically significant hemoptysis (at least 0.5 teaspoon of bright red blood) or tumor bleeding within 2 weeks before the first dose of study intervention.

3. Has radiographic evidence of encasement or invasion of a major blood vessel, or of intratumoral cavitation.

   NOTE:

   - The degree of proximity to major blood vessels should be considered because of the potential risk of severe hemorrhage associated with tumor shrinkage/necrosis following lenvatinib therapy.

   - In the chest, major blood vessels include the main pulmonary artery, the left and right pulmonary arteries, the 4 major pulmonary veins, the superior or inferior vena cava, and the aorta.

4. Has a known history of an additional malignancy, except if the participant has undergone potentially curative therapy with no evidence of that disease recurrence for at least 3 years since initiation of that therapy.

   Note: The time requirement for no evidence of disease for at least 3 years does not apply to the NSCLC for which a participant is enrolled in the study. The time requirement also does not apply to participants who underwent successful definitive resection of basal cell carcinoma of the skin, superficial bladder cancer, squamous cell carcinoma of the skin, in situ cervical cancer, or other in situ cancers.

5. Has an active autoimmune disease that has required systemic treatment in the past 2 years (ie, with the use of disease-modifying agents, corticosteroids, or immunosuppressive drugs). Replacement therapy (eg, thyroxine, insulin, or physiologic corticosteroid replacement therapy for adrenal or pituitary insufficiency, etc.) is not considered a form of systemic treatment and is allowed.

6. Has had an allogeneic tissue/solid organ transplant.

7. Has a known history of human immunodeficiency virus (HIV) infection; HIV testing is not required unless mandated by the local health authority.

8. Has a history of (noninfectious) pneumonitis that required systemic steroids or current pneumonitis/interstitial lung disease.

9. Has a known history of hepatitis B (defined as hepatitis B surface antigen [HBsAg] reactive or hepatitis B virus [HBV]-DNA detected) or known active hepatitis C virus (HCV, defined as HCV-RNA [qualitative] detected or HCV antibody reactive, if HCV-RNA is not the local SOC) infection.

    - Note: No testing for hepatitis B and hepatitis C is required unless mandated by the local health authority.

10. Has a history of a gastrointestinal condition or procedure that in the opinion of the investigator may affect oral study drug absorption.

11. Has significant cardiovascular impairment within 12 months of the first dose of study intervention, such as a history of congestive heart failure greater than New York Heart Association Class II, unstable angina, myocardial infarction, cerebrovascular accident (CVA)/stroke, or cardiac arrhythmia associated with hemodynamic instability.

12. Has not recovered adequately from any toxicity and/or complications from major surgery before starting therapy.

13. Has a known history of active tuberculosis (TB).

14. Has an active infection requiring systemic therapy.

15. Has a known psychiatric or substance abuse disorder that would interfere with the participant's cooperation for the requirements of the study.

16. Previously had a severe hypersensitivity reaction to treatment with an mAb or has a known sensitivity or intolerance to any component of lenvatinib or pembrolizumab.

17. WOCBP who has a positive urine pregnancy test within 24 hours before the first dose of study intervention. If the urine test is positive or cannot be confirmed as negative, a serum pregnancy test will be required.

    Note: If more than 24 hours have elapsed between the screening pregnancy test and the first dose of study intervention, another pregnancy test (urine or serum) must be performed and must be negative for the participant to start receiving study intervention.

Prior/Concomitant Therapy

18. Has received prior systemic chemotherapy or other targeted or biological antineoplastic therapy for their metastatic NSCLC.

    Note: Prior treatment with chemotherapy and/or radiation as part of neoadjuvant/adjuvant or chemoradiation therapy for nonmetastatic NSCLC is allowed as long as therapy was completed at least 6 months before the diagnosis of metastatic NSCLC.

19. Has received prior therapy with an anti-PD-1, anti-PD-L1, or anti-PD-L2 agent or with an agent directed to another stimulatory or co-inhibitory T-cell receptor (eg, CTLA-4, OX 40, CD137) or has received lenvatinib as monotherapy or in combination with anti-PD-1 agents.

20. Has received radiotherapy within 14 days before the first dose of study intervention or received lung radiation therapy of >30 Gy within 6 months before the first dose of study intervention.

    Note: Participants must have recovered from all radiation-related toxicities to Grade 1 or less, not require corticosteroids, and not have had radiation pneumonitis.

21. Has a diagnosis of immunodeficiency or is receiving any form of immunosuppressive therapy within 7 days before the first dose of study intervention.

22. Is receiving systemic steroid therapy (doses exceeding 10 mg daily of prednisone equivalent) within 7 days before the first dose of study intervention.

23. Has received a live or attenuated vaccine within 30 days before the first dose of study intervention. Note: Killed vaccines are allowed.

Prior/Concurrent Clinical Study Experience

Not applicable.

Diagnostic Assessments

24. Participants with proteinuria >1+ on urine dipstick testing/urinalysis will undergo 24-hour urine collection for quantitative assessment of proteinuria. Participants with urine protein ≥1 g/24 hours will be ineligible.

    Prolongation of QTc interval to >480 ms and/or left ventricular ejection fraction (LVEF) below the institutional normal range as determined by a multigated acquisition scan (MUGA) or echocardiogram (ECHO).

Other Exclusions

25. Has a history or current evidence of any condition, therapy, or laboratory abnormality that might confound the results of the study or interfere with the participant's ability to

participate for the full duration of the study, or it is not in the best interest of the participant to participate, in the opinion of the treating investigator.

26. Has had major surgery within 3 weeks prior to first dose of study intervention. Note: Adequate wound healing after major surgery must be assessed clinically, independent of time elapsed for eligibility.

27. Has pre-existing ≥Grade 3 gastrointestinal or nongastrointestinal fistula.

5.3 Lifestyle Considerations

5.3.1 Meals and Dietary Restrictions

Participant should maintain a normal diet unless modifications are required to manage an AE, such as diarrhea, nausea, or vomiting.

5.3.2 Contraception

Based on its mechanism of action, lenvatinib can cause fetal harm when administered to a pregnant woman. Lenvatinib may also result in reduced fertility in females of reproductive potential and may result in damage to male reproductive tissues leading to reduced fertility of unknown duration. In animal reproduction studies, oral administration of lenvatinib during organogenesis at doses below the recommended human dose resulted in embryotoxicity, fetotoxicity, and teratogenicity in rats and rabbits.

Participants should be informed that taking the study intervention may involve unknown risks to the fetus (unborn baby) if pregnancy were to occur during the study. To participate in the study, WOCBP must adhere to the contraception requirement (Appendix 5) from the day of study intervention initiation throughout the study period up to 120 days post pembrolizumab or 30 days post lenvatinib/matching placebo, whichever occurs last. If there is any question that a WOCBP will not reliably comply with the requirements for contraception, that participant should not be entered into the study.

5.3.3 Pregnancy

If a participant inadvertently becomes pregnant while on study intervention, the participant will be immediately discontinued from study intervention. The site will contact the participant at least monthly and document the participant's status until the pregnancy has been completed or terminated. The outcome of the pregnancy will be reported to the Sponsor without delay and within 24 hours if the outcome is an SAE (eg, death, abortion, congenital anomaly, or other disabling or life-threatening complication to the mother or newborn). The study investigator will make every effort to obtain permission to follow the outcome of the pregnancy and report the condition of the fetus or newborn to the Sponsor. If a male participant impregnates a female partner, the study personnel at the site must be informed immediately, and the pregnancy must be reported to the Sponsor and followed as described in Section 8.4.5.

5.3.4 Use in Nursing Women

It is unknown whether lenvatinib or pembrolizumab is excreted in human milk. Since many drugs are excreted in human milk and because of the potential for serious adverse reactions in the nursing infant, participants who are breastfeeding are not eligible for enrollment.

5.4 Screen Failures

Screen failures are defined as participants who consent to participate in the clinical study, but are not subsequently randomized in the study. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any AEs or SAEs meeting reporting requirements as outlined in the data entry guidelines.

5.5 Participant Replacement Strategy

A participant who discontinues from study intervention will not be replaced.

6 STUDY INTERVENTION

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study; the portions of this section referring to these treatments are no longer applicable. This section has been updated accordingly.

Study intervention is defined as any investigational intervention(s), marketed product(s), placebo, or medical device(s) intended to be administered to a study participant according to the study protocol.

Clinical supplies (pembrolizumab) will be packaged to support enrollment as required. Clinical supplies will be affixed with a clinical label in accordance with regulatory requirements.

6.1 Study Intervention(s) Administered

The study interventions to be used in this study are outlined in Table 2.

Table 2 Study Interventions

Arm Name	**Arm Type**	**Intervention Name**	**Type**	**Dose Formulation**	**Unit Dose Strength(s)**	**Dosage Level(s)**	**Route of Administration**	**Regimen/ Treatment Period**	**Use**	**IMP/ NIMP**	**Sourcing**
Arm A and B	Experimental	Pembrolizumab (MK-3475)	Drug	Solution for Infusion	25 mg/mL	200 mg	IV Infusion	Q3W	Standard of care/ Experimental	IMP	Provided centrally by the Sponsor
~~Arm A~~	~~Experimental~~	~~Lenvatinib (E7080/ MK-7902)~~	~~Drug~~	~~Capsule~~	~~10 mg/4 mg~~	~~20 mg~~	~~Oral~~	~~QD~~	~~Experimental~~	~~IMP~~	~~Provided centrally by the Sponsor~~
~~Arm B~~	~~Placebo Comparator~~	~~Placebo~~	~~Drug~~	~~Capsule~~	~~NA~~	~~NA~~	~~Oral~~	~~QD~~	~~Placebo~~	~~IMP~~	~~Provided centrally by the Sponsor~~

Abbreviations: EEA = European Economic Area; IMP = investigational medicinal product; IV = intravenous; NA = not applicable; NIMP = noninvestigational medicinal product; Q3W = every 3 weeks; QD = once daily.
The classification of IMP in this table is based on guidance issued by the European Commission and applies to countries in the EEA. Country differences with respect to the definition/classification of IMP/NIMP may exist. In these circumstances, local legislation is followed.
NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the treatment groups. Participants remaining on treatment will receive open-label pembrolizumab monotherapy as per protocol. The original study interventions and updates are indicated by strikethrough in the table.

All supplies indicated in Table 2 will be provided per the "Sourcing" column depending upon local country operational requirements. If local sourcing, every attempt should be made to source these supplies from a single lot/batch number.

Refer to Section 8.1.8 for details regarding administration of the study intervention.

All placebos were created by the Sponsor to match the active product.

6.1.1 Medical Devices

This section is not applicable.

6.2 Preparation/Handling/Storage/Accountability

6.2.1 Dose Preparation

Details on the preparation and administration of pembrolizumab are provided in the Pharmacy Manual. Lenvatinib is a capsule for oral administration and does not require preparation. Before protocol amendment 007-06, site's staff referred to the Pharmacy Manual for lenvatinib administration.

The rationale for selection of doses to be used in this study is provided in Section 4.3.

6.2.2 Handling, Storage, and Accountability

Details on preparation and administration of IV pembrolizumab are provided in the Pharmacy Manual. Participants/caregivers were given instructions on how to handle and store lenvatinib/matching placebo capsules at home until it was discontinued as per Protocol Amendment 007-06.

The investigator or designee must confirm appropriate temperature conditions have been maintained during transit for all study intervention received, and any discrepancies are reported and resolved before use of the study intervention.

Only participants enrolled in the study may receive study intervention, and only authorized site staff may supply or administer study intervention. All study interventions must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the investigator and authorized site staff.

The investigator, institution, or the head of the medical institution (where applicable) is responsible for study intervention accountability, reconciliation, and record maintenance (ie, receipt, reconciliation, and final disposition records).

For all study sites, the local country Sponsor personnel or designee will provide appropriate documentation that must be completed for drug accountability and return, or local discard and destruction if appropriate. Where local discard and destruction is appropriate, the

investigator is responsible for ensuring that a local discard/destruction procedure is documented.

The study site is responsible for recording the lot number, manufacturer, and expiry date for any locally purchased product (if applicable) as per local guidelines unless otherwise instructed by the Sponsor.

The investigator shall take responsibility for and shall take all steps to maintain appropriate records and ensure appropriate supply, storage, handling, distribution, and usage of study interventions in accordance with the protocol and any applicable laws and regulations.

6.3 Measures to Minimize Bias: Randomization and Blinding

6.3.1 Intervention Assignment

Intervention allocation/randomization will occur centrally using an interactive response technology (IRT) system. There are 2 study intervention arms. Participants will be assigned randomly in a 1:1 ratio to pembrolizumab + lenvatinib or pembrolizumab + matching placebo, respectively.

NOTE: as of Amendment 007-06, the study is unblinded. All participants remaining on treatment will receive open-label pembrolizumab monotherapy.

6.3.2 Stratification

Intervention allocation/randomization will be stratified according to the following factors:

1. Geographic region of the enrolling site: East Asia vs non-East Asia
2. ECOG Performance Score: 0 vs 1
3. TPS: 1% to 49% vs ≥50%

6.3.3 Blinding

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study and the study is unblinded. This section has been updated accordingly.

6.3.3.1 Pembrolizumab

Pembrolizumab will be administered open label; therefore, its identity will be known by the participant, the investigator, the Sponsor, and delegate(s) who are involved in study intervention administration or the clinical evaluation of participants.

6.4 Study Intervention Compliance

Interruptions from the protocol-specified treatment plan for >12 weeks (pembrolizumab) require consultation between the investigator and the Sponsor and written documentation of the collaborative decision on participant management.

6.5 Concomitant Therapy

Medications or vaccinations specifically prohibited in the exclusion criteria are not allowed during the treatment period. If there is a clinical indication for any medication or vaccination specifically prohibited, discontinuation from study intervention may be required. The investigator is to discuss prohibited medication/vaccination with the Sponsor's Clinical Director. The final decision on any supportive therapy or vaccination rests with the investigator and/or the participant's primary physician, but the decision to continue the participant on study intervention requires the mutual agreement of the investigator, the Sponsor, and the participant.

All prior medications (including over-the-counter medications) administered within 30 days before the first dose of study drug and any concomitant therapy administered to the participant during the course of the study (starting at the date of informed consent) until 30 days after the final dose (or 90 days if used to treat an SAE) of study drug will be recorded. Additionally, all diagnostic, therapeutic, or surgical procedures relating to malignancy should be recorded. Any medication that is considered necessary for the participant's health and that is not expected to interfere with the evaluation of or interact with the study interventions may be continued during the study.

6.5.1 Allowed Concomitant Medications

Treatment of complications or AEs, or therapy to ameliorate symptoms (including blood products, blood transfusions, fluid transfusions, antibiotics, and antidiarrheal drugs), may be given at the discretion of the investigator, unless it is expected to interfere with the evaluation of (or to interact with) the study intervention. Antiemetic or any other prophylaxis should be considered in accordance with institutional guidelines.

The following concomitant medications are also allowed:

- Hormone replacement therapy (HRT)
- Thyroid hormone suppressive therapy
- Anticoagulants, including low molecular weight heparin, warfarin, and anti-Xa agents
- Anti-inflammatory agents

- Bisphosphonates or denosumab
- Antihypertensive therapy (including additional antihypertensive treatment as appropriate if BP increases once the participant is enrolled)

6.5.2 Prohibited Concomitant Medications

Participants are prohibited from receiving the following therapies during the screening and treatment phases of the study:

- Concurrent anticancer therapies, such as chemotherapy, targeted therapies, antitumor interventions (surgical resection, surgical debulking of tumor, etc.), or cancer immunotherapy not specified in this protocol
  - Note: Topical anticancer agents to treat skin lesions (eg, in situ melanoma or squamous cell carcinoma) are allowed, excluding skin metastasis of melanoma.
- Other concurrent investigational drugs
- Live or attenuated vaccines within 30 days and while participating in the study. Note: killed vaccines are allowed.
- Systemic glucocorticoids for any purpose other than those listed in Section 6.5.4.1
- Radiation therapy for disease control
  - Note: Palliative radiotherapy is permitted for nontarget lesions if considered medically necessary by the treating physician and upon discussion with the Sponsor.

For participants who, in an assessment by the investigator, require the use of any of the aforementioned treatments for clinical management, continuation of the study intervention and further participation in the study must be discussed and agreed upon with the Sponsor.

If participants receive additional anticancer therapies, this will be judged to represent evidence of PD, and study intervention will be discontinued. These participants should complete all end-of-treatment assessments and continue to be followed for survival during the follow-up period.

6.5.3 Drug Interactions

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. This section is no longer applicable, but has been left unchanged for reference.

There are no DDI-related concomitant medication prohibitions or restrictions.

Lenvatinib is not expected to clinically meaningfully alter exposure to CYP3A4/ P-glycoprotein (Pgp) substrates based on results from a lenvatinib drug-drug interaction

(DDI) study with midazolam (a sensitive CYP3A and Pgp substrate). Clinical studies also showed that co-administration of lenvatinib with either inducers or inhibitors of CYP3A4/Pgp are not of clinical concern.

No drug interaction is expected between pembrolizumab and lenvatinib because of divergent metabolic pathways. Pembrolizumab is a monoclonal antibody and is primarily catabolized like other proteins, while lenvatinib is metabolized by enzymatic (CYP3A and aldehyde oxidase) and nonenzymatic processes (see lenvatinib IB).

6.5.4 Rescue Medications and Supportive Care

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. This section has been updated accordingly.

Participants should receive appropriate supportive care measures as deemed necessary by the treating investigator. Suggested supportive care measures for the management of AEs with potential immunologic etiology are outlined in Section 6.6.1.

6.5.4.1 Systemic Corticosteroid Use

Systemic corticosteroids are permitted in the following situations:

- To mediate potential immune-related adverse events (irAEs) as guided in Table 3.
- As pre- or postmedication to prevent AEs associated with IV contrast.
- Brief, limited use of systemic corticosteroids (≤7 days) are permitted where such use is considered SOC (eg, for chronic obstructive pulmonary disease [COPD] exacerbation).
- Replacement doses of steroids (for example, prednisone 10 mg daily) are permitted while on study, as is the use of local steroid injections and topical steroids.

6.6 Dose Modification

NOTE: As of Amendment 007-06, text in this section relating to dose modification and other allowed dose interruptions of lenvatinib/matching placebo is no longer applicable. This section has been updated accordingly.

6.6.1 Immune-related Events and Dose Modification (Withhold, Treat, Discontinue)

6.6.1.1 Dose Modification and Toxicity Management for Immune-related AEs Associated with Pembrolizumab

AEs associated with pembrolizumab exposure may represent an immune-related response. These irAEs may occur shortly after the first dose or several months after the last dose of pembrolizumab treatment and may affect more than one body system simultaneously. Therefore, early recognition and initiation of treatment is critical to reduce complications.

Based on existing clinical study data, most irAEs were reversible and could be managed with interruptions of pembrolizumab, administration of corticosteroids and/or other supportive care. For suspected irAEs, ensure adequate evaluation to confirm etiology or exclude other causes. Additional procedures or tests such as bronchoscopy, endoscopy, skin biopsy may be included as part of the evaluation. Based on the severity of irAEs, withhold or permanently discontinue pembrolizumab and administer corticosteroids.

Modification and toxicity management guidelines for irAEs associated with pembrolizumab monotherapy, coformulations, or IO combinations are provided in Table 3.

Table 3 Dose Modification and Toxicity Management Guidelines for Immune-related Adverse Events Associated with Pembrolizumab Monotherapy, Coformulations or IO Combinations

General instructions:

1. Severe and life-threatening irAEs should be treated with IV corticosteroids followed by oral steroids. Other immunosuppressive treatment should begin if the irAEs are not controlled by corticosteroids.
2. Pembrolizumab monotherapy, coformulations or IO combinations must be permanently discontinued if the irAE does not resolve or the corticosteroid dose is not ≤10 mg/day within 12 weeks of the last treatment.
3. The corticosteroid taper should begin when the irAE is ≤ Grade 1 and continue at least 4 weeks.
4. If pembrolizumab monotherapy, coformulations or IO combinations have been withheld, treatment may resume after the irAE decreased to ≤ Grade 1 after corticosteroid taper.

irAEs	Toxicity Grade (CTCAEv4.0)	Action With Pembrolizumab Monotherapy, Coformulations or IO Combinations	Corticosteroid and/or Other therapies	Monitoring and Follow-up
Pneumonitis	Grade 2	Withhold	Administer corticosteroids (initial dose of 1-2 mg/kg prednisone or equivalent) followed by taper	Monitor participants for signs and symptoms of pneumonitis Evaluate participants with suspected pneumonitis with radiographic imaging and initiate corticosteroid treatment Add prophylactic antibiotics for opportunistic infections
	Recurrent Grade 2 or Grade 3 or 4	Permanently discontinue		

irAEs	Toxicity Grade (CTCAEv4.0)	Action With Pembrolizumab Monotherapy, Coformulations or IO Combinations	Corticosteroid and/or Other therapies	Monitoring and Follow-up
Diarrhea / Colitis	Grade 2 or 3	Withhold	Administer corticosteroids (initial dose of 1-2 mg/kg prednisone or equivalent) followed by taper	Monitor participants for signs and symptoms of enterocolitis (ie, diarrhea, abdominal pain, blood or mucus in stool with or without fever) and of bowel perforation (ie, peritoneal signs and ileus) Participants with ≥Grade 2 diarrhea suspecting colitis should consider GI consultation and performing endoscopy to rule out colitis Participants with diarrhea/colitis should be advised to drink liberal quantities of clear fluids. If sufficient oral fluid intake is not feasible, fluid and electrolytes should be substituted via IV infusion.
	Recurrent Grade 3 or Grade 4	Permanently discontinue		
AST / ALT Elevation or Increased Bilirubin	Grade 2	Withhold	Administer corticosteroids (initial dose of 0.5-1 mg/kg prednisone or equivalent) followed by taper	Monitor with liver function tests (consider weekly or more frequently until liver enzyme value returned to baseline or is stable)
	Grade 3 or 4	Permanently discontinue	Administer corticosteroids (initial dose of 1-2 mg/kg prednisone or equivalent) followed by taper	
T1DM or Hyperglycemia	New onset T1DM or Grade 3 or 4 hyperglycemia associated with evidence of β-cell failure	Withhold [a]	Initiate insulin replacement therapy for participants with T1DM Administer antihyperglycemic in participants with hyperglycemia	Monitor participants for hyperglycemia or other signs and symptoms of diabetes

irAEs	Toxicity Grade (CTCAEv4.0)	Action With Pembrolizumab Monotherapy, Coformulations or IO Combinations	Corticosteroid and/or Other therapies	Monitoring and Follow-up
Hypophysitis	Grade 2	Withhold	Administer corticosteroids and initiate hormonal replacements as clinically indicated	Monitor for signs and symptoms of hypophysitis (including hypopituitarism and adrenal insufficiency)
	Grade 3 or 4	Withhold or permanently discontinue [a]		
Hyperthyroidism	Grade 2	Continue	Treat with nonselective beta-blockers (eg, propranolol) or thionamides as appropriate	Monitor for signs and symptoms of thyroid disorders
	Grade 3 or 4	Withhold or Permanently discontinue [a]		
Hypothyroidism	Grade 2-4	Continue	Initiate thyroid replacement hormones (eg, levothyroxine or liothyronine) per standard of care	Monitor for signs and symptoms of thyroid disorders
Nephritis and renal dysfunction	Grade 2	Withhold	Administer corticosteroids (prednisone 1-2 mg/kg or equivalent) followed by taper	Monitor changes of renal function
	Grade 3 or 4	Permanently discontinue		
Myocarditis	Grade 1	Withhold	Based on severity of AE administer corticosteroids	Ensure adequate evaluation to confirm etiology and/or exclude other causes
	Grade 2, 3 or 4	Permanently discontinue		

irAEs	Toxicity Grade (CTCAEv4.0)	Action With Pembrolizumab Monotherapy, Coformulations or IO Combinations	Corticosteroid and/or Other therapies	Monitoring and Follow-up
All Other irAEs	Persistent Grade 2	Withhold	Based on severity of AE administer corticosteroids	Ensure adequate evaluation to confirm etiology or exclude other causes
	Grade 3	Withhold or discontinue [b]		
	Recurrent Grade 3 or Grade 4	Permanently discontinue		

AE(s)=adverse event(s); ALT= alanine aminotransferase; AST=aspartate aminotransferase; CTCAE=Common Terminology Criteria for Adverse Events; DRESS=Drug Rash with Eosinophilia and Systemic Symptom; GI=gastrointestinal; IO=immuno-oncology; ir=immune-related; IV=intravenous; SJS=Stevens-Johnson Syndrome; T1DM=type 1 diabetes mellitus; TEN=Toxic Epidermal Necrolysis; ULN=upper limit of normal.

Note: Non-irAE will be managed as appropriate, following clinical practice recommendations.

[a] The decision to withhold or permanently discontinue pembrolizumab monotherapy, coformulations or IO combinations is at the discretion of the investigator or treating physician. If control achieved or ≤ Grade 2, pembrolizumab monotherapy, coformulations or IO combinations may be resumed.

[b] Events that require discontinuation include, but are not limited to: Guillain-Barre Syndrome, encephalitis, myelitis, DRESS, SJS, TEN and other clinically important irAEs (eg, vasculitis and sclerosing cholangitis).

6.6.1.2 Dose Modification and Toxicity Management of Infusion Reactions Related to Pembrolizumab

Pembrolizumab may cause severe or life-threatening infusion reactions, including severe hypersensitivity or anaphylaxis. Signs and symptoms usually develop during or shortly after drug infusion and generally resolve completely within 24 hours of completion of infusion. Dose modification and toxicity management guidelines for pembrolizumab-associated infusion reactions are provided in Table 4.

Table 4 Pembrolizumab Infusion Reaction Dose Modification and Treatment Guidelines

NCI CTCAE Grade	Treatment	Premedication at Subsequent Dosing
Grade 1 Mild reaction; infusion interruption not indicated; intervention not indicated	Increase monitoring of vital signs as medically indicated until the participant is deemed medically stable in the opinion of the investigator.	None
Grade 2 Requires therapy or infusion interruption but responds promptly to symptomatic treatment (eg, antihistamines, NSAIDs, narcotics, IV fluids); prophylactic medications indicated for ≤24 h	**Stop Infusion.** Additional appropriate medical therapy may include but is not limited to the following: - IV fluids - Antihistamines - NSAIDs - Acetaminophen - Narcotics Increase monitoring of vital signs as medically indicated until the participant is deemed medically stable in the opinion of the investigator. If symptoms resolve within 1 hour of stopping drug infusion, the infusion may be restarted at 50% of the original infusion rate (eg, from 100 mL/h to 50 mL/h). Otherwise, dosing will be held until symptoms resolve and the participant should be premedicated for the next scheduled dose. **Participants who develop Grade 2 toxicity despite adequate premedication should be permanently discontinued from further pembrolizumab treatment.**	Participant may be premedicated 1.5 h (± 30 minutes) before infusion of pembrolizumab with: - Diphenhydramine, 50 mg PO (or equivalent dose of antihistamine). - Acetaminophen, 500-1000 mg PO (or equivalent dose of analgesic).

NCI CTCAE Grade	Treatment	Premedication at Subsequent Dosing
Grades 3 or 4 **Grade 3:** Prolonged (ie, not rapidly responsive to symptomatic medication and/or brief interruption of infusion); recurrence of symptoms following initial improvement; hospitalization indicated for other clinical sequelae (eg, renal impairment, pulmonary infiltrates) **Grade 4:** Life-threatening; pressor or ventilatory support indicated	**Stop Infusion.** Additional appropriate medical therapy may include but is not limited to: - Epinephrine** - IV fluids - Antihistamines - NSAIDs - Acetaminophen - Narcotics - Oxygen - Pressors - Corticosteroids Increase monitoring of vital signs as medically indicated until the participant is deemed medically stable in the opinion of the investigator. Hospitalization may be indicated. **In cases of anaphylaxis, epinephrine should be used immediately. **Participant is permanently discontinued from further pembrolizumab treatment.**	No subsequent dosing
Abbreviations: CTCAE = Common Terminology Criteria for Adverse Events; IV = intravenous; NCI = National Cancer Institute; NSAIDs = nonsteroidal anti-inflammatory drugs; PO = by mouth. Appropriate resuscitation equipment should be available at the bedside and a physician readily available during the period of drug administration. For further information, please refer to the CTCAE, version 4.0, at http://ctep.cancer.gov.		

6.6.2 Other Allowed Dose Interruptions

Pembrolizumab may be interrupted for situations other than treatment-related AEs, such as medical/surgical events or logistical reasons not related to study therapy. Participants should be placed back on study therapy within 3 weeks of the scheduled interruption, unless otherwise discussed with the Sponsor. The reason for interruption should be documented in the participant's study record. Scans should not be delayed for delays in cycle treatment.

6.7 Intervention After the End of the Study

There is no study-specified intervention after the end of the study.

6.8 Clinical Supplies Disclosure

NOTE: As of Amendment 007-06, the study is unblinded and this section is no longer applicable.

6.9 Standard Policies

At the close of the study after unblinding, a letter is to be sent by the investigator to those participants who received placebo in the image of the competitor’s product to provide the following advice:

“You have participated in a study conducted by the Sponsor. This letter is to advise you that you were among those who received a look-alike capsule provided by the Sponsor to resemble the drug lenvatinib as much as possible. You did not receive the active drug lenvatinib as provided by Merck.”

7 DISCONTINUATION OF STUDY INTERVENTION AND PARTICIPANT WITHDRAWAL

7.1 Discontinuation of Study Intervention

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study and participants stopped receiving those. This section has been updated accordingly.

Discontinuation of study intervention does not represent withdrawal from the study.

As certain data on clinical events beyond study intervention discontinuation may be important to the study, they must be collected through the participant’s last scheduled follow-up, even if the participant has discontinued study intervention. Therefore, all participants who discontinue study intervention prior to completion of the protocol-specified treatment period will still continue to participate in the study as specified in Sections 1.3 and 8.11.4 unless the participant has withdrawn from the study (Section 7.2).

Participants may discontinue study intervention at any time for any reason or be discontinued from the study intervention at the discretion of the investigator should any untoward effect

occur. In addition, a participant may be discontinued from study intervention by the investigator or the Sponsor if study intervention is inappropriate, the study plan is violated, or for administrative and/or other safety reasons. Specific details regarding procedures to be performed at study intervention discontinuation are provided in Sections 8.1.9 and 8.11.4.

A participant must be discontinued from study intervention but continue to be monitored in the study for any of the following reasons:

- The participant or participant's legally acceptable representative requests to discontinue study intervention.
- The participant interrupts study intervention administration for more than 12 weeks, except if agreed to by the Sponsor.
- The participant has a medical condition or personal circumstance that, in the opinion of the investigator and/or Sponsor, places the participant at unnecessary risk from continued administration of the study intervention.
- The participant has a confirmed positive serum pregnancy test.
- The participant has intercurrent illness that prevents further administration of treatment.
- The participant has any progression or recurrence of any malignancy or occurrence of another malignancy that requires active treatment. Exceptions to malignancy include basal cell carcinoma of the skin, squamous cell carcinoma of the skin, new nonulcerated primary melanoma <1 mm in depth with no nodal involvement, or carcinoma in situ (eg, breast carcinoma, cervical cancer in situ) that have undergone potentially curative therapy. Exceptions should be discussed with the Sponsor before continuing therapy or remaining in follow-up.
- The participant has unacceptable AE(s) or toxicities.
- Any study intervention-related toxicity specified as a reason for permanent discontinuation as defined in the guidelines for dose modification due to AEs in Section 6.6.

7.1.1 Second-course Treatment Phase

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from the study.

7.2 Participant Withdrawal From the Study

A participant must be withdrawn from the study if the participant or participant's legally acceptable representative withdraws consent from the study.

If a participant withdraws from the study, they will no longer receive study intervention or be followed at scheduled protocol visits.

Specific details regarding procedures to be performed at the time of withdrawal from the study are outlined in Section 8.1.9. The procedures to be performed should a participant repeatedly fail to return for scheduled visits and/or if the study site is unable to contact the participant are outlined in Section 7.3.

7.3 Lost to Follow-up

If a participant fails to return to the clinic for a required study visit and/or if the site is unable to contact the participant, the following procedures are to be performed:

- The site must attempt to contact the participant and reschedule the missed visit. If the participant is contacted, the participant should be counseled on the importance of maintaining the protocol-specified visit schedule.

- The investigator or designee must make every effort to regain contact with the participant at each missed visit (eg, telephone calls and/or a certified letter to the participant's last known mailing address or locally equivalent methods). These contact attempts should be documented in the participant's medical record.

- Note: A participant is not considered lost to follow-up until the last scheduled visit for the individual participant. The missing data for the participant will be managed via the prespecified statistical data handling and analysis guidelines.

8 STUDY ASSESSMENTS AND PROCEDURES

- Study procedures and their timing are summarized in the SoA.

- Adherence to the study design requirements, including those specified in the SoA, is essential and required for study conduct.

- The investigator is responsible for ensuring that procedures are conducted by appropriately qualified or trained staff. Delegation of study site personnel responsibilities will be documented in the Investigator Trial File Binder (or equivalent).

- All study-related medical decisions must be made by an investigator who is a qualified physician.

- All screening evaluations must be completed and reviewed to confirm that potential participants meet all eligibility criteria. The investigator will maintain a screening log to record details of all participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

- Procedures conducted as part of the participant's routine clinical management (eg, blood count) and obtained before signing of ICF may be utilized for screening or baseline purposes provided the procedure met the protocol-specified criteria and were performed within the time frame defined in the SoA.

- Additional evaluations/testing may be deemed necessary by the investigator and or the Sponsor for reasons related to participant safety. In some cases, such evaluation/testing may be potentially sensitive in nature (eg, HIV, hepatitis C), and thus local regulations may require that additional informed consent be obtained from the participant. In these cases, such evaluations/testing will be performed in accordance with those regulations.

Repeat or unscheduled samples may be taken for safety reasons or for technical issues with the samples.

8.1 Administrative and General Procedures

8.1.1 Informed Consent

The investigator or medically qualified designee (consistent with local requirements) must obtain documented informed consent from each potential participant or their legally acceptable representative prior to participating in this clinical study or future biomedical research. If there are changes to the participant's status during the study (eg, health or age of majority requirements), the investigator or medically qualified designee must ensure the appropriate documented informed consent is in place.

8.1.1.1 General Informed Consent

Informed consent given by the participant or their legally acceptable representative must be documented on a consent form. The form must include the trial protocol number, trial protocol title, dated signature, and /agreement of the participant (or his/her legally acceptable representative) and of the person conducting the consent discussion.

A copy of the signed and dated informed consent form should be given to the participant (or their legally acceptable representative) before participation in the study.

The initial ICF, any subsequent revised written ICF, and any written information provided to the participant must receive the IRB/IEC's approval/favorable opinion in advance of use. The participant or his/her legally acceptable representative should be informed in a timely manner if new information becomes available that may be relevant to the participant's willingness to continue participation in the study. The communication of this information will be provided and documented via a revised consent form or addendum to the original consent form that captures the participant's or the participant's legally acceptable representative's dated signature.

If the investigator recommends continuation of study intervention beyond disease progression, the participant or his/her legally acceptable representative will be asked to sign consent.

Specifics about the study and the study population are to be included in the study informed consent form.

Informed consent will adhere to IRB/IEC requirements, applicable laws and regulations, and Sponsor requirements.

8.1.2 Inclusion/Exclusion Criteria

All inclusion and exclusion criteria will be reviewed by the investigator who is a qualified physician to ensure that the participant qualifies for the study.

8.1.3 Participant Identification Card

All participants will be given a participant identification card identifying them as participants in a research study. The card will contain study site contact information (including direct telephone numbers) to be used in the event of an emergency. The investigator or qualified designee will provide the participant with a participant identification card immediately after the participant provides documented informed consent. At the time of intervention allocation/randomization, site personnel will add the treatment/randomization number to the participant identification card.

The participant identification card also contains contact information for the emergency unblinding call center so that a healthcare provider can obtain information about study intervention in emergency situations where the investigator is not available.

8.1.4 Medical History

8.1.4.1 General Medical History

A medical history will be obtained by the investigator or qualified designee.

Medical history will include all active conditions, drug allergies, significant medical procedures, and any condition diagnosed within the previous 10 years that are considered to be clinically important by the investigator. Any cancer, other than the cancer under study, will be recorded as medical history, even if diagnosed greater than 10 years before enrollment. Details regarding the cancer under study will be recorded separately and not listed as medical history. The medical history will also include an assessment of smoking history.

8.1.4.2 Oncologic Disease Details

The investigator or qualified designee will obtain historic and current details of the participant's cancer under study. This information will include, but is not limited to, date of

diagnosis, stage, histology, locations of primary lesions, and location of metastases, if applicable.

8.1.5 Prior and Concomitant Medications Review

8.1.5.1 Prior Medications

The investigator or qualified designee will review prior medication use, including any protocol-specified washout requirement, and record prior medication taken by the participant within 30 days before the first dose of study intervention. Treatment for the disease for which the participant has enrolled in the study will be recorded separately and not listed as a prior medication.

8.1.5.2 Prior Oncologic Treatment

The investigator or qualified designee will review and record all treatments for the cancer under study, including systemic and local treatment, vaccinations, radiation, and surgeries. Additional information collected on these treatments will include, but is not limited to, reason for discontinuation, best response, and date of progression after each treatment as applicable.

8.1.5.3 Concomitant Medications

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from the study. This section has been updated accordingly.

The investigator or qualified designee will record medication, if any, taken by the participant during the study. Concomitant medications will be recorded for 30 days after the last dose of study intervention (or 90 days if used to treat an SAE).

Any new anticancer therapy started after the participant's discontinuation from the treatment period will be recorded separately. Additional information collected on this treatment will include, but is not limited to, best response and date of progression.

8.1.6 Assignment of Screening Number

All consented participants will be given a unique screening number that will be used to identify the participant for all procedures that occur prior to randomization. Each participant will be assigned only 1 screening number. Screening numbers must not be re-used for different participants.

Any participant who is screened multiple times will retain the original screening number assigned at the initial screening visit. Specific details on the screening/rescreening visit requirements are provided in Section 8.11.1.

8.1.7 Assignment of Treatment/Randomization Number

All eligible participants will be randomly allocated and will receive a treatment/randomization number. The treatment/randomization number identifies the participant for all procedures occurring after treatment allocation/randomization. Once a treatment/randomization number is assigned to a participant, it can never be re-assigned to another participant.

A single participant cannot be assigned more than 1 treatment/randomization number.

8.1.8 Study Intervention Administration

Study intervention should begin within 3 days of randomization.

Study intervention will be administered by the investigator and/or study staff according to the specifications within the Pharmacy Manual.

8.1.8.1 Timing of Dose Administration

8.1.8.1.1 Lenvatinib/Matching Placebo

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. This section is no longer applicable.

Lenvatinib/matching placebo 20 mg (two 10-mg capsules) QD will be taken orally with water (with or without food) in 21-day cycles at approximately the same time each day. On Day 1 of each pembrolizumab cycle, lenvatinib/matching placebo will be administered 0-4 hours after pembrolizumab.

If a lenvatinib/matching placebo dose is missed and cannot be taken within 12 hours, then that dose should be skipped and the next dose should be taken at the usual time of administration.

For nonvisit days, lenvatinib will be taken at home.

When a participant attends a study visit, he or she will bring any unused capsules.

8.1.8.1.2 Pembrolizumab

Pembrolizumab will be administered as a 30-minute IV infusion on Day 1 of each 21-day cycle. Sites should make every effort to target infusion timing to be as close to 30 minutes as possible. However, given the variability of infusion pumps from site to site, a window of –5 minutes and +10 minutes is permitted (ie, infusion time is 30 minutes: –5 min/+10 min).

After Cycle 1 Day 1, pembrolizumab may be administered up to 3 days before or after the scheduled Day 1 of each subsequent cycle for administrative reasons.

Administration of pembrolizumab will be witnessed by the investigator and/or study staff. The total volume of study treatment infused will be compared with the total volume prepared to determine compliance with each dose administered.

8.1.8.2 Lenvatinib Compliance

Lenvatinib compliance will be calculated by the Sponsor until discontinuation, based on the drug accountability documented by the site staff and monitored by the Sponsor/designee. The objective is 100% compliance, and investigators and their staff should evaluate compliance at each visit and take appropriate steps to optimize compliance.

8.1.9 Discontinuation and Withdrawal

Participants who discontinue study intervention prior to completion of the treatment period should be encouraged to continue to be followed for all remaining study visits as outlined in the SoA and Section 8.11.4.

When a participant withdraws from participation in the study, all applicable activities scheduled for the final study visit should be performed (at the time of withdrawal). Any AEs that are present at the time of withdrawal should be followed in accordance with the safety requirements outlined in Section 8.4.3.

8.1.10 Participant Blinding/Unblinding

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study and the study is unblinded. Participants who remain on treatment will receive open-label pembrolizumab monotherapy as per protocol.

8.1.11 Calibration of Equipment

The investigator or qualified designee has the responsibility to ensure that any device or instrument used for a clinical evaluation/test during a clinical study that provides information about inclusion/exclusion criteria and/or safety or efficacy parameters shall be suitably calibrated and/or maintained to ensure that the data obtained is reliable and/or reproducible. Documentation of equipment calibration must be retained as source documentation at the study site.

8.2 Efficacy/Immunogenicity Assessments

8.2.1 Tumor Scans and Assessment of Disease

NOTE: After amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed; similarly, verification of PD is no longer needed before treatment discontinuation. This section was updated accordingly.

Throughout this section, the term “scan” refers to any medical imaging data used to assess tumor burden and may include cross-sectional imaging (such as CT or MRI), or other methods as specified in this protocol.

The process for scan collection and transmission to the central imaging vendor can be found in the Site Imaging Manual. Tumor scans are strongly preferred to be acquired by computed tomography (CT). For the abdomen and pelvis, contrast-enhanced MRI may be used when CT with iodinated contrast is contraindicated, or when mandated by local practice. The same scan technique regarding modality, ideally the same scanner, and the use of contrast should be used in a participant throughout the study to optimize the reproducibility of the assessment of existing and new tumor burden and improve the accuracy of the assessment of response or progression based on scans. Note: for the purposes of assessing tumor scans, the term "investigator" refers to the local investigator at the site and/or the radiological reviewer located at the site or at an offsite facility.

MRI is preferred for brain scans; however, CT scans will be acceptable, if MRI is medically contraindicated.

Bone scans may be performed to evaluate bone metastases. Any supplemental scans performed to support a positive or negative bone scan, such as plain X-rays acquired for correlation, should be submitted to the central imaging vendor.

Participant eligibility will be determined using local assessment (investigator assessment) based on RECIST 1.1. All scheduled scans for each participant will be submitted to the central imaging vendor. In addition, unscheduled scans to determine PD and scans obtained for other reasons, but demonstrate radiologic progression, are to be submitted to the central imaging vendor.

When the investigator identifies disease progression per RECIST 1.1, but elects to implement iRECIST, the investigator will assess for confirmation of progression by iRECIST at subsequent time points. Images should continue to be submitted to the central imaging vendor.

8.2.1.1 Initial Tumor Scans

Initial tumor scans at screening must be performed within 28 days before the date of randomization. The site study team must review screening scans to confirm the participant has measurable disease per RECIST 1.1.

The screening scans must be submitted to the central imaging vendor for retrospective eligibility review.

Tumor scans performed as part of routine clinical management is acceptable for use as screening tumor scans if it is of diagnostic quality and performed within 28 days before the date of randomization and can be assessed by the central imaging vendor.

If brain scan is performed to document the stability of existing metastases, an MRI scan should be used if possible. If MRI is medically contraindicated, CT with contrast is an acceptable alternative.

8.2.1.2 Tumor Scans During the Study

The first on-study scan assessment should be performed at 9 weeks (63 days ±7 days) from the date of randomization. Subsequent tumor scans should be performed every 9 weeks (63 days ±7 days) or more frequently if clinically indicated. After 54 weeks, participants who remain on treatment will have scans performed every 12 weeks (84 days ±7 days). Scan timing should follow calendar days and should not be adjusted for delays in cycle starts. Scans should continue to be performed until PD is identified by the investigator (unless the investigator elects to continue treatment and follow iRECIST), the start of new anticancer treatment, withdrawal of consent, death, or any other reason for discontinuation of pembrolizumab listed in Section 7.1 is met, whichever occurs first. All supplemental scans must be submitted to the central imaging vendor.

OR should be confirmed by a repeat scan assessment. Tumor scans to confirm PR or CR should be performed at least 4 weeks after the first indication of a response is observed. Participants will then return to regular scheduled scans, starting with the next scheduled time point. Participants who receive additional scans for confirmation do not need to undergo the next scheduled tumor scan if it is less than 4 weeks later; tumor scans may resume at the subsequent scheduled time point. Note: Response does not need to be verified in real time by the central imaging vendor.

Per iRECIST (Section 8.2.1.6), PD should be confirmed by the site 4 to 8 weeks after site-assessed first radiologic evidence of PD in clinically stable participants. Participants who have unconfirmed PD may continue on treatment at the discretion of the investigator until PD is confirmed by the site, provided they have met the conditions detailed in Section 8.2.1.6. Participants who receive confirmatory scans do not need to undergo the next scheduled tumor scan if it is less than 4 weeks later; tumor scans may resume at the subsequent scheduled time point, if clinically stable. Participants who have confirmed PD by iRECIST, as assessed by the site, will discontinue study intervention. Exceptions are detailed in Section 8.2.1.6.

8.2.1.3 End of Treatment and Follow-up Tumor Scans

For participants who discontinue study intervention, tumor scans should be performed at the time of treatment discontinuation (±4-week window). If previous scans were obtained within 4 weeks before the date of discontinuation, then scans at treatment discontinuation is not mandatory. For participants who discontinue study intervention because of documented PD, this is the final required tumor scan.

After amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed; similarly, verification of PD is no longer needed before treatment discontinuation.

8.2.1.4 Second-course Treatment Tumor Scans

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from this study. This section is no longer applicable.

8.2.1.5 RECIST 1.1 Assessment of Disease

RECIST 1.1 will be used by BICR as the primary measure for assessment of tumor response and date of PD, and as a basis for all protocol guidelines related to disease status (eg, discontinuation of study intervention). Although RECIST 1.1 references a maximum of 5 target lesions in total and 2 per organ, this protocol follows a maximum of 10 target lesions and a maximum of 5 target lesions per organ, if clinically relevant to enable a broader sampling of tumor burden.

If disease progression has been assessed by the investigator, the process continues as follows:

- If participant is clinically stable, continue study intervention per protocol
    - Resume imaging per protocol schedule (≥4 weeks to next scan)
    - Send scans to imaging vendor
    - Continue local assessment
    - Do not change investigator assessment of progression
    - If subsequent scan(s) indicate progression, submit scan(s) to imaging vendor
- If the participant is not clinically stable, best medical practice is to be applied

Before stopping study intervention or imaging or starting new anticancer therapy in a participant who is clinically stable, communication with the Sponsor is required.

- Investigator judgment will determine action

  Note: the reconsent addendum may be signed any time after investigator-assessed progression is identified, but must be signed prior to starting study intervention.

    - Obtain scans locally per original protocol schedule
    - Do not send scans to imaging vendor

For the purpose of this decision process, lack of clinical stability is defined as:

- Unacceptable toxicity
- Clinical signs or symptoms indicating clinically significant disease progression
- Decline in performance status
- Rapid disease progression or threat to vital organs or critical anatomical sites (eg, CNS metastasis, respiratory failure due to tumor compression, spinal cord compression) requiring urgent alternative medical intervention

8.2.1.6 iRECIST Assessment of Disease

NOTE: After amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed, verification of PD is no longer needed before treatment discontinuation, and posttreatment imaging is no longer required. This section has been updated accordingly.

iRECIST is based on RECIST 1.1 but adapted to account for the unique tumor response seen with immunotherapeutic drugs. iRECIST will be used by the investigator to assess tumor response and progression and make treatment decisions [Seymour, L., et al 2017]. When clinically stable, participants will continue study intervention beyond RECIST 1.1 PD with continuous assessment of response according to the rules outlined in Appendix 6. iRECIST reflects some participants can have a transient tumor flare after the start of immunotherapy, then experience subsequent disease response. This data will be captured in the clinical database.

Clinical stability is defined as the following:

- Absence of symptoms and signs indicating clinically significant PD
- No decline in ECOG performance status
- No requirements for intensified management, including increased analgesia, radiation, or other palliative care

Any participant deemed clinically unstable should be discontinued from study intervention at site-assessed first radiologic evidence of PD and is not required to have repeat tumor scans for confirmation of PD by iRECIST.

If the investigator decides to continue treatment, the participant may continue to receive study intervention and the tumor assessment should be repeated 4 to 8 weeks later to confirm PD by iRECIST, per investigator assessment. Scans should continue to be sent in to the central imaging vendor .

If repeat scans do not confirm PD per iRECIST, as assessed by the investigator, and the participant continues to be clinically stable, study intervention may continue and follow the regular scan schedule. If PD is confirmed, participants will be discontinued from study intervention.

If a participant has iRECIST-confirmed radiographic progression (iCPD) as defined in Appendix 6, study intervention should be discontinued; however, if the participant is achieving a clinically meaningful benefit, an exception to continue study intervention may be considered following consultation with the Sponsor. In this case, if study intervention is continued, tumor scans should continue to be performed following the intervals as outlined in Section 1 and submitted to the central imaging vendor.

A description of the adaptations and iRECIST process is provided in Appendix 6, with additional details in the iRECIST publication [Seymour, L., et al 2017]. A summary of scans and treatment requirements after first radiologic evidence of progression is provided in Table 5.

Table 5 Summary of Imaging and Treatment Requirements After First Radiologic Evidence of Progression

	Clinically Stable		**Clinically Unstable**	
	Imaging	**Treatment**	**Imaging**	**Treatment**
First radiologic evidence of PD by RECIST 1.1 per investigator assessment	~~Submit the imaging to BICR for verification.~~ Repeat imaging at 4 to 8 weeks to confirm PD per iRECIST.	May continue study intervention at the discretion of the investigator and after obtaining participant's consent.	No additional imaging required.	Discontinue treatment
Repeat tumor imaging confirms PD (iCPD) by iRECIST per investigator assessment	No additional imaging required.	Discontinue treatment (exception is possible upon consultation with Sponsor).	No additional imaging required.	Not applicable
Repeat tumor imaging shows iUPD by iRECIST per investigator assessment	Repeat imaging at 4 to 8 weeks to confirm PD.	Continue study intervention at the investigator's discretion.	No additional imaging required.	Discontinue treatment
Repeat tumor imaging shows iSD, iPR, or iCR by iRECIST per investigator assessment.	Continue regularly scheduled imaging assessments.	Continue study intervention at the investigator's discretion.	No additional imaging required.	Discontinue treatment

Abbreviations: iCPD = iRECIST-confirmed progressive disease; iCR = iRECIST complete response; iRECIST = adjusted Response Evaluation Criteria in Solid Tumors 1.1 for immune-based therapeutics; iSD = iRECIST stable disease; iUPD = iRECIST unconfirmed progressive disease; PD = progressive disease; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors, version 1.1; VOP = verification of progression.

Note: If PD has been centrally verified, further management is performed by the site, based on iRECIST. Any further imaging should still be submitted to the central imaging vendor, but no rapid review will occur. If RECIST 1.1 PD has not been centrally verified, ideally the site should continue treatment. Whether or not treatment continues, imaging should be collected and submitted to the central imaging vendor with VOP request until RECIST 1.1 PD is verified by BICR.

As of amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed for verification of progression. This table was updated accordingly.

8.2.1.7 Tumor Tissue Collection for PD-L1 Status

Participation in this study will be dependent upon supplying tumor tissue for PD-L1 testing from locations not radiated before biopsy; formalin-fixed specimens after the participant has been diagnosed with metastatic disease will be preferred for determination of PD-L1 status before randomization. Biopsies obtained before receipt of adjuvant/neoadjuvant chemotherapy will be permitted if recent biopsy is not feasible.

All participants should submit either a newly obtained core or excisional biopsy or archival tissue (fine-needle aspiration is not adequate for both archival and new tissue samples) to a central laboratory for characterization of PD-L1 status before treatment allocation.

A fine-needle aspirate, frozen sample, plastic-embedded sample, cell block, clot, bone, bone marrow, cytologic specimen, decalcified or formalin-fixed sample that was frozen at any point will not be acceptable for analysis.

Note: Submission of formalin-fixed paraffin-embedded tumor tissue sample blocks are preferred; if submitting unstained slides, the slides should be freshly cut and submitted to the testing laboratory within 14 days from the site slide section date; otherwise, a new specimen will be requested.

If the sample is determined to be nonevaluable before testing by the central laboratory, a new sample should be submitted if available.

The central vendor PD-L1 TPS results of participants will be blinded to the investigator. The Sponsor acknowledges that because of the commercial availability of PD-L1 testing assays, it is possible that the investigator may know a participant's TPS before screening. This risk is seen as acceptable, as the treatment interventions are hypothesized to provide benefit regardless of TPS.

8.2.2 Patient-reported Outcomes

NOTE: As of Amendment 007-06, PROs will no longer be collected. This section is no longer applicable.

The EORTC QLQ-C30, EORTC QLQ-LC13, and EQ-5D-5L questionnaires will be administered by trained site personnel and completed electronically by participants in the following order: EORTC QLQ-C30, EORTC QLQ-LC13, and lastly EQ-5D-5L. The questionnaires should be administered before dosing at every cycle through Cycle 17, then every other cycle through Cycle 35 (eg, Cycles 1-17, 19, 21, 23, 25, 27, 29, 31, 33, and 35), at the Treatment Discontinuation Visit, and at the 30-day Safety Follow-up Visit.

If the Treatment Discontinuation Visit occurs 30 days from the last dose of study intervention, at the time of the mandatory Safety Follow-up Visit, the electronic patient-reported outcomes (ePROs) do not need to be repeated.

It is best practice and strongly recommended that the ePROs are administered to randomized participants before drug administration, AE evaluation, and disease status notification. If the participant does not complete the ePROs at a scheduled time point, the MISS_MODE form must be completed to capture the reason the assessment was not performed.

8.3 Safety Assessments

Details regarding specific safety procedures/assessments to be performed in this study are provided. Planned time points for all safety assessments are provided in the SoA.

8.3.1 Physical Examinations

The investigator or qualified designee will perform a complete physical examination during the screening period. Clinically significant abnormal findings should be recorded as medical history. The time points for full physical examination are described in the SoA (Section 1.3). After the first dose of study treatment, new clinically significant abnormal findings should be recorded as AEs.

For cycles that do not require a full physical examination per the SoA (Section 1.3), the investigator or qualified designee will perform a directed physical examination as clinically indicated before study treatment administration. New clinically significant abnormal findings should be recorded as AEs.

Investigators should pay special attention to clinical signs related to previous serious illnesses.

8.3.2 Vital Signs

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. Repeat BP measurements for monitoring participants receiving lenvatinib are no longer required. This section has been updated accordingly.

The investigator or qualified designee will take vital signs at screening, before the administration of each dose of study treatment and during the follow-up period, as specified in the SoA (Section 1.3). Vital signs include temperature, heart rate, respiratory rate, weight, and blood pressure. Height will be measured at Visit 1 only.

8.3.3 Electrocardiograms

NOTE: As of Amendment 007-06, lenvatinib and matching placebo have been removed from the study. Additional ECG measurements for monitoring participants receiving lenvatinib are no longer required. This section has been updated accordingly.

Electrocardiograms (ECGs) will be obtained as designated in the SoA (Section 1.3). Complete, standardized, 12-lead ECG recordings that permit all 12 leads to be displayed on a single page with an accompanying lead II rhythm strip below the customary 3 × 4 lead format are to be used. In addition to a rhythm strip, a minimum of 3 full complexes should be recorded from each lead simultaneously. Participants must be in the recumbent position for a

period of 5 minutes prior to the ECG. The Fridericia correction method for calculating QTc will be used.

An ECG abnormality may meet the criteria of an AE as described in this protocol (see Appendix 3) and the eCRF Entry Guidelines. In these instances, the AE corresponding to the ECG abnormality will be recorded on the appropriate case report form (CRF).

8.3.4 Echocardiograms or Multigated Acquisition Scans

A MUGA scan (using a technetium-based tracer) or an ECHO will be performed to assess LVEF as designated in the SoA (Section 1.3). Additional assessments may be performed as clinically indicated.

MUGA scans or ECHOs should be performed locally in accordance with the institution's standard practice. MUGA scans are the preferred modality. However, whichever modality is used for an individual participant at baseline should be repeated for all subsequent LVEF assessments for that participant.

LVEF, as assessed by the institution, will be entered in the electronic case report form (eCRF). Investigator assessment will be based on institutional reports.

8.3.5 Clinical Safety Laboratory Assessments

8.3.5.1 Laboratory Safety Evaluations (Hematology, Chemistry and Urinalysis)

- Refer to Appendix 2 for the list of clinical laboratory tests to be performed and to the SoA for the timing and frequency.

- The investigator or medically qualified designee (consistent with local requirements) must review the laboratory report, document this review, and record any clinically relevant changes occurring during the study in the AE section of the case report form (CRF). The laboratory reports must be filed with the source documents. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

- CBC with differential and clinical chemistry results must be reviewed before administration of study intervention. Electrolytes such as potassium, calcium, and magnesium should be monitored and abnormalities, when considered clinically significant, should be corrected in all participants before starting study intervention.

- All protocol-required laboratory assessments, as defined in Appendix 2, must be conducted in accordance with the laboratory manual and the SoA.

- If laboratory values from nonprotocol specified laboratory assessments performed at the institution's local laboratory require a change in study participant management or

are considered clinically significant by the investigator (eg, SAE or AE or dose modification), then the results must be recorded in the appropriate CRF (eg, SLAB).

- For any laboratory tests with values considered clinically significantly abnormal during participation in the study or within 30 days after the last dose of study intervention, every attempt should be made to perform repeat assessments until the values return to normal or baseline or if a new baseline is established as determined by the investigator.

8.3.5.2 Pregnancy Test

All women who are being considered for participation in the study, and who are not surgically sterilized or postmenopausal, must be tested for pregnancy within 24 hours of the first dose of study intervention. If a urine test is positive or not evaluable, a serum test will be required. Participants must be excluded/discontinued from the study in the event of a positive test result. Repeated pregnancy test (such as monthly testing) may be conducted if required by local regulations.

8.4 Adverse Events (AEs), Serious Adverse Events (SAEs), and Other Reportable Safety Events

The definitions of an AE or SAE, as well as the method of recording, evaluating, and assessing causality of AE and SAE and the procedures for completing and transmitting AE, SAE, and other reportable safety event reports can be found in Appendix 3.
Progression of the cancer under study is not considered an AE as described in Section 8.4.6 and Appendix 3.

Adverse events, SAEs, and other reportable safety events will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative).

The investigator and any designees are responsible for detecting, documenting, and reporting events that meet the definition of an AE or SAE as well as other reportable safety events. Investigators remain responsible for following up AEs, SAEs, and other reportable safety events for outcome according to Section 8.4.3.

The investigator, who is a qualified physician, will assess events that meet the definition of an AE or SAE as well as other reportable safety events with respect to seriousness, intensity/toxicity and causality.

8.4.1 Time Period and Frequency for Collecting AE, SAE, and Other Reportable Safety Event Information

All AEs, SAEs, and other reportable safety events that occur after the participant provides documented informed consent but before intervention allocation/randomization must be reported by the investigator if the participant is receiving placebo run-in or other run-in treatment, if the event cause the participant to be excluded from the study, or is the result of a

protocol-specified intervention, including but not limited to washout or discontinuation of usual therapy, diet, or a procedure.

- All AEs from the time of intervention allocation/randomization through 30 days following cessation of study intervention must be reported by the investigator.

- All AEs meeting serious criteria, from the time of intervention allocation/randomization through 90 days following cessation of study intervention or 30 days following cessation of study intervention if the participant initiates new anticancer therapy, whichever is earlier, must be reported by the investigator.

- All pregnancies and exposure during breastfeeding, from the time of treatment allocation/randomization through 120 days following pembrolizumab/placebo or 30 days following cessation of lenvatinib/placebo, whichever occurs last, must be reported by the investigator. If the participant initiates new anticancer therapy following discontinuation of study intervention, the time period for reporting pregnancies and exposure during breastfeeding is reduced to 30 days following cessation of study intervention.

  Exception: A positive pregnancy test at the time of initial screening is not a reportable event unless the participant has received study intervention.

- Additionally, any SAE brought to the attention of an investigator at any time outside the time period specified above must be reported immediately to the Sponsor if the event is considered related to study intervention.

Investigators are not obligated to actively seek AEs or SAEs or other reportable safety events in former study participants. However, if the investigator learns of any SAE, including a death, at any time after a participant has been discharged from the study, and he/she considers the event to be reasonably related to the study intervention or study participation, the investigator must promptly notify the Sponsor.

All initial and follow-up AEs, SAEs, and other reportable safety events will be recorded and reported to the Sponsor or designee within the time frames as indicated in Table 6.

Table 6 Reporting Time Periods and Time Frames for Adverse Events and Other Reportable Safety Events

Type of Event	**Reporting Time Period:** **Consent to Randomization/ Allocation**	**Reporting Time Period:** **Randomization/ Allocation through Protocol-specified Follow-up Period**	**Reporting Time Period:** **After the Protocol-specified Follow-up Period**	**Time Frame to Report Event and Follow-up Information to Sponsor:**
Nonserious Adverse Event (NSAE)	Report if: - due to protocol-specified intervention - causes exclusion - participant is receiving placebo run-in or other run-in treatment	Report all	Not required	Per data entry guidelines
Serious Adverse Event (SAE) including Cancer and Overdose	Report if: - due to protocol-specified intervention - causes exclusion - participant is receiving placebo run-in or other run-in treatment	Report all SAEs, cancer, overdose associated with pembrolizumab and overdose associated with lenvatinib with an AE	Report if: - drug/vaccine related. (Follow ongoing to outcome)	Within 24 hours of learning of event
Pregnancy/ Lactation Exposure	Report if: - participant has been exposed to any protocol-specified intervention (eg, procedure, washout or run-in treatment including placebo run-in) Exception: A positive pregnancy test at the time of initial screening is not a reportable event.	Report all	Previously reported – Follow to completion/termination; report outcome	Within 24 hours of learning of event
Event of Clinical Interest (require regulatory reporting)	Report if: - due to intervention - causes exclusion	Report - potential drug-induced liver injury (DILI) - require regulatory reporting	Not required	Within 24 hours of learning of event
Event of Clinical Interest (do not require regulatory reporting)	Report if: - due to intervention - causes exclusion	Report - non-DILI ECIs and those not requiring regulatory reporting	Not required	Within 5 calendar days of learning of event

8.4.2 Method of Detecting AEs, SAEs, and Other Reportable Safety Events

Care will be taken not to introduce bias when detecting AEs and/or SAEs and other reportable safety events. Open-ended and nonleading verbal questioning of the participant is the preferred method to inquire about AE occurrence.

8.4.3 Follow-up of AE, SAE, and Other Reportable Safety Event Information

After the initial AE/SAE report, the investigator is required to proactively follow each participant at subsequent visits/contacts. All AEs, SAEs, and other reportable safety events including pregnancy and exposure during breastfeeding, events of clinical interest (ECIs), cancer, and overdose will be followed until resolution, stabilization, until the event is otherwise explained, or the participant is lost to follow-up (as defined in Section 7.3). In addition, the investigator will make every attempt to follow all nonserious AEs that occur in randomized participants for outcome. Further information on follow-up procedures is given in Appendix 3.

8.4.4 Regulatory Reporting Requirements for SAE

Prompt notification (within 24 hours) by the investigator to the Sponsor of SAE is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a study intervention under clinical investigation are met.

The Sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study intervention under clinical investigation. The Sponsor will comply with country-specific regulatory requirements and global laws and regulations relating to safety reporting to regulatory authorities, IRB/IECs, and investigators.

Investigator safety reports must be prepared for suspected unexpected serious adverse reactions (SUSARs) according to local regulatory requirements and Sponsor policy and forwarded to investigators as necessary.

An investigator who receives an investigator safety report describing an SAE or other specific safety information (eg, summary or listing of SAE) from the Sponsor will file it along with the IB and will notify the IRB/IEC, if appropriate according to local requirements.

8.4.5 Pregnancy and Exposure During Breastfeeding

Although pregnancy and infant exposure during breastfeeding are not considered AEs, any pregnancy or infant exposure during breastfeeding in a participant (spontaneously reported to the investigator or their designee) that occurs during the study are reportable to the Sponsor.

All reported pregnancies must be followed to the completion/termination of the pregnancy. Pregnancy outcomes of spontaneous abortion, missed abortion, benign hydatidiform mole, blighted ovum, fetal death, intrauterine death, miscarriage, and stillbirth must be reported as serious events (Important Medical Events). If the pregnancy continues to term, the outcome (health of infant) must also be reported.

8.4.6 Disease-related Events and/or Disease-related Outcomes Not Qualifying as AEs or SAEs

Efficacy endpoints as outlined in this section will not be reported to the Sponsor, as described in Section 8.4.1.

Specifically, the suspected/actual events covered in this exception include any event that is progression of the cancer under study.

The Sponsor will monitor unblinded aggregated efficacy endpoint events and safety data to ensure the safety of the participants in the study.

8.4.7 Events of Clinical Interest (ECIs)

Selected nonserious and SAEs are also known as ECIs and must be reported to the Sponsor.

Events of clinical interest for this study include:

1. An overdose of pembrolizumab, as defined in Section 8.5, that is not associated with clinical symptoms or abnormal laboratory results.

2. Any dose over the prescribed dose of lenvatinib associated with an AE.

3. An elevated AST or ALT lab value that is greater than or equal to 3X the upper limit of normal and an elevated total bilirubin lab value that is greater than or equal to 2X the upper limit of normal and, at the same time, an alkaline phosphatase lab value that is less than 2X the upper limit of normal, as determined by way of protocol-specified laboratory testing or unscheduled laboratory testing.*

   *Note: These criteria are based upon available regulatory guidance documents. The purpose of the criteria is to specify a threshold of abnormal hepatic tests that may require an additional evaluation for an underlying etiology. The study site guidance for assessment and follow-up of these criteria can be found in the Investigator Study File Binder (or equivalent).

8.5 Treatment of Overdose

NOTE: After amendment 007-06, lenvatinib and matching placebo were discontinued. This section was left unchanged for reference.

For purposes of this study, an overdose will be defined as any dose exceeding the prescribed dose as follows:

- Pembrolizumab: ≥5 times the protocol-specified dose.

- Lenvatinib: any dose above the protocol-specified dose if associated with an adverse event.

There is no specific antidote for an overdose of lenvatinib. Due to its high degree of plasma protein binding, lenvatinib is not expected to be dialyzable. Adverse reactions in patients receiving single doses of lenvatinib as high as 40 mg were similar to those in clinical studies at the recommended dose for differentiated thyroid cancer and RCC.

No specific information is available on the treatment of overdose of pembrolizumab or lenvatinib. In the event of overdose, the participant should be observed closely for signs of toxicity. Appropriate supportive treatment should be provided if clinically indicated.

All reports of pembrolizumab overdose with and without an AE and all reports of lenvatinib overdose with an AE must be reported by the investigator within 24 hours to the Sponsor either by electronic media or paper.

Reports of pembrolizumab overdose without any associated clinical symptoms or abnormal laboratory results, should be reported using the terminology “accidental or intentional overdose without adverse effect.”

8.6 Pharmacokinetics

As of Protocol Amendment 05, the Sponsor has determined that there is sufficient pembrolizumab/lenvatinib PK data in NSCLC participants. Consequently, no further PK samples for pembrolizumab and lenvatinib or ADA samples for pembrolizumab will be collected.

8.7 Pharmacodynamics

Pharmacodynamic parameters will not be evaluated in this study.

8.8 Planned Genetic Analysis Sample Collection

Samples are to be collected for planned analysis of associations between genetic variants in germline/tumor DNA and drug response. If a documented law or regulation prohibits (or local IRB/IEC does not approve) sample collection for these purposes, then such samples are not be collected at the corresponding sites.

8.9 Biomarkers

To identify novel biomarkers, the following biospecimens to support exploratory analyses of cellular components (eg, protein, RNA, DNA, metabolites) and other circulating molecules will be collected from all participants as specified in the SoA:

- Blood for genetic analysis
- Newly obtained/archival tissue sample for PD-L1 analysis

Sample collection, storage, and shipment instructions for the exploratory biomarker specimens will be provided in the laboratory manual.

8.10 Medical Resource Utilization and Health Economics

All-cause hospitalizations and emergency room visits must be reported in the eCRF, from the time of treatment randomization through 90 days after cessation of study intervention, or 30 days after cessation of study intervention, if the participant initiates new anticancer therapy, whichever is earlier.

8.11 Visit Requirements

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from the study. The following subsections have been updated accordingly.

Visit requirements are outlined in the SoA (Section 1.3). Specific procedure-related details are provided in Section 8.

Unscheduled visits are permitted at any time during the course of the study.

8.11.1 Screening

Documented informed consent must be provided before performing any protocol-specific procedure. Results of a test performed before the participant signs the ICF as part of routine clinical management are acceptable in lieu of a screening test if performed within the specified time frame. Screening procedures are to be completed within 28 days before the first dose of study intervention except for the following:

- Laboratory tests are to be performed per the SoA (Section 1.3).
- Evaluation of ECOG is to be performed at screening, within 7 days before Cycle 1 but before randomization.
- For WOCBP, a urine or serum pregnancy test will be performed within 24 hours before the first dose of study intervention. If urine pregnancy results cannot be confirmed as negative, a serum pregnancy test will be required (performed by the local study site laboratory).

Participants may be rescreened after initially failing to meet the inclusion/exclusion criteria. Results from assessments during the initial screening period are acceptable in lieu of a repeat screening test if performed within the specified time frame and the corresponding inclusion/exclusion criteria are met. Participants who are rescreened will retain their original screening number.

8.11.2 Initial Treatment Phase

Visit requirements are outlined in the SoA (Section 1.3). Specific procedure-related details are provided in Section 8.1.

8.11.3 Second-course Treatment Phase

NOTE: As of Amendment 007-06, the Second-course Treatment Phase has been removed from the study.

8.11.4 Discontinued Participants Continuing to be Monitored in the Study

The Discontinuation Visit should occur at the time study intervention is discontinued for any reason. If the Discontinuation Visit occurs 30 days from the last dose of study intervention, at the time of the mandatory Safety Follow-up Visit, the Discontinuation Visit procedures and any additional safety follow-up procedures should be performed.

Visit requirements are outlined in the SoA. Additional details regarding participant withdrawal and discontinuation are presented in Section 8.1.9.

8.11.5 Poststudy

8.11.5.1 Safety Follow-up Visit

The mandatory Safety Follow-up Visits should be conducted approximately 30 days after the last dose of study intervention or before the initiation of a new anticancer treatment, whichever comes first.

8.11.5.2 Efficacy Follow-up Visits

NOTE: After amendment 007-06, imaging obtained on study will be submitted to the imaging vendor but these will not be assessed, verification of PD is no longer needed before treatment discontinuation, and posttreatment imaging is no longer required.

Participants who complete the protocol-required cycles of study intervention or who discontinue study intervention for a reason other than PD will move into Efficacy Follow-up Phase and should be assessed as outlined in the SoA (Section 1.3) to monitor disease status. Every effort should be made to collect information regarding disease status until the start of new anticancer therapy, PD, death, or end of study. Information regarding poststudy anticancer treatment will be collected if new treatment is initiated. Participants who completed all efficacy assessments and/or will not have further efficacy assessments must enter Survival Follow-up Phase.

8.11.5.3 Survival Follow-up Assessments

Participants who experience confirmed PD or start a new anticancer therapy will move into the Survival Follow-Up Phase and should be contacted by telephone approximately every 12 weeks (or more often as required) to assess survival status until death, withdrawal of consent, or the end of the study, whichever occurs first. Participants may be contacted for survival status at any time during the course of the study.

9 STATISTICAL ANALYSIS PLAN

NOTE: Data from an interim safety and futility eDMC for LEAP-007 (data cutoff: 19-MAY-2021) indicated that the study met the prespecified nonbinding futility criteria for OS for the combination of lenvatinib plus pembrolizumab compared with placebo plus pembrolizumab. This futility analysis was requested at a previous eDMC meeting and conducted as described in the sSAP. Based upon these data and the recommendation of the eDMC, Amendment 007-06 was implemented to unblind the study and remove lenvatinib and matching placebo from the treatment arms. The prespecified interim and final analyses of the study described in the statistical analysis plan (SAP) will not be performed. Selected analyses of safety and ePRO endpoints will be performed at the end of the study; there will be no further analyses of efficacy endpoints.

This section outlines the statistical analysis strategy and procedures for the study. **As of Amendment 007-06, the study has been unblinded.** Changes made to primary and/or key secondary hypotheses, or the statistical methods related to those hypotheses, that occurred prior to study unblinding were documented in previous protocol amendment(s) (consistent with International Council for Harmonisation [ICH] of Technical Requirements for Pharmaceuticals for Human Use Guideline E9). Changes to exploratory or other nonconfirmatory analyses made after the protocol has been finalized, but before unblinding/final database lock, will be documented in a supplemental statistical analysis plan (sSAP) and referenced in the clinical study report (CSR) for the study. Post hoc exploratory analyses will be clearly identified in the CSR. Separate analysis plans (ie, separate documents from the sSAP) will be developed to detail PK and biomarker analyses. The ePRO analysis plan will be included in the sSAP.

9.1 Statistical Analysis Plan Summary

Key elements of the SAP are summarized below. The comprehensive plan is provided in Sections 9.2 through 9.12. **As of Amendment 007-06, the prespecified interim and final analyses of the study described in SAP will not be performed. Selected analyses of safety and ePRO endpoints will be performed at the end of the study; there will be no further analyses of efficacy endpoints. The SAP summary has been updated accordingly.**

Study Design Overview	Phase 3 study of pembrolizumab plus lenvatinib vs pembrolizumab plus placebo for 1L treatment of metastatic NSCLC in participants whose tumors express PD-L1 (TPS ≥1%)
Treatment Assignment	Approximately 620 participants will be randomized in a 1:1 ratio between 2 treatment arms: (1) pembrolizumab + lenvatinib and (2) pembrolizumab + placebo. Stratification factors are as follows: • Geographic region (East Asia vs non-East Asia) • ECOG PS (0 vs 1) • TPS (1%-49% vs ≥50%) **As of Amendment 007-06, the study is unblinded and all ongoing participants will continue treatment with open-label pembrolizumab monotherapy.**

Analysis Populations	Efficacy: Intention-to-Treat (ITT) Safety: All Participants as Treated (APaT)
Primary Endpoints	• PFS per RECIST 1.1 based on BICR • OS
Secondary Endpoints	• OR per RECIST 1.1 based on BICR • Safety and tolerability • Change from baseline in Global health status/QoL, cough, chest pain, dyspnea and physical functioning scores • Time to True Deterioration (TTD) in global health status/QoL, cough, chest pain, dyspnea, and physical functioning • TTD in the composite endpoint of cough, chest pain, or dyspnea items
Statistical Methods for Key Efficacy Analyses	**As of Amendment 007-06, the prespecified interim and final analyses of the study will not be performed. There will be no further analyses of efficacy endpoints.** The primary hypotheses will be evaluated by comparing pembrolizumab + lenvatinib to pembrolizumab + placebo in PFS and OS using a stratified log-rank test and in ORR using the stratified Miettinen and Nurminen method [Miettinen, O. 1985]. The HR will be estimated using a stratified Cox regression model. Event rates over time will be estimated within each treatment group using the Kaplan-Meier method. The difference in ORR will be estimated using the stratified Miettinen and Nurminen method with strata weighting by sample size.
Statistical Methods for Key Safety Analyses	The analysis of safety results will follow a tiered approach. The tiers differ with respect to the analyses that will be performed. There are no events of interest that warrant elevation to Tier 1 events in this study. Tier 2 parameters will be assessed via point estimates with 95% CIs provided for between-group comparisons; only point estimates by treatment group are provided for Tier 3 safety parameters. The 95% CIs for the between-treatment differences in percentages will be provided using the Miettinen and Nurminen method.
Interim Analyses	**As of Amendment 007-06, no interim analyses of the study will be performed.**
Multiplicity	**As of Amendment 007-06, no interim or final analyses of the study will be performed.**
Sample Size and Power	**As of Amendment 007-06, no interim or final analyses of the study will be performed.** The planned sample size is approximately 620 participants. For PFS, based on 416 events, the study has 86.5% power to detect a HR of 0.7 (pembrolizumab + lenvatinib vs pembrolizumab + placebo) at α=0.55% (1-sided). For OS, based on 388 events, the study has 90% power to detect a HR of 0.71 (pembrolizumab + lenvatinib vs pembrolizumab + placebo) at α=1.95% (1-sided).
China Extension Study	**As of Amendment 007-06, enrollment in the Extension Study in China has been stopped.** China participants randomized during the global study phase will be included in all global study analyses (efficacy and safety). China participants randomized during the China extension phase will be excluded from all global study analyses. China participants randomized during global and extension phases will both be included in any China-specific analyses.

9.2 Responsibility for Analyses/In-house Blinding

Note: As of Amendment 007-06, the study is unblinded to the Sponsor, the investigational sites, and the participants. This section has been updated accordingly.

The statistical analysis of the data obtained from this study will be the responsibility of the Clinical Biostatistics department of the Sponsor.

This study was conducted as a double-blind study under in-house blinding procedures until after the eDMC recommended study unblinding.

The Sponsor generated the randomized allocation schedule(s) for study intervention assignment. Blinding issues related to the planned interim analyses are described in Section 9.7, but are no longer applicable as of Amendment 007-06.

Extension Study in China

For all participants in China, including participants randomized in the global study and the extension study, statistician(s)/programmers were blinded to participant level treatment randomization and this information had been clearly documented.

9.3 Hypotheses/Estimation

Objectives and hypotheses of the study are stated in Section 3.

9.4 Analysis Endpoints

Efficacy and safety endpoints that will be evaluated for within- and/or between-treatment differences are listed below.

9.4.1 Efficacy Endpoints

9.4.1.1 Primary Endpoints

***PFS*:** the time from randomization to the first documented PD per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR or death due to any cause, whichever occurs first. See Section 9.6.1 for the definition of censoring.

***OS*:** the time from randomization to death due to any cause.

9.4.1.2 Secondary Endpoints

***OR*:** a confirmed CR or PR per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR.

9.4.1.3 Exploratory Endpoints

***DOR*:** the time from the earliest date of qualifying response until earliest date of PD or death from any cause, whichever comes first.

9.4.2 Safety Endpoints

Safety and tolerability will be assessed by clinical review of all relevant parameters including AEs, SAEs, fatal AEs, laboratory tests, and vital signs. Furthermore, specific events will be collected and designated as ECIs as described in Section 8.4.7.

9.4.3 Patient-reported Outcome Endpoints

The following secondary PRO endpoints will be evaluated as described in Section 4.2.1.3 and analyzed as described in Section 9.6.3:

Change from baseline in

- Global Health Status/QoL scale (QLQ-C30 items 29-30)
- Single-item symptom scales: cough (QLQ-LC13 item 31), chest pain (QLQ-LC13 item 40), and dyspnea (QLQ-C30 item 8)
- Physical functioning scale (QLQ-C30 items 1-5)

Time to true deterioration (TTD) in

- Global Health Status/QoL scale (QLQ-C30 items 29-30)
- Single-item symptom scales: cough (QLQ-LC13 item 31), chest pain (QLQ-LC13 item 40), and dyspnea (QLQ-C30 item 8)
- Physical functioning scale (QLQ-C30 items 1-5)
- Composite symptom endpoint: cough (QLQ-LC13 item 31), chest pain (QLQ-LC13 item 40), or dyspnea (QLQ-C30 item 8)

The TTD in global health status/QoL, cough, chest pain, dyspnea, and physical functioning is defined as the time from baseline to the first onset of a 10 or more points deterioration from baseline with confirmation by the subsequent visit of a 10 or more points deterioration from baseline. The TTD in the composite endpoint of cough, chest pain, or dyspnea is defined as the time to first onset of 10 or more points deterioration from baseline in any one of 3 scale items with confirmation by the subsequent visit of 10 points or more deterioration from baseline in the same scale as the first onset. The EQ-5D-5L will be evaluated as an exploratory endpoint. These analyses and other supportive PRO analyses will be described in the sSAP.

9.5 Analysis Populations

9.5.1 Efficacy Analysis Populations

The analyses of the primary efficacy endpoints are based on the ITT population. All randomized participants will be included in this population. Participants will be analyzed in the treatment group to which they are randomized. Details on the approach to handling missing data are provided in Section 9.6.

Extension Study in China

NOTE: As of Amendment 007-06, enrollment in the Extension Study in China has been stopped. This section has been updated accordingly.

The participants in China who were randomized in the extension study will not be included in the above global study primary efficacy analysis population. The ITT participants in China, including all participants in China randomized in the global study and the extension study, will be analyzed.

9.5.2 Safety Analysis Populations

Safety analyses will be conducted in the APaT population, which consists of all randomized participants who received at least 1 dose of study intervention. Participants will be included in the treatment group corresponding to the study intervention they actually received for the analysis of safety data using the APaT population. This will be the treatment group to which they were randomized, except for participants who take incorrect study intervention for the entire treatment period; such participants will be included in the treatment group corresponding to the study intervention actually received. Any participant who receives the incorrect study intervention for 1 cycle, but receives the randomized treatment for all other cycles, will be analyzed according to the randomized treatment group, and a narrative will be provided for any events that occur during the cycle for which the participant is incorrectly dosed.

At least 1 laboratory, vital sign, or ECG measurement obtained after at least 1 dose of study intervention is required for inclusion in the analysis of the respective safety parameter. To assess change from baseline, a baseline measurement is also required.

Extension Study in China

The participants in China who were randomized and treated in the extension study will not be included in the above global study primary safety analysis population. The APaT participants in China, including all randomized participants in China (in the global study and extension study) who received at least 1 dose of study treatment, will be analyzed.

9.5.3 Patient-reported Outcome Analysis Population

The analyses of PRO endpoints will be based on the PRO full analysis set (FAS) population following the ITT principle and ICH E9 guidelines. The PRO FAS population consists of all

randomized participants who have received at least 1 dose of study intervention and have completed at least 1 PRO assessment.

9.5.4 Pharmacokinetic Analysis Population

The population of PK analysis set includes all the participants who have received at least 1 dose of study intervention with documented dosing history in the lenvatinib + pembrolizumab arm and have measurable plasma levels of lenvatinib or serum levels of pembrolizumab.

9.6 Statistical Methods

NOTE: As of Amendment 007-06, the prespecified interim and final analyses of the study will not be performed. Selected analyses of safety and ePRO endpoints will be performed at end of study; there will be no further analyses of efficacy endpoints. The subsections below are retained for reference.

9.6.1 Statistical Methods for Efficacy Analyses

This section describes the statistical methods that address the primary and secondary efficacy objectives. Methods related to exploratory objectives will be described in the sSAP. Efficacy results that will be deemed to be statistically significant after consideration of the Type I error control strategy are described in Section 9.8. Nominal *p* values will be computed for other efficacy analyses, but should be interpreted with caution because of potential issues of multiplicity.

The stratification factors used for randomization (see Section 6.3.2) will be applied to all stratified analyses, in particular, the stratified log-rank test, stratified Cox model, and stratified Miettinen and Nurminen method [Miettinen, O. 1985]. If there are small strata, for the purpose of analysis, strata will be combined to ensure sufficient number of participants, responses and events in each stratum. Details regarding the pooling strategy will be prespecified in the sSAP prior to the database lock for the first analysis when each applicable endpoint will be analyzed, and decisions regarding the pooling will be based on a blinded review of response and event counts by stratum.

The efficacy analyses for OR, DOR and PFS will include responses and documented progression events that occur prior to Second-course treatment.

A summary of the primary analysis strategy for the key efficacy endpoints is provided in Table 7.

Table 7 Analysis Methods for Key Efficacy Endpoints

Endpoint/Variable	Statistical Method	Analysis Population	Missing Data Approach
Primary Analyses:			
PFS (RECIST 1.1) by BICR	Testing: stratified log-rank test Estimation: Stratified Cox model with Efron's tie-handling method	ITT	Censored according to rules in Table 8
OS	Testing: stratified log-rank test Estimation: Stratified Cox model with Efron's tie-handling method	ITT	Censored at last known alive date
Secondary Analyses:			
ORR (RECIST 1.1) by BICR	Testing and estimation: stratified Miettinen and Nurminen method	ITT	Participants with missing data are considered nonresponders
Abbreviations: BICR = blinded independent central review; ITT = intention-to-treat; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors.			

The strategy to address multiplicity issues with regard to multiple endpoints and interim analyses is described in Section 9.7, Interim Analyses, and in Section 9.8, Multiplicity.

9.6.1.1 Progression-free Survival

The nonparametric Kaplan-Meier method will be used to estimate the PFS curve in each treatment group. The hypotheses of treatment difference in PFS will be tested by the stratified log-rank test. A stratified Cox proportional hazard model with Efron's method of tie handling will be used to estimate the magnitude of the treatment difference (ie, HR) between the treatment arms. The HR and its 95% CI from the stratified Cox model with Efron's method of tie handling and with a single treatment covariate will be reported.

Since PD is assessed periodically, PD can occur any time in the time interval between the last assessment where PD was not documented and the assessment when PD is documented. For the primary analysis, for the participants who have PD, the true date of PD will be approximated by the date of the first assessment at which PD is objectively documented per RECIST 1.1 by the BICR vendor, regardless of discontinuation of study drug. Additional analyses will be performed for PFS per iRECIST.

To evaluate the robustness of the PFS endpoint per RECIST 1.1 via BICR by the imaging vendor, 1 primary and 2 sensitivity analyses with a different set of censoring rules will be performed. For the primary analysis, if the events (PD or death) are after more than 1 missed disease assessment, the data are censored at the last disease assessment before missing visits. Also, data after new anticancer therapy are censored at the last disease assessment before the initiation of new anticancer therapy. The first sensitivity analysis follows ITT principles (ie,

PDs/deaths are counted as events regardless of missed study visits or initiation of new anticancer therapy). The second sensitivity analysis considers discontinuation of treatment or initiation of an anticancer treatment after discontinuation of study-specified treatments, whichever occurs later, to be a PD event for participants without documented PD or death. If a participant meets multiple criteria for censoring, the censoring criterion that occurs earliest will be applied. The censoring rules for primary and sensitivity analyses are summarized in Table 8.

Table 8 Censoring Rules for Primary and Sensitivity Analyses of Progression-free Survival

Situation	Primary Analysis	Sensitivity Analysis 1	Sensitivity Analysis 2
No PD and no death; new anticancer treatment is not initiated	Censored at last disease assessment	Censored at last disease assessment	Progressed at treatment discontinuation for reasons other than CR; otherwise, censored at last disease assessment if still on study intervention or completed study intervention
No PD and no death; new anticancer treatment is initiated	Censored at last disease assessment before new anticancer treatment	Censored at last disease assessment	Progressed at date of new anticancer treatment
PD or death documented after ≤1 missed disease assessment and before new anticancer therapy	Progressed at date of documented PD or death	Progressed at date of documented PD or death	Progressed at date of documented PD or death
PD or death documented immediately after ≥2 consecutive missed disease assessments or after new anticancer therapy	Censored at last disease assessment before the earlier date of ≥2 consecutive missed disease assessment and new anticancer therapy	Progressed at date of documented PD or death	Progressed at date of documented PD or death
Abbreviation: CR = complete response; PD = progressive disease.			

9.6.1.2 Overall Survival

The nonparametric Kaplan-Meier method will be used to estimate the survival curves. The treatment difference in survival will be assessed by the stratified log-rank test (based on the stratification factors defined in Section 6.3.2). A stratified Cox proportional hazard model with Efron's method of tie handling will be used to assess the magnitude of the treatment difference (ie, the HR). The HR and its 95% CI from the stratified Cox model with a single treatment covariate will be reported. The stratification factors used for randomization (Section 6.3.2) will be applied to both the stratified log-rank test and the stratified Cox model. Participants without documented death at the time of analysis will be censored at the

date of last known contact. The analysis using the restricted mean survival time method may be conducted for OS to account for the possible nonproportional hazards effect.

9.6.1.3 Objective Response Rate

The stratified Miettinen and Nurminen method will be used for comparison of ORR between the treatment groups. The difference in ORR and its 95% CI from the stratified Miettinen and Nurminen method with strata weighting by sample size will be provided.

9.6.1.4 Duration of Response

For participants who demonstrate CR or PR, DOR is defined as the time from first documented evidence of CR or PR until PD or death due to any cause, whichever occurs first.

The nonparametric Kaplan-Meier method will be used to summarize the DOR. The median and range of DOR will be provided.

Censoring rules for DOR are summarized in Table 9.

Table 9 Censoring Rules for Duration of Response

Situation	Date of Progression or Censoring	Outcome
No PD or death, no new anticancer therapy initiated	Last adequate disease assessment	Censor (nonevent)
No PD or death, new anticancer therapy initiated	Last adequate disease assessment before new anticancer therapy initiated	Censor (nonevent)
Death or PD immediately after ≥2 consecutive missed disease assessments or after new anticancer therapy, if any	Earlier date of last adequate disease assessment before ≥2 missed adequate disease assessments and new anticancer therapy, if any	Censor (nonevent)
Death or PD after ≤1 missed disease assessments and before new anticancer therapy, if any	PD or death	End of response (event)

Abbreviation: PD = progressive disease.
Note: A missed disease assessment includes any assessment that is not obtained or is considered inadequate for evaluation of response.

9.6.2 Statistical Methods for Safety Analyses

Safety and tolerability will be assessed by clinical review of all relevant parameters including AEs, laboratory tests, vital signs, and ECG measurements.

The analysis of safety results will follow a tiered approach (Table 10). The tiers differ with respect to the analyses that will be performed. AEs (specific terms as well as system organ

class terms) and events that meet predefined limits of change in laboratory, vital signs, and ECG parameters are either prespecified as "Tier 1" endpoints or will be classified as belonging to "Tier 2" or "Tier 3," based on observed proportions of participants with an event.

Tier 1 Events

Safety parameters or AEs of special interest that are identified a priori constitute "Tier 1" safety endpoints that will be subject to inferential testing for statistical significance. AEs that are immune-mediated or potentially immune-mediated are well documented and will be evaluated separately; however, these events have been characterized consistently throughout the pembrolizumab clinical development program, and determination of statistical significance is not expected to add value to the safety evaluation. Similarly, the combination of pembrolizumab and lenvatinib has not been associated with any new safety signals. Finally, there are no known AEs associated with participants with NSCLC for which determination of a *p* value is expected to impact the safety assessment. Therefore, there are no Tier 1 events for this protocol.

Tier 2 Events

Tier 2 parameters will be assessed via point estimates with 95% CIs provided for differences in the proportion of participants with events using the Miettinen and Nurminen method, an unconditional, asymptotic method [Miettinen, O. 1985].

Membership in Tier 2 requires that at least 10% of participants in any treatment group exhibit the event; all other AEs and predefined limits of change will belong to Tier 3. The threshold of at least 10% of participants was chosen for Tier 2 events because the population randomized in this study are in critical conditions and usually experience various AEs of similar types regardless of treatment; events reported less frequently than 10% of participants would obscure the assessment of the overall safety profile and add little to the interpretation of potentially meaningful treatment differences. In addition, Grade 3 to 5 AEs (≥5% of participants in one of the treatment groups) and SAEs (≥5% of participants in one of the treatment groups) will be considered Tier 2 endpoints. Because many 95% CIs may be provided without adjustment for multiplicity, the CIs should be regarded as a helpful descriptive measure to be used in safety review, not a formal method for assessing the statistical significance of the between-group differences.

Tier 3 Events

Safety endpoints that are not Tier 1 or 2 events are considered Tier 3 events. Only point estimates by treatment group are provided for Tier 3 safety parameters.

Continuous Safety Measures

For continuous measures such as changes from baseline in laboratory, vital signs, and ECG parameters, summary statistics for baseline, on-treatment, and change from baseline values will be provided by treatment group in table format.

Table 10 Analysis Strategy for Safety Parameters

Safety Tier	Safety Endpoint	95% CI for Treatment Comparison	Descriptive Statistics
Tier 2	Any AE (≥10% of participants in one of the treatment groups)	X	X
	Any Grade 3 to 5 AE (≥5% of participants in one of the treatment groups)	X	X
	Any serious AE (≥5% of participants in one of the treatment groups)	X	X
Tier 3	Any AE		X
	Change from baseline results (laboratory test toxicity grade)		X
Abbreviations: AE = adverse event; CI = confidence interval; X = results will be provided.			

9.6.3 Statistical Methods for Patient-reported Outcome Analyses

Mean change from baseline

The time point for the mean change from baseline will be determined based on blinded data review prior to the database lock for any PRO analysis and documented in the sSAP.

To assess the treatment effects on the PRO score change from baseline in the PRO endpoints defined in Section 9.4.3, a constrained longitudinal data analysis (cLDA) model proposed by Liang and Zeger [Liang, Kung-Yee and Zeger, Scott L. 2000] will be applied, with the PRO score as the response variable, and treatment, time, the treatment by time interaction, and stratification factors used for randomization (See Section 6.3.2) as covariates. The treatment difference in terms of least square (LS) mean change from baseline will be estimated from this model together with 95% CI. Model-based LS mean with 95% CI will be provided by treatment group for PRO scores at baseline and postbaseline time point.

Time to Deterioration (TTD)

For the TTD endpoints defined in Section 9.4.3, the Kaplan-Meier method will be used to estimate the TTD curve for each treatment group. The estimate of median time to deterioration and its 95% CI will be obtained from the Kaplan-Meier estimates. The treatment difference in TTD will be assessed by the stratified log-rank test. A stratified Cox proportional hazard model with Efron's method of tie handling and with a single treatment covariate will be used to assess the magnitude of the treatment difference (ie, HR). The HR and its 95% CI will be reported. The same stratification factors used for the stratified PFS and OS analyses will be used as the stratification factors in both the stratified log-rank test and the stratified Cox model.

9.6.4 Demographic and Baseline Characteristics

The comparability of the treatment groups for each relevant demographic and baseline characteristic will be assessed by the use of tables and/or graphs. No statistical hypothesis tests will be performed on these characteristics. The number and percentage of participants screened and randomized and the primary reason for screening failure and discontinuation will be displayed. Demographic variables, baseline characteristics, primary and secondary diagnoses, and prior and concomitant therapies will be summarized by treatment, either by descriptive statistics or categorical tables.

9.7 Interim Analyses

NOTE: As of Amendment 007-06, no interim analyses of the study will be performed. This section is no longer applicable.

9.7.1 Safety Interim Analyses

The DMC conducted regular safety monitoring. The timing of the safety monitoring was specified in the DMC charter. No further analysis is warranted.

9.8 Multiplicity

NOTE: As of Amendment 007-06, no interim or final analyses of the study will be performed. This section is no longer applicable.

9.9 Sample Size and Power Calculations

NOTE: As of Amendment 007-06, no interim or final analyses of the study will be performed. This section is retained for reference.

The study will randomize 620 participants in a 1:1 ratio into the pembrolizumab + lenvatinib and pembrolizumab + placebo arms. PFS and OS are primary endpoints for the study, with OR as the secondary endpoint.

For the PFS endpoint, based on a target number of 416 events at the final PFS analysis, the study has approximately 86.5% power to detect an HR of 0.7 at the initially allocated α=0.0055 (1-sided).

For the OS endpoint, based on a target number of 388 events and one interim analysis at approximately 78% of the target number of events, the study has approximately 90% power to detect an HR of 0.71 at the initially allocated α=0.0195 (1-sided).

Based on KEYNOTE-042 data, the above sample size and power calculations for PFS and OS assume the following:

PFS follows a piecewise exponential distribution, with a median of 5 months up to 6.5 months and then a median of 12 months thereafter for the control group.

OS follows an exponential distribution, with a median of 16.7 months for the control group.

Enrollment period is approximately 21.7 months.

Annual drop-out rate is 20% and 2% for PFS and OS, respectively.

Follow-up period is approximately 8.8 and 18.5 months for PFS and OS, respectively, after the last participant is randomized.

The sample size and power calculations were performed in R (package "gsDesign") and EAST 6.4.

Extension Study in China

As of amendment 007-06, enrollment in the Extension Study in China has been stopped.

In order to evaluate the consistency of efficacy and safety in the subpopulation in China compared with the global population, following completion of global study enrollment, participants in China will continue to be randomized in a 1:1 ratio into the pembrolizumab + lenvatinib arm and pembrolizumab + placebo arm until the planned sample size of approximately 120 participants in China is reached. Participants in China randomized after completion of enrollment in the global study will not be included in the analysis of the global study.

9.10 Subgroup Analyses

To determine whether the treatment effect is consistent across various subgroups, the between-group treatment effect for OS, PFS, and OR (with a nominal 95% CI) will be estimated and plotted by treatment group within each category of the following subgroup variables:

- Geographic region (East Asia, non-East Asia)
- ECOG performance status (0, 1)
- Predominant tumor histology (squamous, nonsquamous)
- TPS (1%-49%, ≥50%)
- Age category (<65, ≥65 years)
- Sex (female, male)
- Race (white, nonwhite)
- Smoking status (never, former/current smoker)

- Brain metastasis at baseline (presence, absence)
- Liver metastasis at baseline (presence, absence)

The consistency of the treatment effect will be assessed using descriptive statistics for each category of the subgroup variables listed above. If the number of participants in a category of a subgroup variable is less than 10% of the ITT population, the subgroup analysis will not be performed for this category of the subgroup variable, and this subgroup variable will not be displayed in the forest plot. The subgroup analyses for PFS and OS will be conducted using unstratified Cox model, and the subgroup analyses for OR will be conducted using unstratified Miettinen and Nurminen method.

9.11 Compliance (Medication Adherence)

Drug accountability data for study intervention will be collected during the study. Any deviation from protocol-directed administration will be reported.

9.12 Extent of Exposure

Extent of exposure for a participant is defined as number of cycles in which the participant receives the study intervention. Summary statistics will be provided on extent of exposure for the APaT population.

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 Appendix 1: Regulatory, Ethical, and Study Oversight Considerations

10.1.1 Code of Conduct for Clinical Trials

Merck Sharp and Dohme Corp., a subsidiary of Merck & Co., Inc. (MSD)

Code of Conduct for Interventional Clinical Trials

I. Introduction

A. Purpose

MSD, through its subsidiaries, conducts clinical trials worldwide to evaluate the safety and effectiveness of our products. As such, we are committed to designing, implementing, conducting, analyzing, and reporting these trials in compliance with the highest ethical and scientific standards. Protection of participants in clinical trials is the overriding concern in the design and conduct of clinical trials. In all cases, MSD clinical trials will be conducted in compliance with local and/or national regulations (including all applicable data protection laws and regulations), and International Council for Harmonisation Good Clinical Practice (ICH-GCP), and also in accordance with the ethical principles that have their origin in the Declaration of Helsinki.

B. Scope

Highest ethical and scientific standards shall be endorsed for all clinical interventional investigations sponsored by MSD irrespective of the party (parties) employed for their execution (eg, contract research organizations, collaborative research efforts). This Code is not intended to apply to trials that are observational in nature, or which are retrospective. Further, this Code does not apply to investigator-initiated trials, which are not under the full control of MSD.

II. Scientific Issues

A. Trial Conduct

1. Trial Design

Except for pilot or estimation trials, clinical trial protocols will be hypothesis-driven to assess safety, efficacy, and/or pharmacokinetic or pharmacodynamic indices of MSD or comparator products. Alternatively, MSD may conduct outcomes research trials, trials to assess or validate various endpoint measures, or trials to determine patient preferences, etc.

The design (ie, participant population, duration, statistical power) must be adequate to address the specific purpose of the trial and shall respect the data protection rights of all participants, trial site staff and, where applicable, third parties. Participants must meet protocol entry criteria to be enrolled in the trial.

2. Site Selection

MSD selects investigative sites based on medical expertise, access to appropriate participants, adequacy of facilities and staff, previous performance in clinical trials, as well as budgetary considerations. Prior to trial initiation, sites are evaluated by MSD personnel to assess the ability to successfully conduct the trial.

3. Site Monitoring/Scientific Integrity

Investigative trial sites are monitored to assess compliance with the trial protocol and Good Clinical Practice (GCP). MSD reviews clinical data for accuracy, completeness, and consistency. Data are verified versus source documentation according to standard operating procedures. Per MSD policies and procedures, if

potential fraud, scientific/research misconduct, privacy incidents/breaches or Clinical Trial-related Significant Quality Issues are reported, such matters are investigated. When necessary, appropriate corrective and/or preventative actions are defined and regulatory authorities and/or ethics review committees are notified.

B. Publication and Authorship

Regardless of trial outcome, MSD commits to publish primary and secondary results of its registered trials of marketed products in which treatment is assigned, according to the prespecified plans for data analysis. To the extent scientifically appropriate, MSD seeks to publish the results of other analyses it conducts that are important to patients, physicians, and payers. Some early phase or pilot trials are intended to be hypothesis-generating rather than hypothesis testing, in such cases, publication of results may not be appropriate since the trial may be underpowered and the analyses complicated by statistical issues such as multiplicity.

MSD's policy on authorship is consistent with the recommendations published by the International Committee of Medical Journal Editors (ICMJE). In summary, authorship should reflect significant contribution to the design and conduct of the trial, performance or interpretation of the analysis, and/or writing of the manuscript. All named authors must be able to defend the trial results and conclusions. MSD funding of a trial will be acknowledged in publications.

III. Participant Protection

A. Ethics Committee Review (Institutional Review Board [IRB]/Independent Ethics Committee [IEC])

All clinical trials will be reviewed and approved by an IRB/IEC before being initiated at each site. Significant changes or revisions to the protocol will be approved by the ethics committee prior to implementation, except changes required urgently to protect participant safety that may be enacted in anticipation of ethics committee approval. For each site, the ethics committee and MSD will approve the participant informed consent form.

B. Safety

The guiding principle in decision-making in clinical trials is that participant welfare is of primary importance. Potential participants will be informed of the risks and benefits of, as well as alternatives to, trial participation. At a minimum, trial designs will take into account the local standard of care.

All participation in MSD clinical trials is voluntary. Participants enter the trial only after informed consent is obtained. Participants may withdraw from an MSD trial at any time, without any influence on their access to, or receipt of, medical care that may otherwise be available to them.

C. Confidentiality

MSD is committed to safeguarding participant confidentiality, to the greatest extent possible, as well as all applicable data protection rights. Unless required by law, only the investigator, Sponsor (or individuals acting on behalf of MSD), ethics committee, and/or regulatory authorities will have access to confidential medical records that might identify the participant by name.

D. Genomic Research

Genomic research will only be conducted in accordance with a protocol and informed consent authorized by an ethics committee.

IV. Financial Considerations

A. Payments to Investigators

Clinical trials are time- and labor-intensive. It is MSD's policy to compensate investigators (or the sponsoring institution) in a fair manner for the work performed in support of MSD trials. MSD does not pay incentives to enroll participants in its trials. However, when enrollment is particularly challenging, additional payments may be made to compensate for the time spent in extra recruiting efforts.

MSD does not pay for participant referrals. However, MSD may compensate referring physicians for time spent on chart review to identify potentially eligible participants.

B. Clinical Research Funding

Informed consent forms will disclose that the trial is sponsored by MSD and that the investigator or sponsoring institution is being paid or provided a grant for performing the trial. However, the local ethics committee may wish to alter the wording of the disclosure statement to be consistent with financial practices at that institution. As noted above, all publications resulting from MSD trials will indicate MSD as a source of funding.

C. Funding for Travel and Other Requests

Funding of travel by investigators and support staff (eg, to scientific meetings, investigator meetings, etc.) will be consistent with local guidelines and practices.

V. Investigator Commitment

Investigators will be expected to review MSD's Code of Conduct as an appendix to the trial protocol, and in signing the protocol, agree to support these ethical and scientific standards.

10.1.2 Financial Disclosure

Financial Disclosure requirements are outlined in the US Food and Drug Administration Regulations, Financial Disclosure by Clinical Investigators (21 CFR Part 54). It is the Sponsor's responsibility to determine, based on these regulations, whether a request for Financial Disclosure information is required. It is the investigator's/subinvestigator's responsibility to comply with any such request.

The investigator/subinvestigator(s) agree, if requested by the Sponsor in accordance with 21 CFR Part 54, to provide his/her financial interests in and/or arrangements with the Sponsor to allow for the submission of complete and accurate certification and disclosure statements. The investigator/subinvestigator(s) further agree to provide this information on a Certification/Disclosure Form, commonly known as a financial disclosure form, provided by the Sponsor. The investigator/subinvestigator(s) also consent to the transmission of this information to the Sponsor in the United States for these purposes. This may involve the transmission of information to countries that do not have laws protecting personal data.

10.1.3 Data Protection

Participants will be assigned a unique identifier by the Sponsor. Any participant records or datasets that are transferred to the Sponsor will contain the identifier only; participant names or any information that would make the participant identifiable will not be transferred.

The participant must be informed that his/her personal study-related data will be used by the Sponsor in accordance with local data protection law. The level of disclosure must also be explained to the participant.

The participant must be informed that his/her medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the Sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

10.1.3.1 Confidentiality of Data

By signing this protocol, the investigator affirms to the Sponsor that information furnished to the investigator by the Sponsor will be maintained in confidence, and such information will be divulged to the IRB, IEC, or similar or expert committee; affiliated institution and employees, only under an appropriate understanding of confidentiality with such board or committee, affiliated institution and employees. Data generated by this study will be considered confidential by the investigator, except to the extent that it is included in a publication as provided in the Publications section of this protocol.

10.1.3.2 Confidentiality of Participant Records

By signing this protocol, the investigator agrees that the Sponsor (or Sponsor representative), IRB/IEC, or regulatory authority representatives may consult and/or copy study documents to verify worksheet/CRF data. By signing the consent form, the participant agrees to this process. If study documents will be photocopied during the process of verifying worksheet/CRF information, the participant will be identified by unique code only; full names/initials will be masked prior to transmission to the Sponsor.

By signing this protocol, the investigator agrees to treat all participant data used and disclosed in connection with this study in accordance with all applicable privacy laws, rules and regulations.

10.1.3.3 Confidentiality of IRB/IEC Information

The Sponsor is required to record the name and address of each IRB/IEC that reviews and approves this study. The Sponsor is also required to document that each IRB/IEC meets regulatory and ICH GCP requirements by requesting and maintaining records of the names and qualifications of the IRB/IEC members and to make these records available for regulatory agency review upon request by those agencies.

10.1.4 Committees Structure

10.1.4.1 Steering Committee

This study will be conducted in consultation with a Steering Committee. The Steering Committee is composed of the following:

- Sponsor personnel
- Eisai personnel
- Investigators participating in the study
- Consulting therapeutic area and clinical study experts

The Steering Committee will provide guidance on the operational aspects of the study and provide input with respect to study design, interpretation of results, and subsequent peer-reviewed scientific publications.

Specific details regarding responsibilities and governance of the Steering Committee will be described in a separate charter.

10.1.4.2 Executive Oversight Committee

The EOC is composed of members of Sponsor Senior Management. The EOC will receive and decide upon any recommendations made by the DMC or Steering Committee regarding the study.

10.1.4.3 External Data Monitoring Committee

To supplement the routine study monitoring outlined in this protocol, an external DMC will monitor the interim data from this study. The voting members of the committee are external to the Sponsor. The members of the DMC must not be involved with the study in any other way (eg, they cannot be study investigators) and must have no competing interests that could affect their roles with respect to the study.

The DMC will make recommendations to the EOC regarding steps to ensure both participant safety and the continued ethical integrity of the study. Also, the DMC will review interim study results, consider the overall risk and benefit to study participants (Section 9.7, Interim Analysis) and recommend to the EOC whether the study should continue in accordance with the protocol.

Specific details regarding composition, responsibilities, and governance, including the roles and responsibilities of the various members and the Sponsor protocol team OR external collaborating organization protocol team; meeting facilitation; the study governance structure; and requirements for and proper documentation of DMC reports, minutes, and recommendations will be described in the DMC charter that is reviewed and approved by all the DMC members.

10.1.5 Publication Policy

The results of this study may be published or presented at scientific meetings. The Sponsor will comply with the requirements for publication of study results. In accordance with standard editorial and ethical practice, the Sponsor will generally support publication of multicenter studies only in their entirety and not as individual site data. In this case, a coordinating investigator will be designated by mutual agreement.

If publication activity is not directed by the Sponsor, the investigator agrees to submit all manuscripts or abstracts to the Sponsor before submission. This allows the Sponsor to protect proprietary information and to provide comments.

Authorship will be determined by mutual agreement and in line with International Committee of Medical Journal Editors authorship requirements.

10.1.6 Compliance with Study Registration and Results Posting Requirements

Under the terms of the Food and Drug Administration Amendments Act (FDAAA) of 2007 and the European Medicines Agency (EMA) clinical trial Directive 2001/20/EC, the Sponsor of the study is solely responsible for determining whether the study and its results are subject to the requirements for submission to http://www.clinicaltrials.gov, www.clinicaltrialsregister.eu or other local registries. MSD, as Sponsor of this study, will review this protocol and submit the information necessary to fulfill these requirements. MSD entries are not limited to FDAAA or the EMA clinical trial directive mandated trials. Information posted will allow participants to identify potentially appropriate studies for their disease conditions and pursue participation by calling a central contact number for further information on appropriate study locations and study site contact information.

By signing this protocol, the investigator acknowledges that the statutory obligations under FDAAA, the EMA clinical trials directive, or other locally mandated registries are that of the Sponsor and agrees not to submit any information about this study or its results to those registries.

10.1.7 Compliance with Law, Audit, and Debarment

By signing this protocol, the investigator agrees to conduct the study in an efficient and diligent manner and in conformance with this protocol; generally accepted standards of GCP (eg, International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use GCP: Consolidated Guideline and other generally accepted standards of GCP); and all applicable federal, state and local laws, rules and regulations relating to the conduct of the clinical study.

The Code of Conduct, a collection of goals and considerations that govern the ethical and scientific conduct of clinical investigations sponsored by MSD, is provided in this appendix under the Code of Conduct for Clinical Studies.

The investigator agrees not to seek reimbursement from participants, their insurance providers, or from government programs for procedures included as part of the study reimbursed to the investigator by the Sponsor.

The investigator will promptly inform the Sponsor of any regulatory authority inspection conducted for this study.

The investigator agrees to provide the Sponsor with relevant information from inspection observations/findings to allow the Sponsor to assist in responding to any citations resulting from regulatory authority inspection and will provide the Sponsor with a copy of the proposed response for consultation before submission to the regulatory authority.

Persons debarred from conducting or working on clinical studies by any court or regulatory authority will not be allowed to conduct or work on this Sponsor's studies. The investigator will immediately disclose in writing to the Sponsor if any person who is involved in conducting the study is debarred or if any proceeding for debarment is pending or, to the best of the investigator's knowledge, threatened.

10.1.8 Data Quality Assurance

All participant data relating to the study will be recorded on printed or electronic CRF unless transmitted to the Sponsor or designee electronically (eg, laboratory data). The investigator or qualified designee is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

Detailed information regarding Data Management procedures for this protocol will be provided separately.

The investigator must maintain accurate documentation (source data) that supports the information entered in the CRF.

The investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

Study documentation will be promptly and fully disclosed to the Sponsor by the investigator upon request and also shall be made available at the study site upon request for inspection, copying, review, and audit at reasonable times by representatives of the Sponsor or any regulatory authorities. The investigator agrees to promptly take any reasonable steps that are requested by the Sponsor or any regulatory authorities as a result of an audit or inspection to cure deficiencies in the study documentation and worksheets/CRFs.

The Sponsor or designee is responsible for the data management of this study including quality checking of the data.

Study monitors will perform ongoing source data review and verification to confirm that data entered into the CRF by authorized site personnel are accurate, complete, and verifiable from source documents; that the safety and rights of participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH GCP, and all applicable regulatory requirements.

Records and documents, including participants' documented informed consent, pertaining to the conduct of this study must be retained by the investigator for 15 years after study completion unless local regulations or institutional policies require a longer retention period. No records may be destroyed during the retention period without the written approval of the Sponsor. No records may be transferred to another location or party without written notification to the Sponsor.

10.1.9 Source Documents

Source documents provide evidence for the existence of the participant and substantiate the integrity of the data collected. The investigator/institution should maintain adequate and accurate source documents and study records that include all pertinent observations on each of the site's participants. Source documents and data should be attributable, legible, contemporaneous, original, accurate, and complete. Changes to source data should be traceable, should not obscure the original entry, and should be explained if necessary (eg, via an audit trail). Source documents are filed at the investigator's site.

Data reported on the CRF or entered in the eCRF that are transcribed from source documents must be consistent with the source documents or the discrepancies must be explained. The investigator/institution may need to request previous medical records or transfer records, depending on the study. Also, current medical records must be available.

10.1.10 Study and Site Closure

The Sponsor or its designee may stop the study or study site participation in the study for medical, safety, regulatory, administrative, or other reasons consistent with applicable laws, regulations, and GCP.

In the event the Sponsor prematurely terminates a particular study site, the Sponsor will promptly notify that study site's IRB/IEC.

10.2 Appendix 2: Clinical Laboratory Tests

- The laboratory tests detailed in Table 11 will be performed by the local laboratory.
- Protocol-specific requirements for inclusion or exclusion of participants are detailed in Section 5 of the protocol.
- Additional tests may be performed at any time during the study if determined necessary by the investigator or required by local regulations.
- Pregnancy testing requirements for study inclusion are described in Section 5.1 and Section 8.3.5.2.
- Pregnancy testing (urine or serum as required by local regulations) should be conducted at monthly intervals during intervention.
- Pregnancy testing (urine or serum as required by local regulations) should be conducted at the end of relevant systemic exposure and correspond with the time frame for female participant contraception in Section 5.1, Inclusion Criteria.
- Additional serum or urine pregnancy tests may be performed, as determined necessary by the investigator or required by local regulation, to establish the absence of pregnancy at any time during the participant's participation in the study.

Table 11 Protocol-Required Safety Laboratory Assessments

Hematology	Comprehensive Chemistry Panel	Urinalysis	Other
Hematocrit	Albumin	Blood	FSH[a]
Hemoglobin	Alkaline phosphatase	Glucose	Pregnancy test (highly sensitive serum or urine) human chorionic gonadotropin (hCG)[b]
Platelet count	ALT	Protein	PT/INR and aPTT/PTT
RBC WBC count with differential (absolute or percentage): Neutrophils Lymphocytes Monocytes Eosinophils Basophils	AST	Specific gravity	Total T3 (or FT3), total T4 (or FT4), and TSH[c,d]
	CO_2 or bicarbonate[e]	Microscopic examination, if abnormal results are noted	HIV testing[f]
	Calcium		Hepatitis B and C testing[f,g]
	Chloride		Blood for genetic analysis
	Creatinine		
	Glucose		
	Magnesium		
	Phosphorus		
	Potassium		Blood for serum biomarkers
	Sodium		Blood for RNA analysis
	Total bilirubin		Blood for circulating tumor nucleic acids
	Direct bilirubin[h]		Blood for plasma biomarkers
	Total protein		
	BUN or urea[i]		
	Amylase		
	Lipase		

Abbreviations: ALT = alanine aminotransferase; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; BUN = blood urea nitrogen; C1D1 = Cycle 1 Day 1; CO_2 = carbon dioxide; FSH = follicle-stimulating hormone; FT3 = free triiodothyronine; FT4 = free thyroxine; HIV = human immunodeficiency virus; INR = international normalized ratio; PT = prothrombin time; PTT = partial thromboplastin time; RBC = red blood cells; SOC = standard of care; T3 = triiodothyronine; T4 = thyroxine; TSH = thyroid-stimulating hormone; WBC = white blood cell; WOCBP = women of childbearing potential.

[a] As needed, FSH to be performed at screening to confirm postmenopausal status.

[b] Perform on WOCBP only 24 hours before first dose. Pregnancy tests must be repeated before every cycle.

[c] T3 is preferred; if not available, free T3 may be tested.

[d] Participants may be dosed in subsequent cycles after C1D1 while thyroid function tests are pending.

[e] Bicarbonate only if available as SOC in the region.

[f] If mandated by local health authority.

[g] Hepatitis B surface antigen (HBsAg) or hepatitis B virus (HBV)-DNA. Hepatitis C virus (HCV)-RNA (qualitative) or HCV antibody.

[h] If total bilirubin is elevated above the upper limit of normal.

[i] BUN or urea (one or the other should be collected per institutional standard).

The investigator (or medically qualified designee) must document their review of each laboratory safety report.

10.3 Appendix 3: Adverse Events: Definitions and Procedures for Recording, Evaluating, Follow-up, and Reporting

10.3.1 Definition of AE

AE definition

- An AE is any untoward medical occurrence in a clinical study participant, temporally associated with the use of study intervention, whether or not considered related to the study intervention.

- NOTE: An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of a study intervention.

- NOTE: For purposes of AE definition, study intervention (also referred to as Sponsor's product) includes any pharmaceutical product, biological product, vaccine, diagnostic agent, or protocol specified procedure whether investigational or marketed (including placebo, active comparator product, or run-in intervention), manufactured by, licensed by, provided by, or distributed by the Sponsor for human use in this study.

Events meeting the AE definition

- Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (eg, ECG, radiological scans, vital signs measurements), including those that worsen from baseline, considered clinically significant in the medical and scientific judgment of the investigator.

- Exacerbation of a chronic or intermittent pre-existing condition including either an increase in frequency and/or intensity of the condition.

- New conditions detected or diagnosed after study intervention administration even though it may have been present before the start of the study.

  Note: Congenital disorders (eg, present from birth) not detected or diagnosed prior to study intervention administration do not qualify for reporting as AE.

- Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.

- Signs, symptoms, or the clinical sequelae of a suspected overdose of either study intervention or a concomitant medication.

- For all reports of overdose (whether accidental or intentional) with an associated AE, the AE term should reflect the clinical symptoms or abnormal test result. An overdose without any associated clinical symptoms or abnormal laboratory results is reported using the terminology "accidental or intentional overdose without adverse effect."

Events NOT meeting the AE definition

- Medical or surgical procedure (eg, endoscopy, appendectomy): the condition that leads to the procedure is the AE.
- Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).
- Anticipated day-to-day fluctuations of pre-existing disease(s) or condition(s) present or detected at the start of the study that do not worsen.
- Surgery planned prior to informed consent to treat a pre-existing condition that has not worsened.
- Refer to Section 8.4.6 for protocol-specific exceptions.

10.3.2 Definition of SAE

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met.

An SAE is defined as any untoward medical occurrence that, at any dose:

a. Results in death

b. Is life-threatening

- The term "life-threatening" in the definition of "serious" refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe.

c. Requires inpatient hospitalization or prolongation of existing hospitalization

- Hospitalization is defined as an inpatient admission, regardless of length of stay, even if the hospitalization is a precautionary measure for continued observation. (Note: Hospitalization for an elective procedure to treat a pre-existing condition that has not worsened is not an SAE. A pre-existing condition is a clinical condition that is diagnosed prior to the use of an MSD product and is documented in the participant's medical history.

d. Results in persistent or significant disability/incapacity

- The term disability means a substantial disruption of a person's ability to conduct normal life functions.
- This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza,

and accidental trauma (eg, sprained ankle) that may interfere with or prevent everyday life functions but do not constitute a substantial disruption.

e. Is a congenital anomaly/birth defect

- In offspring of participant taking the product regardless of time to diagnosis.

f. Other important medical events

- Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent 1 of the other outcomes listed in the above definition. These events should usually be considered serious.

- Examples of such events include invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

10.3.3 Additional Events Reported in the Same Manner as SAE

Additional events that require reporting in the same manner as SAE

In addition to the above criteria, AEs meeting either of the below criteria, although not serious per ICH definition, are reportable to the Sponsor in the same time frame as SAEs to meet certain local requirements. Therefore, these events are considered serious by the Sponsor for collection purposes.

- Is a new cancer (that is not a condition of the study)

- Is associated with an overdose of pembrolizumab

10.3.4 Recording AE and SAE

AE and SAE recording

- When an AE/SAE occurs, it is the responsibility of the investigator to review all documentation (eg, hospital progress notes, laboratory, and diagnostics reports) related to the event.

- The investigator will record all relevant AE/SAE information on the AE CRFs/worksheets at each examination.

- It is not acceptable for the investigator to send photocopies of the participant's medical records to the Sponsor in lieu of completion of the AE CRF page.

- There may be instances when copies of medical records for certain cases are requested by the Sponsor. In this case, all participant identifiers, with the exception of the participant number, will be blinded on the copies of the medical records before submission to the Sponsor.

- The investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. In such cases, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

Assessment of intensity/toxicity

- An event is defined as "serious" when it meets at least 1 of the predefined outcomes as described in the definition of an SAE, not when it is rated as severe.

1. The investigator will make an assessment of intensity for each AE and SAE (and other reportable safety event) according to the NCI Common Terminology for Adverse Events (CTCAE), version 4. Any AE that changes CTCAE grade over the course of a given episode will have each change of grade recorded on the AE CRFs/worksheets.

    - Grade 1: Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.

    - Grade 2: Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL).

    - Grade 3: Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL.

    - Grade 4: Life threatening consequences; urgent intervention indicated.

    - Grade 5: Death related to AE.

Assessment of causality

1. Did the Sponsor's product cause the AE?

2. The determination of the likelihood that the Sponsor's product caused the AE will be provided by an investigator who is a qualified physician. The investigator's signed/dated initials on the source document or worksheet that supports the causality noted on the AE form, ensures that a medically qualified assessment of causality was done. This initialed document must be retained for the required regulatory time frame. The criteria below are intended as reference guidelines to assist the investigator in assessing the likelihood of a relationship between the test product and the AE based upon the available information.

3. **The following components are to be used to assess the relationship between the Sponsor's product and the AE; the greater the correlation with the components and their respective elements (in number and/or intensity), the more likely the Sponsor's product caused the AE:**

   - **Exposure:** Is there evidence that the participant was actually exposed to the Sponsor's product such as: reliable history, acceptable compliance assessment (pill count, diary, etc.), expected pharmacologic effect, or measurement of drug/metabolite in bodily specimen?

   - **Time Course:** Did the AE follow in a reasonable temporal sequence from administration of the Sponsor's product? Is the time of onset of the AE compatible with a drug-induced effect (applies to studies with investigational medicinal product)?

   - **Likely Cause:** Is the AE not reasonably explained by another etiology such as underlying disease, other drug(s)/vaccine(s), or other host or environmental factors.

   - **Dechallenge:** Was the Sponsor's product discontinued or dose/exposure/frequency reduced?

   - If yes, did the AE resolve or improve?

   - If yes, this is a positive dechallenge.

   - If no, this is a negative dechallenge.

   - (Note: This criterion is not applicable if: (1) the AE resulted in death or permanent disability; (2) the AE resolved/improved despite continuation of the Sponsor's product; (3) the study is a single-dose drug study; or (4) Sponsor's product(s) is/are only used 1 time.)

   - **Rechallenge:** Was the participant re-exposed to the Sponsor's product in this study?

   - If yes, did the AE recur or worsen?

   - If yes, this is a positive rechallenge.

   - If no, this is a negative rechallenge.

   (Note: This criterion is not applicable if: (1) the initial AE resulted in death or permanent disability, or (2) the study is a single-dose drug study; or (3) Sponsor's product(s) is/are used only 1 time.)

NOTE: IF A RECHALLENGE IS PLANNED FOR AN AE THAT WAS SERIOUS AND MAY HAVE BEEN CAUSED BY THE SPONSOR'S PRODUCT, OR IF RE-EXPOSURE

TO THE SPONSOR'S PRODUCT POSES ADDITIONAL POTENTIAL SIGNIFICANT RISK TO THE PARTICIPANT THEN THE RECHALLENGE MUST BE APPROVED IN ADVANCE BY THE SPONSOR CLINICAL DIRECTOR AS PER DOSE MODIFICATION GUIDELINES IN THE PROTOCOL, AND IF REQUIRED, THE INIRB/IEC.

4. **Consistency with study intervention profile:** Is the clinical/pathological presentation of the AE consistent with previous knowledge regarding the Sponsor's product or drug class pharmacology or toxicology?

5. The assessment of relationship will be reported on the case report forms/worksheets by an investigator who is a qualified physician according to his/her best clinical judgment, including consideration of the above elements.

6. Use the following scale of criteria as guidance (not all criteria must be present to be indicative of a Sponsor's product relationship).

   - Yes, there is a reasonable possibility of Sponsor's product relationship:

   - There is evidence of exposure to the Sponsor's product. The temporal sequence of the AE onset relative to the administration of the Sponsor's product is reasonable. The AE is more likely explained by the Sponsor's product than by another cause.

   - No, there is not a reasonable possibility of Sponsor's product relationship:

   - Participant did not receive the Sponsor's product OR temporal sequence of the AE onset relative to administration of the Sponsor's product is not reasonable OR the AE is more likely explained by another cause than the Sponsor's product. (Also entered for a participant with overdose without an associated AE.)

7. For each AE/SAE, the investigator must document in the medical notes that he/she has reviewed the AE/SAE and has provided an assessment of causality.

8. There may be situations in which an SAE has occurred and the investigator has minimal information to include in the initial report to the Sponsor. However, it is very important that the investigator always make an assessment of causality for every event before the initial transmission of the SAE data to the Sponsor.

9. The investigator may change his/her opinion of causality in light of follow-up information and send an SAE follow-up report with the updated causality assessment.

10. The causality assessment is 1 of the criteria used when determining regulatory reporting requirements.

11. For studies in which multiple agents are administered as part of a combination regimen, the investigator may attribute each AE causality to the combination regimen or to a single agent of the combination. In general, causality attribution should be assigned to the combination regimen (ie, to all agents in the regimen). However,

causality attribution may be assigned to a single agent if in the investigator's opinion, there is sufficient data to support full attribution of the AE to the single agent.

Follow-up of AE and SAE

- The investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by Sponsor to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.

- New or updated information will be recorded in the CRF.

- The investigator will submit any updated SAE data to the Sponsor within 24 hours of receipt of the information.

10.3.5 Reporting of AEs, SAEs, and Other Reportable Safety Events to the Sponsor

AE, SAE, and other reportable safety event reporting to Sponsor via electronic data collection tool

- The primary mechanism for reporting to the Sponsor will be the electronic data collection (EDC) tool.
    - Electronic reporting procedures can be found in the EDC data entry guidelines (or equivalent).
    - If the electronic system is unavailable for more than 24 hours, then the site will use the paper AE Reporting form.
        - Reference Section 8.4.1 for reporting time requirements.
- The site will enter the SAE data into the electronic system as soon as it becomes available.
- After the study is completed at a given site, the EDC tool will be taken off-line to prevent the entry of new data or changes to existing data.
- If a site receives a report of a new SAE from a study participant or receives updated data on a previously reported SAE after the EDC tool has been taken off-line, then the site can report this information on a paper SAE form or by telephone (see next section).
- Contacts for SAE reporting can be found in the Investigator Study File Binder (or equivalent).

SAE reporting to the Sponsor via paper CRF

- If the EDC tool is not operational, facsimile transmission or secure e-mail of the SAE paper CRF is the preferred method to transmit this information to the Sponsor.

- In rare circumstances and in the absence of facsimile equipment, notification by telephone is acceptable with a copy of the SAE data collection tool sent by overnight mail or courier service.

- Initial notification via telephone does not replace the need for the investigator to complete and sign the SAE CRF pages within the designated reporting time frames.

- Contacts and instructions for SAE reporting and paper reporting procedures can be found in the Investigator Study File Binder (or equivalent).

10.4 Appendix 4: Device Events, Adverse Device Events, and Medical Device Incidents: Definitions, Collection, and Documentation

This section is not applicable.

10.5 Appendix 5: Contraceptive Guidance and Pregnancy Testing

10.5.1 Definitions

Women of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming postmenopausal unless permanently sterile (see below).

If fertility is unclear (eg, amenorrhea in adolescents or athletes) and a menstrual cycle cannot be confirmed before first dose of study intervention, additional evaluation should be considered.

Women in the following categories are not considered WOCBP:

- Premenarchal
- Premenopausal female with 1 of the following:
  - Documented hysterectomy
  - Documented bilateral salpingectomy
  - Documented bilateral oophorectomy

For individuals with permanent infertility due to an alternate medical cause other than the above (eg, Mullerian agenesis, androgen insensitivity), investigator discretion should be applied to determining study entry.

Note: Documentation can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.

- Postmenopausal female
  - A postmenopausal state is defined as no menses for 12 months without an alternative medical cause.
    - A high follicle-stimulating hormone (FSH) level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or HRT. However, in the absence of 12 months of amenorrhea, confirmation with 2 FSH measurements in the postmenopausal range is required.
    - Females on HRT and whose menopausal status is in doubt will be required to use one of the nonhormonal highly effective contraception methods if they wish to continue their HRT during the study. Otherwise, they must discontinue HRT to allow confirmation of postmenopausal status before study enrollment.

10.5.2 Contraception Requirements

Contraceptives allowed during the study include[a]:
Highly Effective Contraceptive Methods That Have Low User Dependency *Failure rate of <1% per year when used consistently and correctly.*
• Progestogen-only subdermal contraceptive implant[b] • Intrauterine hormone-releasing system (IUS)[c] • Intrauterine device (IUD) • Bilateral tubal occlusion
• Azoospermic partner (Vasectomized or secondary to medical cause) • This is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used. A spermatogenesis cycle is approximately 90 days. Note: Documentation of azoospermia can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.
Sexual Abstinence • Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the study intervention. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.
a. Contraceptive use by men or women should be consistent with local regulations regarding the use of contraceptive methods for participants of clinical studies. b. If locally required, in accordance with Clinical Trial Facilitation Group (CTFG) guidelines, acceptable contraceptive implants are limited to those which inhibit ovulation. c. IUS is a progestin-releasing IUD. Note: The following are not acceptable methods of contraception: – Periodic abstinence (calendar, symptothermal, postovulation methods), withdrawal (coitus interruptus), spermicides only, and lactational amenorrhea method (LAM). – Male condom with cap, diaphragm, or sponge with spermicide. – Male and female condom should not be used together (due to risk of failure with friction).

10.5.3 Pregnancy Testing

WOCBP should only be included after a negative highly sensitive urine or serum pregnancy test. Refer to Section 8.3.5.2 and Appendix 2 for further details.

A urine pregnancy test will be obtained per the SoA (Section 1.3). Additional urine/serum testing may be performed if clinically warranted and/or as defined by local regulations. If a urine pregnancy test cannot be confirmed as negative, a serum pregnancy test is required.

10.6 Appendix 6: Description of the iRECIST Process for Assessment of Disease Progression

iRECIST is based on RECIST 1.1 but adapted to account for the unique tumor response seen with immunotherapeutic drugs. iRECIST will be used by the investigator to assess tumor response and progression, and to guide decisions about changes in management.

Assessment at Screening and Before RECIST 1.1 Progression

Until radiographic PD based on RECIST 1.1, there is no distinct iRECIST assessment.

Assessment and Decision at RECIST 1.1 Progression

For participants who show radiological PD by RECIST 1.1, the investigator will decide whether to continue a participant on study intervention until repeat scans are obtained, as described in Section 8.2.1.6.

Tumor flare may manifest as any factor causing radiographic progression per RECIST 1.1, including:

- Increase in the sum of diameters of target lesion(s) identified at baseline to ≥20% and ≥5 mm from nadir
- Note: the iRECIST publication uses the terminology “sum of measurements,” but “sum of diameters” will be used in this protocol, consistent with the original RECIST 1.1 terminology
- Unequivocal progression of nontarget lesion(s) identified at baseline
- Development of new lesion(s)

iRECIST defines new response categories, including iUPD (unconfirmed PD) and iCPD (confirmed PD). For purposes of iRECIST assessment, the first visit showing PD according to RECIST 1.1 will be assigned a visit (overall) response of iUPD, regardless of which factors caused the PD.

At this visit, target and nontarget lesions identified at baseline by RECIST 1.1 will be assessed as usual.

New lesions will be classified as measurable or nonmeasurable, using the same size thresholds and rules as for baseline lesion assessment in RECIST 1.1. From measurable new lesions, up to 5 lesions total (up to 2 per organ), may be selected as New Lesions – Target. The sum of diameters of these lesions will be calculated and kept distinct from the sum of diameters for target lesions at baseline. All other new lesions will be followed qualitatively as New Lesions – Nontarget.

Assessment at the Confirmatory Imaging

On the confirmatory imaging, the participant will be classified as PD confirmed (with an overall response of iCPD), or as showing persistent unconfirmed PD (with an overall response of iUPD), or as showing disease stability or response (iSD/iPR/iCR).

Confirmation of PD

PD is considered confirmed, and the overall response will be iCPD, if ANY of the following occur:

- Any of the factors that were the basis for the iUPD at the previous visit show worsening
  - For target lesions, worsening is a further increase in the sum of diameters of ≥5 mm, compared with any prior iUPD time point
  - For nontarget lesions, worsening is any significant growth in lesions overall, compared with a prior iUPD time point; this does not have to meet the "unequivocal" standard of RECIST 1.1
  - For new lesions, worsening is any of these:
    - An increase in the new lesion sum of diameters by ≥5 mm from a prior iUPD time point
    - Visible growth of new nontarget lesions
    - The appearance of additional new lesions
- Any new factor appears that would have triggered PD by RECIST 1.1

Persistent iUPD

PD is considered not confirmed, and the overall response remains iUPD, if:

- None of the PD-confirming factors identified above occurs AND
- The target lesion sum of diameters (initial target lesions) remains above the initial PD threshold (by RECIST 1.1)

Additional imaging for confirmation should be scheduled 4 to 8 weeks from the imaging on which iUPD is seen. This may correspond to the next visit in the original visit schedule. The assessment of the subsequent confirmation imaging proceeds in an identical manner, with possible outcomes of iCPD, iUPD, and iSD, iPR, and iCR.

Resolution of iUPD

PD is considered not confirmed, and the overall response becomes iSD, iPR, or iCR, if:

- None of the PD-confirming factors identified above occurs, AND
- The target lesion sum of diameters (initial target lesions) is not above the initial PD threshold.

The response is classified as iSD or iPR (depending on the sum of diameters of the target lesions) or iCR if all lesions resolve.

In this case, the initial iUPD is considered to be pseudoprogression, and the level of suspicion for progression is "reset." This means that the next visit that shows radiographic progression, whenever it occurs, is again classified as iUPD by iRECIST, and the confirmation process is repeated before a response of iCPD can be assigned.

Management Following the Confirmatory Imaging

If repeat imaging does not confirm PD per iRECIST, as assessed by the investigator, and the participant continues to be clinically stable, study intervention may continue and follow the regular imaging schedule. If PD is confirmed, participants will be discontinued from study intervention.

NOTE: If a participant has confirmed radiographic progression (iCPD) as defined above, but the participant is achieving a clinically meaningful benefit, an exception to continue study intervention may be considered following consultation with the Sponsor. In this case, if study intervention is continued, tumor imaging should continue to be performed following the intervals as outlined in the SoA (Section 1.3).

Detection of Progression at Visits After Pseudo-progression Resolves

After resolution of pseudoprogression (ie, achievement of iSD, iPR, iCR, iUPD is indicated by any of the following events):

- Target lesions
  - Sum of diameters reaches the PD threshold (≥20% and ≥5 mm increase from nadir) either for the first time, or after resolution of previous pseudoprogression. The nadir is always the smallest sum of diameters seen during the entire trial, either before or after an instance of pseudoprogression.
- Nontarget lesions
  - If nontarget lesions have never shown unequivocal PD, their doing so for the first time results in iUPD.

  - If nontarget lesions have shown previous unequivocal PD, and this PD has not resolved, iUPD results from any significant further growth of nontarget lesions, taken as a whole.

- New lesions
  - New lesions appear for the first time
  - Additional new lesions appear
  - Previously identified new target lesions show an increase of ≥5 mm in the new lesion sum of diameters, from the nadir value of that sum
  - Previously identified nontarget lesions show any significant growth

If any of the events above occur, the overall response for that visit is iUPD, and the iUPD evaluation process (see Assessment at the Confirmatory Imaging above) is repeated. PD must be confirmed before iCPD can occur.

The decision process is identical to the iUPD confirmation process for the initial PD, with 1 exception: if new lesions occurred at a prior instance of iUPD, and at the confirmatory imaging the burden of new lesions has increased from its smallest value (for new target lesions, the sum of diameters is ≥5 mm increased from its nadir), then iUPD cannot resolve to iSD or iPR. It will remain iUPD until either a decrease in the new lesion burden allows resolution to iSD or iPR, or until a confirmatory factor causes iCPD.

Additional details about iRECIST are provided in the iRECIST publication [Seymour, L., et al 2017].

10.7 Appendix 7: Country-specific Requirements

10.7.1 Japan-specific Requirements

For the assistance to early diagnosis of pneumonitis/interstitial lung disease (ILD) in study participants, the following items, such as pulse oximetry monitoring (peripheral capillary oxygen saturation [SpO_2]), C-reactive protein (CRP), Krebs von den Lungen-6 (KL-6), and surfactant protein-D (SP-D), will be measured in this study. These items should be measured as follows:

- SpO_2 at the timing of vital sign assessment
- CRP, KL-6, and SP-D at screening*, predose on Day 1 of every cycle, end of treatment, and the Safety Follow-up Visit (30 days after last dose)

*Should be measured at the timing of clinical laboratory tests (such as hematology/chemistry).

If pneumonitis/ILD occurs regardless of causality with study intervention, an independent ILD evaluation committee will conduct adjudication of cases of the pneumonitis/ILD. For this purpose, relevant data, such as chest imaging (from the baseline to the recovery of pneumonitis/ILD) will be submitted to MSD K.K.

10.7.2 France-specific Requirements

Treatment with pembrolizumab should be permanently discontinued in cases of confirmed Stevens-Johnson syndrome or toxic epidermal necrolysis.

10.7.3 Canada-specific Requirements

See Section 6.6.1 Immune-related Events and Dose Modification (Withhold, Treat, Discontinue)

Treatment with pembrolizumab should be permanently discontinued in cases of confirmed Stevens-Johnson syndrome or toxic epidermal necrolysis.

10.8 Appendix 8: Abbreviations

Abbreviation	Expanded Term
1L	first-line
ADA	antidrug antibody
ADL	activities of daily living
AE	adverse event
AJCC	American Joint Committee on Cancer
ALK	anaplastic lymphoma kinase
APaT	All Participants as Treated
BICR	blinded independent central review
CI	confidence interval
CL	clearance
COPD	chronic obstructive pulmonary disease
CR	complete response
CRF	case report form
CRP	C-reactive protein
CSR	clinical study report
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTFG	Clinical Trial Facilitation Group
CVA	cerebrovascular accident
DILI	drug-induced liver injury
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DOR	duration of response
ECG	electrocardiogram
ECI	event of clinical interest
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data collection
EGFR	epidermal growth factor receptor
EMA	European Medicines Agency

Abbreviation	Expanded Term
EOC	Executive Oversight Committee
EORTC	European Organization for Research and Treatment of Cancer
ePRO	electronic patient-reported outcome
EQ-5D-5L	EuroQoL-5D-5L
FAS	Full Analysis Set
FDA	Food and Drug Administration
FDAAA	Food and Drug Administration Amendments Act
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor
FSH	follicle-stimulating hormone
GCP	Good Clinical Practice
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HR	hazard ratio
HRQoL	health-related quality of life
HRT	hormone replacement therapy
IA	interim analysis
IB	Investigator's Brochure
ICF	informed consent form
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
iCPD	confirmed PD
iCR	confirmed CR
IEC	Independent Ethics Committee
irAE	immune-related adverse event
iPR	confirmed PR
IRB	Institutional Review Board
iRECIST	adjusted RECIST 1.1 for Immune-based Therapeutics
iSD	confirmed SD
ITT	Intention to Treat

Abbreviation	Expanded Term
IUD	intrauterine device
iUPD	unconfirmed PD
IUS	intrauterine hormone-releasing system
IV	intravenous
KL-6	Krebs von den Lungen-6
LEAP	lenvatinib and pembrolizumab
LVEF	left ventricular ejection fraction
mAb	monoclonal antibody
MRI	magnetic resonance imaging
MTD	maximum tolerated dose
MUGA	multigated acquisition scan
NCI	National Cancer Institute
NSCLC	non–small cell lung cancer
OR	objective response
ORR	objective response rate
OS	overall survival
PD	progressive disease
PD-1	programmed cell death protein
PDGFR	platelet-derived growth factor receptor
PD-L1	programmed cell death ligand 1
PD-L2	programmed cell death ligand 2
PFS	progression-free survival
Pgp	P-glycoprotein
PK	pharmacokinetic
PopPK	population pharmacokinetics
PR	partial response
PRES	posterior reversible encephalopathy syndrome
PRO	patient-reported outcomes
Q3W	every 3 weeks
QD	once daily
QLQ-LC13	Quality of Life Questionnaire and Lung Cancer Module 13
QoL	quality of life

Abbreviation	Expanded Term
RCC	renal cell carcinoma
RECIST 1.1	Response Evaluation Criteria in Solid Tumors 1.1
RNA	ribonucleic acid
RP2D	recommended Phase 2 dose
RPLS	reversible posterior leukoencephalopathy syndrome
RTKi	receptor tyrosine kinase inhibitor
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SoA	schedule of activities
SOC	standard of care
SP-D	surfactant protein-D
SpO_2	peripheral capillary oxygen saturation
sSAP	supplemental statistical analysis plan
TAM	tumor-associated macrophage
TB	tuberculosis
TPS	tumor proportion score
TTD	time to true deterioration
UPCR	urine protein-to-creatinine ratio
V	volume of distribution
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
WOCBP	woman/women of childbearing potential

11 REFERENCES

[Aaronson, N. K., et al 1993]	Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365-76.	[03Q3QL]
[Bergman, B., et al 1994]	Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M. The EORTC QLQ-LC13: a modular supplement to the EORTC core quality of life questionnaire (QLQ-C30) for use in lung cancer clinical trials. Eur J Cancer 1994;30A(5):635-42.	[00VTZB]
[Cross, M. J. 2001]	Cross MJ, Claesson-Welsh L. FGF and VEGF function in angiogenesis: signalling pathways, biological responses and therapeutic inhibition. Trends Pharmacol Sci. 2001 Apr;22(4):201-7.	[04XKP6]
[Disis, M. L. 2010]	Disis ML. Immune regulation of cancer. J Clin Oncol 2010;28(29):4531-8.	[058SQL]
[Dudley, M. E., et al 2005]	Dudley ME, Wunderlich JR, Yang JC, Sherry RM, Topalian SL, Restifo NP, et al. Adoptive cell transfer therapy following non-myeloablative but lymphodepleting chemotherapy for the treatment of patients with refractory metastatic melanoma. J Clin Oncol 2005;23(10):2346-57.	[00VMPR]
[Ellis, L. M. 2008]	Ellis LM, Hicklin DJ. VEGF-targeted therapy: mechanisms of anti-tumour activity. Nat Rev Cancer. 2008 Aug;8:579-91.	[04XKPD]

[Ferrara, N., et al 2003]	Ferrara N, Gerber HP, LeCouter J. The biology of VEGF and its receptors. Nat Med. 2003 Jun;9(6):669-76.	[04XKQ2]
[Havel, L., et al 2014]	Havel L, Lee JS, Lee KH, Bidoli P, Kim JH, Ferry D, et al. E7080 (lenvatinib) in addition to best supportive care (BSC) versus BSC alone in third-line or greater nonsquamous, non-small cell lung cancer (NSCLC) [abstract]. Presented at: The American Society of Clinical Oncology (ASCO) Annual Meeting; 2014 May 31; Alexandria, VA. J Clin Oncol. 2014 May 20;32(15 suppl). Abstract no. 8043.	[04X49Z]
[Hodi, F. S., et al 2014]	Hodi FS, Ribas A, Daud A, Hamid O, Robert C, Kefford R, et al. Patterns of response in patients with advanced melanoma treated with Pembrolizumab (MK-3475) and evaluation of immune-related response criteria (irRC). J Immunother Cancer. 2014;2(Suppl 3):P103.	[0465RW]
[Hunder, N. N., et al 2008]	Hunder NN, Wallen H, Cao J, Hendricks DW, Reilly JZ, Rodmyre R, et al. Treatment of metastatic melanoma with autologous CD4+ T cells against NY-ESO-1. N Engl J Med 2008;358(25):2698-703.	[00VMPX]
[IB Edition 15-Eisai 2018]	Lenvatinib (LENVIMA®, KISPLYX®) Investigator's Brochure, Edition Number 15	[04Z6DS]
[IB Edition 16 2018]	KEYTRUDA (pembrolizumab) Investigator Brochure, Edition Number 16, 29-JUN-2018.	[04W7QW]

[Kato, Y., et al 2019]	Kato Y, Tabata K, Kimura T, Yachie-Kinoshita A, Ozawa Y, Yamada K, et al. Lenvatinib plus anti-PD-1 antibody combination treatment activates CD8(+) T cells through reduction of tumor-associated macrophage and activation of the interferon pathway. PLoS One. 2019 Feb 27;14(2):e0212513.	[057JXY]
[Kimura, T., et al 2018]	Kimura T, Kato Y, Ozawa Y, Kodama K, Ito J, Ichikawa K, et al. Immunomodulatory activity of lenvatinib contributes to antitumor activity in the Hepa1-6 hepatocellular carcinoma model. Cancer Sci. 2018;109:3993-4002.	[057LPH]
[Liang, Kung-Yee and Zeger, Scott L. 2000]	Liang K-Y, Zeger SL. Longitudinal data analysis of continuous and discrete responses for pre-post designs. Sankhya: The Indian Journal of Statistics 2000;62(Series B, Pt. 1):134-48.	[00QJ0M]
[Lieu, C., et al 2011]	Lieu C, Heymach J, Overman M, Tran H, Kopetz S. Beyond VEGF: inhibition of the fibroblast growth factor pathway and antiangiogenesis. Clin Cancer Res. 2011 Oct 1;17(19):6130-9.	[04XKQM]
[Limaverde-Sousa, G., et al 2014]	Limaverde-Sousa G, Sternberg C, Ferreira CG. Antiangiogenesis beyond VEGF inhibition: a journey from antiangiogenic single-target to broad-spectrum agents. Cancer Treat Rev. 2014;40:548-57.	[04XKQW]

[Lopes, G., et al 2018]	Lopes G, Wu Y, Kudaba I, Kowalski D, Cho BC, Castro G, et al. Pembrolizumab (pembro) versus platinum-based chemotherapy (chemo) as first-line therapy for advanced/metastatic NSCLC with a PD-L1 tumor proportion score (TPS) ≥ 1%: open-label, phase 3 KEYNOTE-042 study [abstract]. Presented at: The American Society of Clinical Oncology (ASCO) Annual Meeting; 2018 Jun 1-5; Chicago, IL. J Clin Oncol. 2018;36(suppl). Abstract no. LBA4.	[04Y5G9]
[Makker, V., et al 2018]	Makker V, Rasco D, Vogelzang NJ, Messing M, Brose MS, Cohn AL, et al. Lenvatinib + Pembrolizumab in patients with advanced endometrial cancer: updated results. Poster session presented at: 2018 American Society of Clinical Oncology (ASCO) Annual Meeting; 2018 Jun 1-5; Chicago, IL. Poster no. 5596.	[04YJ9J]
[Miettinen, O. 1985]	Miettinen O, Nurminen M. Comparative analysis of two rates. Stat Med 1985;4:213-26.	[00VMQY]
[National Cancer Institute 2018]	SEER*Stat [Internet]. Bethesda (MD): National Cancer Institute. 2018. Cancer stat facts: lung and bronchus cancer; [about 12 screens]. Available from: https://seer.cancer.gov/statfacts/html/lungb.html.	[0507C2]
[Nishio, M., et al 2013]	Nishio M, Horai T, Horiike A, Nokihara H, Yamamoto N, Takahashi T, et al. Phase 1 study of lenvatinib combined with carboplatin and paclitaxel in patients with non-small-cell lung cancer. Br J Cancer. 2013;109:538-44.	[04ZK34]

[Pickard, A. S., et al 2007]	Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes 2007;5:1-8.	[00W0FM]
[Rabin, R. 2001]	Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol group. Ann Med 2001;33:337-43.	[03QM46]
[Reck, M., et al 2016]	Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2016 Oct 8. [Epub ahead of print].	[04JNM3]
[Seymour, L., et al 2017]	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. Lancet Oncol. 2017 Mar;18(3):e143-52.	[04P9RV]
[Sulpher, J. A., et al 2013]	Sulpher JA, Owen SP, Hon H, Tobros K, Shepherd FA, Sabri E, et al. Factors influencing a specific pathologic diagnosis of non-small-cell lung carcinoma. Clin Lung Cancer. 2013 May;14(3):238-44.	[040565]
[Tammela, T. 2010]	Tammela T, Alitalo K. Lymphangiogenesis: molecular mechanisms and future promise. Cell. 2010 Feb 19;140:460-76.	[04XKRB]
[Wolchok, J. D., et al 2009]	Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, LebbéC, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009;15(23):7412-20.	[00VMNZ]

[World Health Organization 2012]	Globocan 2012: Lung Cancer Estimated Incidence, Mortality and Prevalence Worldwide in 2012 [Internet]. Lyon (FR): World Health Organization, International Agency for Research on Cancer (IARC). c2014 – [updated 2014 Oct 3]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx	[0402K4]
[Yamamoto, Y., et al 2014]	Yamamoto Y, Matsui J, Matsushima T, Obaishi H, Miyazaki K, Nakamura K, et al. Lenvatinib, an angiogenesis inhibitor targeting VEGFR/FGFR, shows broad antitumor activity in human tumor xenograft models associated with microvessel density and pericyte coverage. Vasc Cell. 2014;6:18.	[04XQ0S]

Supplemental Statistical Analysis Plan (sSAP)

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

This supplemental SAP (sSAP) is a companion document to the protocol. In addition to the information presented in the protocol SAP, which provides the principal features of confirmatory analyses for this trial, this supplemental SAP provides additional statistical analysis details/data derivations and documents modifications or additions to the analysis plan that are not "principal" in nature and result from information that was not available at the time of protocol finalization.

2. SUMMARY OF CHANGES

As the observed PFS event number at the July 2021 DMC was more than the expected PFS event number at the IA (PFS final analysis), the sSAP was amended to update the non-binding futility criteria for July 2021 DMC in Appendix 5.1..

3. ANALYTICAL AND METHODOLOGICAL DETAILS FOR GLOBAL STUDY

3.1. Statistical Analysis Plan Summary

Study Design Overview	Phase 3 study of pembrolizumab plus lenvatinib vs pembrolizumab plus placebo for 1L treatment of metastatic NSCLC in participants whose tumors express PD-L1 (TPS ≥1%)
Treatment Assignment	Approximately 620 participants will be randomized in a 1:1 ratio between 2 treatment arms: (1) pembrolizumab + lenvatinib and (2) pembrolizumab + placebo. Stratification factors are as follows: • Geographic region (East Asia vs non-East Asia) • ECOG PS (0 vs 1) • TPS (1%-49% vs ≥50%)
Analysis Populations	Efficacy: Intention to Treat (ITT) Safety: All Participants as Treated (APaT) PRO: Full Analysis Set (FAS)
Primary Endpoints	• PFS per RECIST 1.1 based on BICR • OS
Secondary Endpoints	• OR per RECIST 1.1 based on BICR • Safety and tolerability • Mean change from baseline in Global health status/QoL, cough, chest pain, dyspnea and physical functioning scores • Time to True Deterioration (TTD) in global health status/QoL, cough, chest pain, dyspnea, physical functioning, and composite endpoint (combination of cough, chest pain, or dyspnea items)
Statistical Methods for Key Efficacy Analyses	The primary hypotheses will be evaluated by comparing pembrolizumab + lenvatinib to pembrolizumab + placebo in PFS and OS using a stratified log-rank test and in ORR using the stratified Miettinen and Nurminen method [Miettinen, O. 1985]. The HR will be estimated using a stratified Cox regression model. Event rates over time will be estimated within each treatment group using the Kaplan-Meier method. The difference in ORR will be estimated using the stratified Miettinen and Nurminen method with strata weighting by sample size.

Statistical Methods for Key Safety Analyses	The primary hypotheses will be evaluated by comparing pembrolizumab + lenvatinib to pembrolizumab + placebo in PFS and OS using a stratified log-rank test and in ORR using the stratified Miettinen and Nurminen method [Miettinen, O. 1985]. The HR will be estimated using a stratified Cox regression model. Event rates over time will be estimated within each treatment group using the Kaplan-Meier method. The difference in ORR will be estimated using the stratified Miettinen and Nurminen method with strata weighting by sample size.

Interim Analyses	One interim analysis and one final analysis are planned for this study. Results will be reviewed by an external Data Monitoring Committee (DMC). Details are provided in Section 9.7.
	• Interim analysis (IA): to be performed after ~416 PFS events are observed and ~8.8 months after last participant randomized
	  ○ Primary purpose: final PFS analysis and interim OS analysis
	• Final analysis: to be performed after ~388 deaths are observed and ~18.5 months after last participant randomized
	  ○ Primary purpose: final OS analysis
	Note that for the IA and the FA, if the events accrue slower than expected such that the targeted number of events cannot be reached in the anticipated timeframe, the Sponsor may conduct the analysis with an additional 2 months and 7 months of follow-up for the IA and FA respectively, or the specified number of events is observed, whichever occurs first.

Multiplicity	The overall Type I error rate over the multiple endpoints will be strongly controlled at 2.5% (1-sided). A 0.55% (1-sided) will be initially allocated to test PFS and 1.95% (1-sided) allocated to test OS. The graphical approach of Maurer and Bretz [Maurer, W., et al 2011] will be applied to reallocate α among the hypotheses for ORR, PFS, and OS. Lan-DeMets [Lan, K. K. G. and DeMets, D. L. 1983] and O'Brien-Fleming [O'Brien, P. C. and Fleming, T. R. 1979] group sequential methods will be used to allocate α among the interim and final analysis for the OS endpoint.
Sample Size and Power	The planned sample size is approximately 620 participants. For PFS, based on 416 events, the study has 86.5% power to detect a HR of 0.7 (pembrolizumab + lenvatinib vs pembrolizumab + placebo) at α=0.55% (1-sided). For OS, based on 388 events, the study has 90% power to detect a HR of 0.71 (pembrolizumab + lenvatinib vs pembrolizumab + placebo) at α=1.95% (1-sided).
China Extension Study	China participants randomized during the global study phase will be included in all global study analyses (efficacy and safety). China participants randomized during the China extension phase will be excluded from all global study analyses. China participants randomized during global and extension phases will both be included in the China-specific analyses.

3.2. Responsibility for Analyses/In-house Blinding

The statistical analysis of the data obtained from this study will be the responsibility of the Clinical Biostatistics department of the Sponsor.

This study will be conducted as a double-blind study under in-house blinding procedures. The official, final database will not be unblinded until medical/scientific review has been performed, protocol deviations have been identified, and data have been declared final and complete.

The Sponsor will generate the randomized allocation schedule(s) for study intervention assignment.

Blinding issues related to the planned interim analyses are described in Section 3.7.

3.3. Hypotheses/Estimation

Objectives and hypotheses of the study are stated in Section 3 of the study protocol.

3.4. Analysis Endpoints

Efficacy and safety endpoints that will be evaluated for within- and between-treatment differences are listed below.

3.4.1. Efficacy Endpoints

3.4.1.1. Primary Endpoints

***PFS*:** the time from randomization to the first documented PD per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR or death due to any cause, whichever occurs first. See Section 9.6.1 for the definition of censoring.

***OS*:** the time from randomization to death due to any cause.

3.4.1.2. Secondary Endpoints

***OR*:** a confirmed CR or PR per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR.

3.4.1.3. Exploratory Endpoints

***DOR*:** the time from the earliest date of qualifying response until earliest date of PD or death from any cause, whichever comes first.

3.4.2. Safety Endpoints

Safety and tolerability will be assessed by clinical review of all relevant parameters including AEs, SAEs, fatal AEs, laboratory tests, and vital signs. Furthermore, specific events will be collected and designated as ECIs as described in Section 8.4.7 of the study protocol.

3.4.3. Patient-reported Outcome Endpoints

Secondary

- Mean change from baseline in EORTC QLQ-C30 global health status/QoL (items 29-30), physical functioning (items 1-5), and dyspnea scores (item 8), and EORTC QLQ-LC13 cough (item 31) and chest pain (item 40) scores.

- TTD as measured by each of the EORTC QLQ-C30 global health status/QoL, physical functioning, and dyspnea scores, and EORTC QLQ-LC13 cough and chest pain scores.

- TTD as measured by a composite symptom endpoint: cough (QLQ-LC13 item 31), chest pain (QLQ-LC13 item 40), or dyspnea (QLQ-C30 item 8)

Based on prior literature (Osoba D et al., 1998; King et al., 1996, Maringwa et al., 2011), a ≥10 point or greater worsening from baseline for each scale represents a minimally important difference (MID) that represents a clinically relevant deterioration. TTD in the individual scores specified above is defined as the time to first onset of ≥10 (out of 100) point deterioration from baseline in a given scale/subscale/item and confirmed by a second adjacent ≥10 point deterioration from baseline. TTD in the composite endpoint is defined as the time to first onset of a ≥10 point deterioration (i.e., increase in score) from baseline in any one of the 3 lung cancer symptom scale items (EORTC-QLQ-C30 [dyspnea] and EORTC-QLQ-LC13 [cough and chest pain]) with confirmation by the subsequent visit of a ≥10 point deterioration (i.e., increase in score) from baseline in the same scale as the first onset. Changes from baseline in EORTC QLQ-C30 scores will also be interpreted according to recent subscale-specific guidelines, which indicate that clinically meaningful differences vary by scale (Cocks et al., 2012)

Exploratory

- Change from baseline in the additional scales of the EORTC QLQ-C30, EORTC QLQ-LC13, and EuroQoL EQ-5D visual analogue scale (VAS) scores. Specifically:

    - Each EORTC QLQ-C30 functional scales: role functioning, emotional functioning, cognitive functioning, and social functioning

    - Each of three EORTC QLQ-C30 symptom scales (fatigue, nausea/vomiting, and pain), and single item measures (sleep disturbance, appetite loss, constipation, diarrhea, and financial difficulties)

    - Each EORTC QLQ-LC13 item: site-specific pain (pain in arm or shoulder, and pain in other parts), sore mouth, dysphagia, peripheral neuropathy, alopecia, and hemoptysis

    - The EQ-5D VAS

- Proportions of improvement / stability / stability + improvement / deterioration in EORTC QLQ-C30 global health status / QoL, physical functioning, and dyspnea scores, and EORTC QLQ-LC13 cough and chest pain scores where

    - Improvement

        - Improvement for the global health status/QoL and physical functioning scores is defined by an increase of 10 points or more in score from baseline at any time during the study and confirmed by an increase of 10 points or more in score at the next consecutive visit.

- Improvement for the dyspnea, cough, and chest pain scores is defined by a decrease of 10 points or more in score from baseline at any time during the study and confirmed by a decrease of 10 points or more in score at the next consecutive visit.

- Stability is defined as follows, when the criteria for improvement are not met:
  - an improvement confirmed by a less than 10 points change from baseline at the next consecutive visit
  - less than 10 points change in score confirmed by a less than 10 points change in score at the next consecutive visit
  - less than 10 points change in score and confirmed by an improvement at the next consecutive visit
- Stability + improvement is defined as the sum of proportions with stability and improvement as specified above.
- Deterioration is defined as a 10 points or greater worsening from baseline at any time during the study when the criteria for improvement or stability are not met.

3.5. Analysis Populations

3.5.1. Efficacy Analysis Populations

The analyses of the primary efficacy endpoints are based on the ITT population. All randomized participants will be included in this population. Participants will be analyzed in the treatment group to which they were randomized. Based on actual enrollment, all participants who were randomized on or prior to January 28, 2021 will be included in the ITT population of the global study. All participants who failed screening in addition to all randomized participants on or prior to January 28, 2021will be included in the screening population. Details on the approach to handling missing data are provided in Section 3.6.

3.5.2. Safety Analysis Populations

Safety analyses will be conducted in the APaT population, which consists of all randomized participants who received at least 1 dose of study intervention. Participants will be included in the treatment group corresponding to the study intervention they actually received for the analysis of safety data using the APaT population. This will be the treatment group to which they are randomized, except for participants who take incorrect study intervention for the entire treatment period; such participants will be included in the treatment group corresponding to the study intervention actually received. Any participant who receives the incorrect study intervention for 1 cycle, but receives the randomized treatment for all other cycles, will be analyzed according to the randomized treatment group, and a narrative will be provided for any events that occur during the cycle for which the participant is incorrectly dosed.

At least 1 laboratory, vital sign, or ECG measurement obtained subsequent to at least 1 dose of study intervention is required for inclusion in the analysis of the respective safety parameter. To assess change from baseline, a baseline measurement is also required.

3.5.3. Patient-reported Outcome Analysis Population

The PRO analyses are based on the PRO Full Analysis Set (PRO FAS) population, defined as all randomized participants who have at least one PRO assessment available and have received at least one dose of the study intervention. Participants will be analyzed in the treatment group to which they were randomized.

3.5.4. Pharmacokinetic Analysis Population

The population of PK analysis set includes all the participants who have received at least 1 dose of study intervention with documented dosing history in the lenvatinib + pembrolizumab arm and have measurable plasma levels of lenvatinib or serum levels of pembrolizumab.

3.6. Statistical Methods

3.6.1. Statistical Methods for Efficacy Analyses

This section describes the statistical methods that address the primary and secondary efficacy objectives. Efficacy results that will be deemed to be statistically significant after consideration of the Type I error control strategy are described in Section 3.8. Nominal *p* values will be computed for other efficacy analyses, but should be interpreted with caution because of potential issues of multiplicity.

The stratification factors used for randomization (see Section 6.3.2 in the protocol) will be applied to all stratified analyses, in particular, the stratified log-rank test, stratified Cox model, and stratified Miettinen and Nurminen method [Miettinen, O. 1985].

The efficacy analyses for OR, DOR and PFS will include responses and documented progression events that occur prior to second course treatment.

A summary of the primary analysis strategy for the key efficacy endpoints is provided in Table 1.

Table 1 Analysis Methods for Key Efficacy Endpoints

Endpoint/Variable	Statistical Method	Analysis Population	Missing Data Approach
Primary Analyses:			
PFS (RECIST 1.1) by BICR	Testing: stratified log-rank test Estimation: Stratified Cox model with Efron's tie-handling method	ITT	Censored according to rules in Table 2
OS	Testing: stratified log-rank test Estimation: Stratified Cox model with Efron's tie-handling method	ITT	Censored at last known alive date
Secondary Analyses:			
ORR (RECIST 1.1) by BICR	Testing and estimation: stratified Miettinen and Nurminen method	ITT	Participants with missing data are considered nonresponders
Abbreviations: BICR = blinded independent central review; ITT = intention-to-treat; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors.			

The strategy to address multiplicity issues with regard to multiple endpoints and interim analyses is described in Section 3.7, Interim Analyses, and in Section 3.8, Multiplicity.

3.6.1.1. Progression-free Survival

The nonparametric Kaplan-Meier method will be used to estimate the PFS curve in each treatment group. The hypotheses of treatment difference in PFS will be tested by the stratified log-rank test. A stratified Cox proportional hazard model with Efron's method of tie handling will be used to estimate the magnitude of the treatment difference (ie, HR) between the treatment arms. The HR and its 95% CI from the stratified Cox model with Efron's method of tie handling and with a single treatment covariate will be reported. The stratification factors used for randomization (Section 6.3.2 in the protocol) will be applied to both the stratified log-rank test and the stratified Cox model.

Since PD is assessed periodically, PD can occur any time in the time interval between the last assessment where PD was not documented and the assessment when PD is documented. For the primary analysis, for the participants who have PD, the true date of PD will be approximated by the date of the first assessment at which PD is objectively documented per RECIST 1.1 by the BICR vendor, regardless of discontinuation of study drug.

To evaluate the robustness of the PFS endpoint per RECIST 1.1 via BICR by the imaging vendor, 1 primary and 2 sensitivity analyses with a different set of censoring rules will be performed. For the primary analysis, if the events (PD or death) are immediately after more than 1 consecutive missed disease assessment, the data are censored at the last disease assessment before missing visits. Also data after new anticancer therapy are censored at the last disease assessment before the initiation of new anticancer therapy. The first sensitivity analysis follows ITT principles (ie, PDs/deaths are counted as events regardless of missed study visits or initiation

of new anticancer therapy). The second sensitivity analysis considers discontinuation of treatment or initiation of an anticancer treatment subsequent to discontinuation of study-specified treatments, whichever occurs later, to be a PD event for participants without documented PD or death. If a participant meets multiple criteria for censoring, the censoring criterion that occurs earliest will be applied. The censoring rules for primary and sensitivity analyses are summarized in Table 2.

Table 2 Censoring Rules for Primary and Sensitivity Analyses of Progression-free Survival

Situation	Primary Analysis	Sensitivity Analysis 1	Sensitivity Analysis 2
No PD and no death; new anticancer treatment is not initiated	Censored at last disease assessment	Censored at last disease assessment	Progressed at treatment discontinuation for reasons other than CR; otherwise, censored at last disease assessment if still on study intervention or completed study intervention
No PD and no death; new anticancer treatment is initiated	Censored at last disease assessment before new anticancer treatment	Censored at last disease assessment	Progressed at date of new anticancer treatment
PD or death documented after ≤1 missed disease assessment and before new anticancer therapy	Progressed at date of documented PD or death	Progressed at date of documented PD or death	Progressed at date of documented PD or death
PD or death documented immediately after ≥2 consecutive missed disease assessments or after new anticancer therapy	Censored at last disease assessment before the earlier date of ≥2 consecutive missed disease assessment and new anticancer therapy	Progressed at date of documented PD or death	Progressed at date of documented PD or death
Abbreviation: CR = complete response; PD = progressive disease.			

In case the proportional hazards assumption is not valid, Restricted Mean Survival Time (RMST) method may be conducted for PFS to account for the possible non-proportional hazards effect as a sensitivity analysis.

3.6.1.2. Overall Survival

The nonparametric Kaplan-Meier method will be used to estimate the survival curves. The treatment difference in survival will be assessed by the stratified log-rank test (based on the stratification factors defined in Section 6.3.2 in the protocol). A stratified Cox proportional hazard model with Efron's method of tie handling will be used to assess the magnitude of the treatment difference (ie, the HR). The HR and its 95% CI from the stratified Cox model with a single treatment covariate will be reported. The stratification factors used for randomization (Section 6.3.2 in the protocol) will be applied to both the stratified log-rank test and the stratified

Cox model. Participants without documented death at the time of analysis will be censored at the date of last known contact.

In case the proportional hazards assumption is not valid, RMST method may be conducted for OS to account for the possible non-proportional hazards effect as a sensitivity analysis.

3.6.1.3. Objective Response Rate

The stratified Miettinen and Nurminen method will be used for comparison of ORR between the treatment groups. The difference in ORR and its 95% CI from the stratified Miettinen and Nurminen method with strata weighting by sample size will be provided. The same stratification factors used for randomization (Section 6.3.2 in the protocol) will be used as stratification factors in the analysis.

3.6.1.4. Duration of Response

For subjects who demonstrate CR or PR, DOR is defined as the time from first documented evidence of CR or PR until PD or death due to any cause, whichever occurs first.

The nonparametric Kaplan-Meier method will be used to summarize the DOR. The median and range of DOR will be provided.

Censoring rules for DOR are summarized in Table 3.

Table 3 Censoring Rules for Duration of Response

Situation	Date of Progression or Censoring	Outcome
No PD or death, no new anticancer therapy initiated	Last adequate disease assessment	Censor (nonevent)
No PD or death, new anticancer therapy initiated	Last adequate disease assessment before new anticancer therapy initiated	Censor (nonevent)
Death or PD immediately after ≥2 consecutive missed disease assessments or after new anticancer therapy, if any	Earlier date of last adequate disease assessment before ≥2 missed adequate disease assessments and new anticancer therapy, if any	Censor (nonevent)
Death or PD after ≤1 missed disease assessments and before new anticancer therapy, if any	PD or death	End of response (event)
Abbreviation: PD = progressive disease. Note: A missed disease assessment includes any assessment that is not obtained or is considered inadequate for evaluation of response.		

3.6.2. Statistical Methods for Safety Analyses

Safety and tolerability will be assessed by clinical review of all relevant parameters including AEs, laboratory tests, vital signs, and ECG measurements.

The analysis of safety results will follow a tiered approach (Table *4*). The tiers differ with respect to the analyses that will be performed. AEs (specific terms as well as system organ class terms) and events that meet predefined limits of change in laboratory, vital signs, and ECG parameters are either prespecified as "Tier 1" endpoints or will be classified as belonging to "Tier 2" or "Tier 3," based on observed proportions of participants with an event.

Tier 1 Events

Safety parameters or AEs of special interest that are identified a priori constitute "Tier 1" safety endpoints that will be subject to inferential testing for statistical significance. AEs that are immune-mediated or potentially immune-mediated are well documented and will be evaluated separately; however, these events have been characterized consistently throughout the pembrolizumab clinical development program, and determination of statistical significance is not expected to add value to the safety evaluation. Similarly, the combination of pembrolizumab and lenvatinib has not been associated with any new safety signals. Finally, there are no known AEs associated with participants with NSCLC for which determination of a *p* value is expected to impact the safety assessment. Therefore, there are no Tier 1 events for this protocol.

Tier 2 Events

Tier 2 parameters will be assessed via point estimates with 95% CIs provided for differences in the proportion of participants with events using the Miettinen and Nurminen method, an unconditional, asymptotic method [Miettinen, O. 1985].

Membership in Tier 2 requires that at least 10% of participants in any treatment group exhibit the event; all other AEs and predefined limits of change will belong to Tier 3. The threshold of at least 10% of participants was chosen for Tier 2 events because the population randomized in this study are in critical conditions and usually experience various AEs of similar types regardless of treatment; events reported less frequently than 10% of participants would obscure the assessment of the overall safety profile and add little to the interpretation of potentially meaningful treatment differences. In addition, Grade 3 to 5 AEs (≥5% of participants in one of the treatment groups) and SAEs (≥5% of participants in one of the treatment groups) will be considered Tier 2 endpoints. Because many 95% CIs may be provided without adjustment for multiplicity, the CIs should be regarded as a helpful descriptive measure to be used in safety review, not a formal method for assessing the statistical significance of the between-group differences.

Tier 3 Events

Safety endpoints that are not Tier 1 or 2 events are considered Tier 3 events. Only point estimates by treatment group are provided for Tier 3 safety parameters.

Continuous Safety Measures

For continuous measures such as changes from baseline in laboratory, vital signs, and ECG parameters, summary statistics for baseline, on-treatment, and change from baseline values will be provided by treatment group in table format.

Table 4 Analysis Strategy for Safety Parameters

Safety Tier	Safety Endpoint	p-Value	95% CI for Treatment Comparison	Descriptive Statistics
Tier 2	Grade 3-5 AE (incidence ≥5% of participants in one of the treatment groups)		X	X
	Serious AE (incidence ≥5% of participants in one of the treatment groups)		X	X
	AEs (incidence ≥10% of participants in one of the treatment groups)		X	X
Tier 3	Any AE			X
	Any Grade 3-5 AE			X
	Any Serious AE			X
	Any Drug-Related AE			X
	Any Serious and Drug-Related AE			X
	Any Grade 3-5 and Drug-Related AE			X
	Discontinuation due to AE			X
	Death			X
	Specific AEs, SOCs (incidence <10% of participants in all of the treatment groups)			X
	Change from Baseline Results (lab toxicity shift)			X
Abbreviations: AE = adverse event; CI = confidence interval; SOC = system organ class.				

Exposure-Adjusted Approach

To properly account for the potential difference in follow-up time between the study arms, AE incidence adjusted for treatment exposure analyses may be performed as appropriate.

Time to Grade 3-5 AE

In addition to the tiered approach, exploratory analysis may be performed on the time to the first Grade 3-5 AE. The time to the first Grade 3-5 AE is defined as the time from the first day of study drug to the first event of a Grade 3-5 AE. Summary statistics will be provided.

3.6.3. Statistical Methods for Patient-reported Outcome Analyses

3.6.3.1. Scoring Algorithm

EORTC QLQ-C30:

For each scale or item, a linear transformation will be applied to standardize the score as between 0 and 100, according to the corresponding scoring standard. For global health status/quality of life and all functional scales, a higher value indicates a better level of function; for symptom scales and items, a higher value indicates increased severity of symptoms.

The scoring method for each subscale of the EORTC-QLQ-C30 will be performed according to the scoring manual (Scott, et al., 2008). According to the EORTC QLQ-C30 manual, if items $I_1, I_2, \cdots, I_n$ are included in a scale, the linear transformation procedure is as follows:

1. Compute the raw score:

$$RS = (I_1 + I_1 + \cdots + I_n)/n$$

2. Linear transformation to obtain the score *S*:

$$\text{Functional scales: } S = \left(1 - \frac{RS-1}{Range}\right) \times 100$$

$$\text{Symptom scales/items: } S = \left(1 - \frac{RS-1}{Range}\right) \times 100$$

$$\text{Global health status/quality of life scale: } S = \left(1 - \frac{RS-1}{Range}\right) \times 100$$

The range is defined as the difference between the maximum possible value of *RS* and the minimum possible value for RS. If more than half of the items within one scale are missing, then the scale is considered missing; otherwise, the score will be calculated as the average score of those available items.

EORTC QLQ-LC13:

The lung cancer questionnaire module is a validated self-reported PRO questionnaire that is intended to be used in conjunction with the QLQ-C30 and includes 13 questions that share the four level ordinal response with the QLQ-C30 and can be summarized and scored as described above (1= “not at all”; 2= “a little”; 3= “quite a bit”; 4= “very much”). The LC13 comprises both multi-item and single-item measures of lung cancer-associated symptoms (i.e. coughing, hemoptysis, dyspnea, and pain) and treatment related symptoms (i.e., sore mouth, dysphagia, peripheral neuropathy, and alopecia). A linear transformation will be applied to standardize the scores between 0 (least severe symptom) and 100 (most severe symptom) as described above for the EORTC QLQ-C30 symptom scales/items scoring.

EQ-5D:

The EQ-5D-5L is primarily designed for self-completion and consists of 2 parts: a descriptive system and the VAS. The EQ-5D-5L descriptive system includes 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The 5 dimensions each have 5 levels: no problems, slight, moderate, severe problems, and extreme problems. The responses patterns on the 5 dimensions are scored using country-specific population weights to provide an aggregate index score anchored at 0 (death) and 1 (perfect health); depending on the algorithm used some states may be considered worse than death.

The EQ-5D VAS records the respondent’s self-rated health on a vertical, visual analogue scale (0-100), with endpoints labeled ‘the best health you can imagine’ and ‘the worst health you can

imagine'. The recall period is current health today (the day of completion). For the EQ-5D VAS scale, A ≥7-point change from baseline in VAS is considered to be a MID (Packard et al., 2007).

3.6.3.2. PRO Compliance Summary

Completion and compliance of EORTC QLQ-C30, EORTC-QLQ-LC13, and EQ-5D by visit and by treatment will be described. Numbers and percentages of complete and missing data at each visit will be summarized. An instrument is considered complete if at least one valid score is available according to the missing item rules outlined in the scoring manual for each respective instrument.

Completion rate of treated participants (CR-T) at a specific time point is defined as the number of treated participants who complete at least one item over the number of treated participants in the PRO analysis population.

$$\text{CR-T} = \frac{Number\ of\ treated\ participants\ who\ complete\ at\ least\ one\ item}{Number\ of\ treated\ participants\ in\ the\ PRO\ analysis\ population}$$

The completion rate is expected to shrink in the later visit during study period due to the participants who discontinued early. Therefore, another measurement, compliance rate of eligible participants (CR-E) will also be employed as the support for completion rate. CR-E is defined as the number of treated participants who complete at least one item over number of eligible participants who are expected to complete the PRO assessment, not including the participants missing by design such as death, discontinuation, translation not available.

$$\text{CR-E} = \frac{Number\ of\ treated\ participants\ who\ complete\ at\ least\ one\ item}{Number\ of\ eligible\ participants\ who\ are\ expected\ to\ complete}$$

The reasons of non-completion and non-compliance will be provided in supplementary table:

- Completed as scheduled
- Not completed as scheduled
- Off-study: not scheduled to be completed.

In addition, reasons for non-completion as scheduled of these measures will be collected using "miss_mode" forms filled by site personnel and will be summarized in table format. The schedule (study visits and estimated study times) and mapping of study visit to analysis visit for PRO data collection is provided in Table 5. If there are multiple PRO collections within any of the stated time windows, the assessment completed closest to the target collection day will be used in the analyses.

Table 5 PRO Data Collection Schedule and Mapping of Study visit to Analysis Visit

Treatment Week	Target Day	Range
0	1	-28 to 1
3	22	2 to 32
6	43	33 to 53
9	64	54 to 74
12	85	75 to 95
15	106	96 to 116
18	127	117 to 137
21	148	138 to 158
24	169	159 to 179
27	190	180 to 200
30	211	201 to 221
33	232	222 to 242
36	253	243 to 263
39	274	264 to 284
42	295	285 to 305
45	316	306 to 326
48	337	327 to 358
54	379	359 to 400
60	421	401 to 442
66	463	443 to 484
72	505	485 to 526
78	547	527 to 568
84	589	569 to 610
90	631	611 to 652
96	673	653 to 694
102	715	695 to 736

3.6.3.3. Mean Change from Baseline

The time point for the mean change from baseline analysis will be defined as the latest time point at which approximately CR-T ≥ 60% and CR-E ≥ 80% based on blinded data review prior to the database lock for any PRO analysis.

To assess the treatment effects on the PRO score change from baseline in the global health status/QoL, physical functioning and dyspnea scores of the EORTC QLQ-C30 and the cough and chest pain scores of the EORTC QLQ-LC13 scales a constrained longitudinal data analysis (cLDA) model proposed by Liang and Zeger [Liang, K-Y. and Zeger, S. L. 2000] will be applied, with the PRO score as the response variable, and treatment, time, the treatment by time interaction, and stratification factors used for randomization (See Section 6.3.2 in the protocol) as covariates. The treatment difference in terms of least square (LS) mean change from baseline will be estimated from this model together with 95% CI and nominal two-sided p-value. Model-based LS mean with 95% CI will be provided by treatment group for PRO scores at baseline and post-baseline time point.

The cLDA model assumes a common mean across treatment groups at baseline and a different mean for each treatment at each of the post-baseline time points. In this model, the response vector consists of baseline and the values observed at each post-baseline time point. Time is treated as a

categorical variable so that no restriction is imposed on the trajectory of the means over time. The cLDA model is specified as follows:

$$E(Y_{ijt}) = \gamma_0 + \gamma_{jt} I(t > 0) + \beta X_i, j = 1,2,,3,..,n;\ t = 0,1,2,3,..k$$

where Y_{ijt} is the PRO score for participant i, with treatment assignment j at visit t; γ_0 is the baseline mean for all treatment groups, γ_{jt} is the mean change from baseline for treatment group j at time t; X_i is the stratification factor (binary) vector for this participant, and β is the coefficient vector for stratification factors. An unstructured covariance matrix will be used to model the correlation among repeated measurements. If the unstructured covariance model fails to converge with the default algorithm, then Fisher scoring algorithm or other appropriate methods can be used to provide initial values of the covariance parameters. In the rare event that none of the above methods yield convergence, a structured covariance such as Toeplitz can be used to model the correlation among repeated measurements. In this case, the asymptotically unbiased sandwich variance estimator will be used. The cLDA model implicitly treats missing data as missing at random (MAR).

Line plots for the empirical mean change from baseline in the EORTC QLQ-C30 global health status/QoL, physical functioning and dyspnea scores and EORTC QLQ-LC13 cough and chest pain scores (as specified in section 3.4.3) will be provided across all time points as a supportive analysis.

In addition, the model-based LS mean change from baseline to the specified post-baseline time point together with 95% CI will be plotted in bar charts for the EORTC QLQ-C30 global health status/QoL, physical functioning and symptom scales, and EORTC QLQ-LC13 symptom scales (as specified in section 3.4.3).

3.6.3.4. Time-to-True Deterioration (TTD) Analysis

The non-parametric Kaplan-Meier method will be used to estimate the deterioration curve in each group. The estimate of median time to deterioration and its 95% confidence interval will be obtained from the Kaplan-Meier survival estimates. The treatment difference in TTD will be assessed by the stratified log-rank test, and nominal two-sided p-values will be reported. A stratified Cox proportional hazard model with Efron's method of tie handling will be used to assess the magnitude of the treatment difference (hazard ratio) between treatment arms. Stratification factors used for randomization (See Section 6.3.2 in the protocol) will be applied to the analysis will be used in the stratified Cox PH model.

The approach for the time-to-deterioration analysis will be based on the assumption of non-informative censoring. The participants who do not have deterioration on the last date of evaluation will be censored. Table 6 provides censoring rule for TTD analysis.

Table 6 Censoring Rules for Time-to-Deterioration

Scenario	Outcome
Deterioration documented	Event observed at time of assessment (first deterioration)
Ongoing, death or discontinued from study without deterioration	Right censored at time of last assessment
No baseline assessments	Right censored at treatment start date

3.6.3.5. Overall Improvement and Overall Improvement/Stability

Overall improvement rate will be analyzed, which is defined as the proportion of participants who have achieved an improvement as defined in Section 3.4.3 PRO Endpoints. The point estimate of overall proportions of participants who have achieved an improvement, stability, and deterioration will be provided by treatment group together with 95% CI using exact binomial method by Clopper and Pearson (1934). Stratified Miettinen and Nurminen's method will be used for comparison of the overall improvement rate between the treatment groups. The difference in overall improvement rate and its 95% CI, along with nominal two-sided p-values, from the stratified Miettinen and Nurminen's method with strata weighting by sample size will be provided. The stratification factors used for randomization (see Section 6.3.2 in protocol) will be applied to the analysis.

The same method will be used to analyze overall improvement/stability rate, which is defined as the proportion of participants who have achieved improvement/stability as defined in Section 3.4.3 PRO Endpoints.

3.6.3.6. Analysis Strategy for Key PRO Endpoints

A summary of the analysis strategy for the key PRO endpoints is provided in Table 7.

Table 7 Analysis Strategy for Key PRO Endpoints

Endpoint/Variable	Statistical Method	Analysis Population	Missing Data Approach
Mean change from baseline in – EORTC QLQ-C30 • Global health status/QoL • Physical functioning • dyspnea – EORTC QLQ-LC13 • Cough • Chest pain	cLDA model	FAS	Model-based.= cLDA
TTD in – EORTC QLQ-C30 • Global health status/QoL • Physical functioning • dyspnea – EORTC QLQ-LC13 • Cough • Chest pain	Stratified log-rank test and HR estimation using stratified Cox model with Efron's tie handling method	FAS	Censored according to rules in Table *6*.
Overall improvement and overall improvement/stability in – EORTC QLQ-C30 • Global health status/QoL • Physical functioning • dyspnea – EORTC QLQ-LC13 • Cough • Chest pain	Stratified Miettinen and Nurminen method	FAS	Participants with missing data are considered not achieving improvement/stability.
Abbreviations: cLDA = constrained longitudinal data analysis, FAS = full analysis set, QoL = quality of life.			

3.6.4. Demographic and Baseline Characteristics

The comparability of the treatment groups for each relevant demographic and baseline characteristic will be assessed by the use of tables and/or graphs. No statistical hypothesis tests will be performed on these characteristics. The number and percentage of participants screened and randomized and the primary reason for screening failure and discontinuation will be displayed. Demographic variables, baseline characteristics, primary and secondary diagnoses, and prior and concomitant therapies will be summarized by treatment, either by descriptive statistics or categorical tables.

3.7. Interim Analyses

Blinding to treatment assignment will be maintained at all investigational sites. The results of interim analyses will not be shared with the investigators before the completion of the study. Participant-level unblinding will be restricted to an external unblinded statistician and scientific

programmer performing the interim analysis, who will have no other responsibilities associated with the study.

An external DMC will serve as the primary reviewer of the results of the interim analyses of the study and will make recommendations for discontinuation of the study or protocol modifications to the study Executive Oversight Committee (EOC). If the DMC recommends modifications to the design of the protocol or discontinuation of the study, the EOC (and potentially other limited Sponsor personnel) may be unblinded to results at the treatment level to act on these recommendations. The extent to which individuals are unblinded with respect to results of interim analyses will be documented by the unblinded statistician. Additional logistical details will be provided in the DMC charter. Key aspects of the interim analyses are described below.

Treatment-level results from the interim analysis will be provided to the DMC by the unblinded statistician. The unblinded statistician will not be involved in any discussions regarding modifications to the protocol, statistical methods, identification of protocol deviations, or data validation efforts after the interim analyses.

If the study is positive at an interim analysis for both primary endpoints, additional analyses, including but not limited to the protocol-specified final analysis, may be carried out for exploratory purpose or upon regulatory request.

3.7.1. Efficacy Interim Analysis

One interim analysis is planned in addition to the final analysis for this study. For the interim and final analyses, all the randomized participants will be included. Results of the interim analysis will be reviewed by the DMC. Details on the boundaries for establishing statistical significance with regard to efficacy are discussed further in Section 3.8.

The analyses planned, endpoints evaluated, and drivers of timing are summarized in Table *8*.

Table 8 Summary of Interim and Final Analyses Strategy

Analyses	Key Endpoints	Timing	Estimated Time After First Participant Randomized	Primary Purpose of Analysis
IA	PFS OS	Both ~416 PFS events have been observed and ~8.8 months after last participant randomized	~30.5 months	• Final PFS analysis • Interim OS analysis
Final analysis	OS	Both ~388 deaths have occurred and ~18.5 months after last participant randomized	~40.2 months	• Final OS analysis
Note that for the IA and the FA, if the events accrue slower than expected such that the targeted number of events cannot be reached in the anticipated timeframe, the Sponsor may conduct the analysis with an additional 2 months and 7 months of follow-up for the IA and FA respectively, or the specified number of events is observed, whichever occurs first. Abbreviations: IA = interim analysis; OS = overall survival; PFS = progression-free survival.				

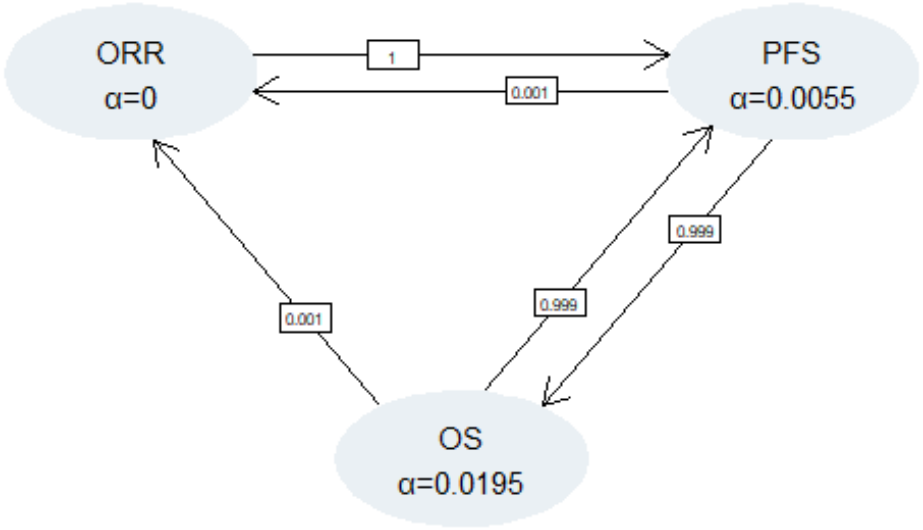
Non-binding futility analyses will be conducted in July 2021 (data cutoff in May 2021) to coincide with an anticipated safety DMC meeting and at the IA as per DMC recommendation. Further details of any evaluation are specified in the Appendix 5.1.

3.7.2. Safety Interim Analyses

The DMC will conduct regular safety monitoring. The timing of the safety monitoring will be specified in the DMC charter.

3.8. Multiplicity

The study uses the graphical method of Maurer and Bretz [Maurer, W., et al 2011] to control multiplicity for multiple hypotheses as well as interim analysis. According to this approach, study hypotheses may be tested more than once, and when a particular null hypothesis is rejected, the α allocated to that hypothesis can be reallocated to other hypothesis tests. Figure *1* shows the initial 1-sided α allocation for each hypothesis in the ellipse representing the hypothesis. The weights for reallocation from each hypothesis to the others are represented in the boxes on the lines connecting the hypotheses.



Abbreviations: ORR = objective response rate; OS = overall survival; PFS = progression-free survival.

Note: If both PFS and OS null hypotheses are rejected, the reallocation strategy allows retesting of ORR at α = 0.025.

Figure 1 Multiplicity Graph for Type I Error Control

3.8.1. Objective Response Rate

No initial alpha is allocated to test ORR. If superiorities for both the PFS and OS hypotheses are declared, α = 0.025 will be rolled over to the hypothesis for ORR and the test statistics computed at the IA for the ORR hypothesis will be used for inferential testing with an updated alpha level of 0.025. The power at the α-level of 0.025 is 99.9%, with an approximate treatment difference (ΔORR) required for reaching the efficacy bound being 7.5%, assuming underlying 28% and 48% ORR in the control and experimental groups, respectively.

3.8.2. Progression-free Survival

The study will test PFS only at the interim analysis, at an initial α level of 0.0055. Table 9 shows the boundary properties for each of these α-levels for the PFS analysis. Note that the final row indicates the total power to reject the null hypothesis for PFS at each α level. Also note that if the OS null hypothesis is rejected at the FA, the previously computed PFS test statistics may be used for inferential testing with its updated bounds considering the α re-allocation from the OS hypothesis. Following the multiplicity strategy as outlined in Figure 1, if superiority for both the ORR and OS hypotheses are declared, α = 0.025 will be rolled over to the hypothesis for PFS.

Table 9 Efficacy Boundaries and Properties for Progression-free Survival Analyses

Analysis	Value	α=0.0055	α=0.025
IA: 100%[a] N: 620 Events: 416 Month: 30.5[f]	Z	2.5427	1.9600
	p (1-sided)[b]	0.0055	0.0250
	~HR at bound[c]	0.7793	0.8252
	P(Cross) if HR=1[d]	0.0055	0.0250
	P(Cross) if HR=0.7[e]	0.8650	0.9536

Abbreviations: HR = hazard ratio; IA = interim analysis.
The number of events and timings are estimated.
[a] Percentage of total planned events at the interim analysis.
[b] The nominal α for testing.
[c] The approximate HR required to reach an efficacy bound
[d] The probability of crossing a bound under the null hypothesis
[e] The probability of crossing a bound under the alternative hypothesis
[f] The approximate number of months since first participant randomized

3.8.3. Overall Survival

The study will test OS at IA and FA. Following the multiplicity strategy as outlined in Figure *1*, the OS hypothesis may be tested at α=0.0195 (initially allocated α) or α=0.025 (if the null hypotheses for both PFS and ORR are rejected). Table 10 shows the boundary properties for each of these α-levels for OS at IA and FA, which were derived using a Lan-DeMets and O'Brien-Fleming spending function. Note that if the PFS null hypothesis is rejected at IA, OS interim and final analysis test may be compared with its updated bounds considering the α reallocation from the PFS hypothesis.

Table 10 Efficacy Boundaries and Properties for Overall Survival Analyses

Analysis	**Value**	**α=0.0195**	**α=0.025**
IA: 78%[a] N: 620 Events: 302 Month:30.5 [f]	Z	2.4029	2.2869
	p (1-sided)[b]	0.0081	0.0111
	~HR at bound[c]	0.7584	0.7686
	P(Cross) if HR=1[d]	0.0081	0.0111
	P(Cross) if HR=0.71[e]	0.7173	0.7548
Final N: 620 Events: 388 Month: 40.2 [f]	Z	2.1181	2.0193
	p (1-sided)[b]	0.0171	0.0217
	~HR at bound[c]	0.8064	0.8145
	P(Cross) if HR=1[d]	0.0195	0.0250
	P(Cross) if HR=0.71[e]	0.9000	0.9165

Abbreviations: HR = hazard ratio; IA = interim analysis.
The number of events and timings are estimated.

[a] Percentage of total planned events at each interim analysis.
[b] The nominal α for group sequential testing.
[c] The approximate HR required to reach an efficacy bound
[d] The probability of crossing a bound under the null hypothesis
[e] The probability of crossing a bound under the alternative hypothesis
[f] The approximate number of months since first participant randomized

The bounds provided in Table *10* are based on the assumption that the expected number of events at the IA and FA for OS are 302 and 388, respectively. At the time of an analysis, the observed number of events may differ substantially from the expected. To avoid overspending at an interim analysis and leave reasonable alpha for the final analysis, the minimum alpha spending strategy will be adopted. At an IA, the information fraction used in Lan-DeMets spending function to determine the alpha spending at the IA will be based on the minimum of the expected information fraction and the actual information fraction at each analysis. Specifically,

- In the scenario that the events accrue slower than expected and the observed number of events is less than the expected number of events at a given analysis, the information fraction will be calculated as the observed number of events at the interim analysis over the target number of events at FA.

- In the scenario that the events accrue faster than expected and the observed number of events exceeds the expected number of events at a given analysis, then the information fraction will be calculated as the expected number of events at the interim analysis over the target number of events at FA.

The final analysis will use the remaining Type I error that has not been spent at the earlier analyses. The event counts for all analyses will be used to compute correlations.

Of note, while the information fraction used for alpha spending calculation will be the minimum of the actual information fraction and the expected information fraction, the correlations required for deriving the bounds will still be computed using the actual information fraction based on the observed number of events at each analysis over the target number of events at FA.

The minimum spending approach assumes timing is not based on any observed Z-value and thus the Z test statistics used for testing conditioned on timing are multivariate normal. Given the probabilities derived with the proposed spending method, the correlations based on actual event counts are used to compute bounds that control the Type I error at the specified alpha level for a given hypothesis conditioned on the interim analysis timing. Since this is true regardless of what is conditioned on, the overall Type I error for a given hypothesis unconditionally is controlled at the specified level. By using more conservative spending early in the study, power can be retained to detect situations where the treatment effect may be delayed.

3.8.4. Safety Analyses

The DMC has responsibility for assessment of overall risk:benefit. When prompted by safety concerns, the DMC can request corresponding efficacy data. DMC review of efficacy data to assess the overall risk:benefit to study participants will not require a multiplicity adjustment typically associated with a planned efficacy interim analysis.

3.9. Sample Size and Power Calculations

The study will randomize 620 participants in a 1:1 ratio into the pembrolizumab + lenvatinib and pembrolizumab + placebo arms. PFS and OS are primary endpoints for the study, with OR as the secondary endpoint.

For the PFS endpoint, based on a target number of 416 events at the final PFS analysis , the study has approximately 86.5% power to detect an HR of 0.7 at the initially allocated α=0.0055 (1-sided).

For the OS endpoint, based on a target number of 388 events and one interim analysis at approximately 78% of the target number of events, the study has approximately 90% power to detect an HR of 0.71 at the initially allocated α=0.0195 (1-sided).

Based on KEYNOTE-042 data, the above sample size and power calculations for PFS and OS assume the following:

- PFS follows a piecewise exponential distribution, with a median of 5 months up to 6.5 months and then a median of 12 months thereafter for the control group
- OS follows an exponential distribution, with a median of 16.7 months for the control group
- Enrollment period is approximately 21.7 months
- Annual drop-out rate is 20% and 2% for PFS and OS, respectively

- Follow-up period is approximately 8.8 and 18.5 months for PFS and OS, respectively, after the last participant is randomized

The sample size and power calculations were performed in R (package "gsDesign") and EAST 6.4.

3.10. Subgroup Analyses

To determine whether the treatment effect is consistent across various subgroups, the between-group treatment effect for OS, PFS, and OR (with a nominal 95% CI) will be estimated and plotted by treatment group within each category of the following subgroup variables:

- Geographic region (East Asia, non-East Asia)
- ECOG performance status (0, 1)
- Predominant tumor histology (squamous, nonsquamous)
- TPS (1%-49%, ≥50%)
- Age category (<65, ≥65 years)
- Sex (female, male)
- Race (white, nonwhite)
- Smoking status (never, former/current smoker)
- Brain metastasis at baseline (presence, absence)
- Liver metastasis at baseline (presence, absence)

The consistency of the treatment effect will be assessed using descriptive statistics for each category of the subgroup variables listed above. If the number of participants in a category of a subgroup variable is less than 10% of the ITT population, the subgroup analysis will not be performed for this category of the subgroup variable, and this subgroup variable will not be displayed in the forest plot. The subgroup analyses for PFS and OS will be conducted using unstratified Cox model, and the subgroup analyses for OR will be conducted using unstratified Miettinen and Nurminen method.

3.11. Compliance (Medication Adherence)

Drug accountability data for study intervention will be collected during the study. Any deviation from protocol-directed administration will be reported.

3.12. Extent of Exposure

Extent of exposure for a participant is defined as number of cycles in which the participant receives the study intervention. Summary statistics will be provided on extent of exposure for the APaT population.

4. STATISTICAL ANALYSIS PLAN FOR EXTENSION

This section outlines the statistical analysis strategy and procedures for China specific subgroup analysis, which is required by the Chinese regulatory authority.

4.1. Introduction

After the global study enrollment is closed, subjects from China will continue to be enrolled in an extension portion designed to meet China local registration needs. The extension portion will be identical to the global study (e.g., inclusion and exclusion criteria, primary and secondary endpoints, study procedures) in general, with the additional statistical analysis plan for the Chinese subpopulation. The purpose of this extension portion is to evaluate the consistency of efficacy and safety in the Chinese subpopulation to the global population. Country-specific analysis may also be conducted per local regulatory requirement.

After the enrollment for the global study is completed, subjects in China will continue to be enrolled in a 1:1 ratio into the pembrolizumab + lenvatinib arm and pembrolizumab + placebo arm until the sample size for the Chinese subpopulation reaches approximately 120 in total between the global and extension portion.

After the cut-off date for the primary analyses of the global study (including interim and final analyses), all Chinese subjects, including subjects enrolled in the global study and the extension portion, will continue their randomized treatment and continue to be followed up for PFS and OS events for China registration purpose. The extension portion will be completed after target number of PFS or OS events (Section 4.7) has been observed between the two arms in the Chinese subpopulation.

However, if the target number of PFS or OS events in the Chinese subpopulation is reached before an IA for the global study, the corresponding analysis for Chinese subpopulation will occur at the same time as the global IA or the final analysis (global study).

4.2. Responsibility for Analyses/In-House Blinding

For all Chinese subjects, including subjects randomized in the global study and the extension portion, patient level treatment randomization information will be blinded for the statistician(s)/programmer(s) responsible for the China extension portion analysis until the extension portion data base lock is achieved. The extent to which individuals are unblinded to the results will be limited. Blinded and unblinded members will be clearly documented with blinding status along with time information.

4.3. Hypotheses/Estimation

No hypothesis testing is planned for the China extension portion. Nominal p values will be reported for efficacy analyses if applicable but should be interpreted with caution due to potential issues of multiplicity.

After succeeding in the global study, the consistency of efficacy and safety in the Chinese subpopulation to the global population will be evaluated. Consistency of efficacy will be evaluated using the percentage of risk reduction preserved in the Chinese subpopulation from the empirical

risk reduction from the global primary efficacy analyses (based on point estimates). Sample size is designed to provide about 80% chance of observing the point estimate of Chinese subpopulation preserves ≥ approximately 50% of empirical risk reduction from the global primary efficacy analysis assuming the same hazard ratio used in the sample size and power calculation for the global study.

4.4. The Analysis Endpoints

4.4.1. Efficacy Endpoints

4.4.1.1. Primary Endpoints

***PFS*:** the time from randomization to the first documented PD per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR or death due to any cause, whichever occurs first. See Section 3.6.1 for the definition of censoring.

***OS*:** the time from randomization to death due to any cause.

4.4.1.2. Secondary Endpoints

***OR*:** a confirmed CR or PR per RECIST 1.1, adjusted to follow a maximum of 10 target lesions and a maximum of 5 target lesions per organ, based on BICR.

4.4.1.3. Exploratory Endpoints

***DOR*:** the time from the earliest date of qualifying response until earliest date of PD or death from any cause, whichever comes first.

4.4.2. Safety Endpoints

Safety and tolerability will be assessed by clinical review of all relevant parameters including AEs, SAEs, fatal AEs, laboratory tests, and vital signs. Furthermore, specific events will be collected and designated as ECIs as described in Section 8.4.7 of the study protocol.

4.4.3. Patient-reported Outcome Endpoints

Secondary

- Mean change from baseline in EORTC QLQ-C30 global health status/QoL (items 29-30), physical functioning (items 1-5), and dyspnea scores (item 8), and EORTC QLQ-LC13 cough (item 31) and chest pain (item 40) scores.
- TTD as measured by each of the EORTC QLQ-C30 global health status/QoL, physical functioning, and dyspnea scores, and EORTC QLQ-LC13 cough and chest pain scores.
- TTD as measured by a composite symptom endpoint: cough (QLQ-LC13 item 31), chest pain (QLQ-LC13 item 40), or dyspnea (QLQ-C30 item 8)

Based on prior literature (Osoba D et al., 1998; King et al., 1996, Marringwa et al., 2011), a ≥10 point or greater worsening from baseline for each scale represents a minimally important difference (MID) that represents a clinically relevant deterioration. TTD in the individual scores specified above is defined as the time to first onset of ≥10 (out of 100) point deterioration from baseline in a given scale/subscale/item and confirmed by a second adjacent ≥10 point deterioration from baseline. TTD in the composite endpoint is defined as the time to first onset of a ≥10 point deterioration (i.e., increase in score) from baseline in any one of the 3 lung cancer symptom scale items (EORTC-QLQ-C30 [dyspnea] and EORTC-QLQ-LC13 [cough and chest pain]) with confirmation by the subsequent visit of a ≥10 point deterioration (i.e., increase in score) from baseline in the same scale as the first onset. Changes from baseline in EORTC QLQ-C30 scores will also be interpreted according to recent subscale-specific guidelines, which indicate that clinically meaningful differences vary by scale (Cocks et al., 2012)

4.5. Analysis Populations

4.5.1. Efficacy Analysis Populations

Efficacy analysis will be carried out in the intention-to-treat (ITT) China subpopulation. This population will include all Chinese participants who are randomized in the global study and all participants who are randomized in the extension portion.

Participants from China enrolled in the extension portion of this study after completion of the global enrollment will not be included in the primary efficacy analysis population for the global study.

4.5.2. Safety Analysis Populations

Safety analysis will be carried out in the All Patients as Treated (APaT) China subpopulation, i.e., all randomized Chinese participants (in the global study and extension portion) who received at least 1 dose of study treatment.

Participants from China randomized and treated in the extension portion of this study after completion of the global enrollment will not be included in the primary safety analysis population for the global study.

4.5.3. Patient-reported Outcome Analysis Population

The PRO analyses are based on the PRO Full Analysis Set (PRO FAS) China subpopulation, defined as all randomized participants who have at least one PRO assessment available for the specific endpoint and have received at least one dose of the study intervention. Participants will be analyzed in the treatment group to which they are randomized. The PRO FAS China subpopulation will include all Chinese participants (in the global study and extension portion) in this population.

Participants from China enrolled in the extension portion of this study after completion of the global enrollment will not be included in the primary PRO analysis population for the global study.

4.5.4. Pharmacokinetic Analysis Population

The analysis of PK data will be carried out in the population PK analysis set China subpopulation. The Population PK Analysis set will include all participants who have received at least 1 dose of study intervention with documented dosing history in the lenvatinib + pembrolizumab arm and have measurable plasma levels of lenvatinib or serum levels of pembrolizumab. Population PK Analysis Set China subpopulation will include all Chinese participants (in the global study and extension portion) in this population.

Participants from China enrolled in the extension portion of this study after completion of the global enrollment will not be included in the primary PK analysis population for the global study.

4.6. Statistical Methods

Regarding the analysis for extension, no formal hypothesis testing is planned. No multiplicity adjustment will be applied to the analysis for extension.

4.6.1. Statistical Methods for Efficacy Analyses

Analyses regarding objective response and disease progression will include all events in all participants excluding events that occur in the Second Course Treatment.

4.6.1.1. Progression-free Survival

Analysis of PFS for extension is the same to that for the global study if applicable.

The nonparametric Kaplan-Meier method will be used to estimate the PFS curve in each treatment group. The treatment difference in PFS will be assessed by the stratified log-rank test (based on the stratification factors defined in Section 6.3.2 in the protocol). . A stratified Cox proportional hazard model with Efron's method of tie handling will be used to estimate the magnitude of the treatment difference (ie, HR) between the treatment arms. The HR and its 95% CI from the stratified Cox model with Efron's method of tie handling and with a single treatment covariate will be reported. The same stratification factors used in the global study will be used (see section 6.3.2 in the protocol). For the Chinese subgroup analyses, the stratified method will only be used if applicable. The factor of geographic region (East Asia vs. non-East Asia) will not be included in the Chinese subgroup stratified analyses. An analysis using the restricted mean survival time method may be conducted for PFS to account for the possible nonproportional hazards effect.

Consistency of efficacy will be evaluated using the percentage of risk reduction preserved in the Chinese subpopulation from the empirical risk reduction from the global primary efficacy analyses (based on point estimates). Sample size is designed to provide about 80% chance of observing the point estimate of Chinese subpopulation preserves ≥ approximately 50% of empirical risk

reduction from the global primary efficacy analysis assuming the same hazard ratio used in the sample size and power calculation for the global study. The primary analysis for PFS will be conducted in the Chinese subpopulation when approximately 75 PFS events have been collected.

4.6.1.2. Overall Survival

Analysis of OS for extension is the same to that for the global study if applicable.

The nonparametric Kaplan-Meier method will be used to estimate the survival curves. The treatment difference in OS will be assessed by the stratified log-rank test (based on the stratification factors defined in Section 6.3.2 in the protocol). A stratified Cox proportional hazard model with Efron's method of tie handling will be used to assess the magnitude of the treatment difference (ie, the HR). The HR and its 95% CI from the stratified Cox model with a single treatment covariate will be reported. The stratification factors used for randomization (Section 6.3.2 in the protocol) will be applied to both the stratified log-rank test and the stratified Cox model. The same stratification factors used in the global study will be used. For the Chinese subgroup analysis, the stratified method will only be used if applicable. The factor of Geography (East Asia vs. non-East Asia) will not be included in the stratified analysis for the Chinese subgroup analysis. Participants without documented death at the time of analysis will be censored at the date of last known contact. An analysis using the restricted mean survival time method may be conducted for OS to account for the possible nonproportional hazards effect.

Consistency in OS will be evaluated similarly as that in PFS. The primary analysis for OS will be conducted in the Chinese subpopulation when approximately 75 OS events have been collected.

4.6.1.3. Objective Response Rate

Analysis of ORR or DOR for extension is the same to that for the global study if applicable.

The stratified Miettinen and Nurminen method will be used for comparison of ORR between the treatment groups. The difference in ORR and its 95% CI from the stratified Miettinen and Nurminen method with strata weighting by sample size will be provided. The same stratification factors used for randomization (Section 6.3.2 of the study protocol) will be used as stratification factors in the analysis. The same stratification factors used in the global study will be used. For the Chinese subpopulation analysis, the stratified method will only be used if applicable. The factor of Geography (East Asia vs. non-East Asia) will not be included in the stratified analysis for Chinese subgroup analysis.

4.6.1.4. Duration of Response

For subjects who demonstrate CR or PR, DOR is defined as the time from first documented evidence of CR or PR until PD or death due to any cause, whichever occurs first.

The nonparametric Kaplan-Meier method will be used to summarize the DOR. The median and range of DOR will be provided.

4.6.2. Statistical Methods for Safety Analyses

Safety analyses for extension are the same to that for the global study as described in Section 3.6.2.

4.6.3. Statistical Methods for PRO Analysis

PRO Analysis are the same for extension to that for the global study as described in Section 3.6.3.

4.6.4. Summaries of Baseline Characteristics, Demographics, and Other Analyses

They are the same for extension to that for the global study as described in Section 3.6.4.

4.7. Interim Analysis and Final analysis

The primary analysis for PFS will be conducted in the Chinese subpopulation when approximately 75 PFS events have been collected. OS will also be analyzed.

The primary analysis for OS will be conducted in the Chinese subpopulation when approximately 75 OS events have been collected.

4.8. Multiplicity

No multiplicity adjustment will be applied to the analysis of China.

4.9. Sample Size and Power Calculations

After the completion of global study enrollment, the extension portion will continue to enroll participants and randomize eligible participants until the sample size for the overall randomized Chinese subpopulation reaches approximately 120. Participants from China enrolled in the extension portion of this study after completion of the global enrollment will not be included in the primary efficacy analysis population for the global study.

The extension portion will complete after approximately 75 OS events have been observed between the two arms in the Chinese subpopulation. With 75 OS events and a true hazard ratio of 0.7, the extension portion has ~80% chance to observe a point estimate of OS that preserves ≥ approximately 50% of the empirical risk reduction from the global analysis in the Chinese subpopulation. With 75 PFS events and a true hazard ratio of 0.7, the extension portion has ~80% chance to observe a point estimate of PFS that preserves ≥ approximately 50% of the empirical risk reduction from the global analysis in the Chinese subpopulation.

The above calculations for the consistency evaluation in OS and PFS are based on the same assumptions in the global study for sample size and power evaluation as specified in Section 3.9.

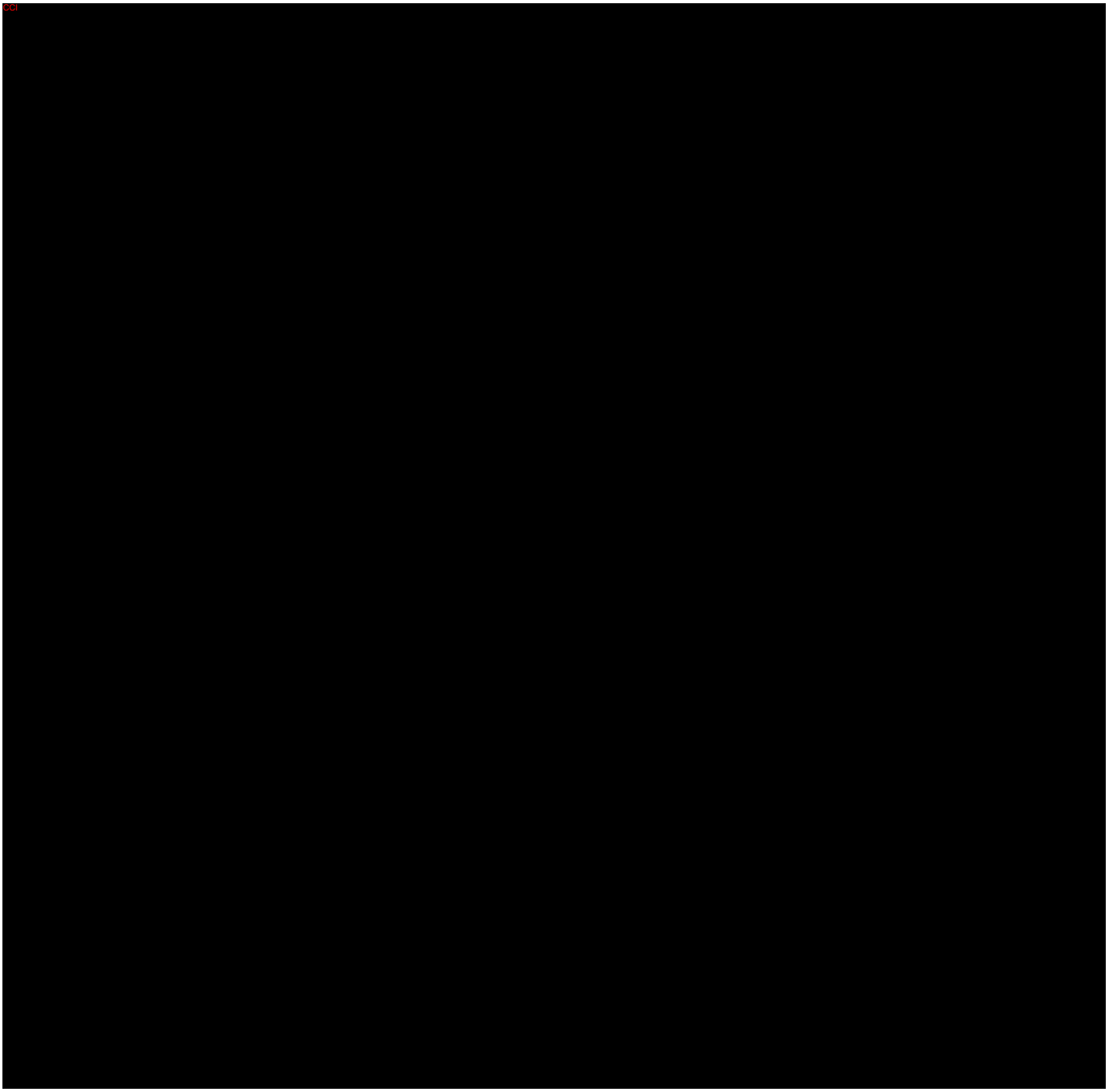
5. APPENDICES:

5.1. Rules for Non-Binding Futility Analyses

Non-binding futility analyses will be conducted to coincide with the safety DMC meeting in July 2021 (data cutoff in May 2021) and at the IA as per DMC recommendation. For each futility analysis, the DMC committee will evaluate the totality of the data and discuss with the EOC. The futility boundary for OS is based upon the Hwang-Shih-DeCani (HSD) beta-spending function with parameter of -15. The futility criteria are non-binding and only intended to provide guidance.

CCI

[redacted]



5.2. ABBREVIATIONS

Abbreviation	Expanded Term
1L	First-line
APaT	All Participants as Treated
BICR	blinded independent central review
CI	confidence interval

Abbreviation	Expanded Term
cLDA	constrained longitudinal data analysis
CR	complete response
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
DMC	data monitoring committee
DOR	duration of response
ECG	electrocardiogram
ECI	event of clinical interest
EOC	Executive Oversight Committee
EORTC	European Organisation for Research and Treatment of Cancer
ePRO	electronic patient-reported outcome
EQ-5D	EuroQoL-5D-5L
FAS	Full Analysis Set
HR	hazard ratio
HRQoL	health-related quality of life
HSD	Hwang-Shih-DeCani
IA	interim analysis
ICH	International Conference on Harmonization
iRECIST	Modified RECIST 1.1 for Immune-based Therapeutics
ITT	Intention to Treat
NPH	Non-proportional Hazards
PH	Proportional Hazards
NSCLC	non-small cell lung cancer
OR	objective response
ORR	objective response rate
OS	overall survival
PD	progressive disease
PD-L1	programmed cell death ligand 1
PFS	progression-free survival
PH	proportional hazards

Abbreviation	Expanded Term
PK	pharmacokinetic
PR	partial response
PRO	patient-reported outcomes
QLQ-LC13	Quality of Life Questionnaire and Lung Cancer Module 13
QoL	quality of life
RECIST 1.1	Response Evaluation Criteria in Solid Tumors 1.1
SAE	serious adverse event
SAP	statistical analysis plan
sSAP	supplemental statistical analysis plan
TPS	tumor proportion score
TTD	time to true deterioration

6. REFERENCES

Cocks K, King MT, Velikova G, de Castro G , Martyn St-James M, Fayers PM, et al. Evidence-based guidelines for interpreting change scores for the European organisation for the research and treatment of cancer quality of life questionnaire core 30. Eur J Cancer. 2012 Jul;48(11):1713-21.

Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2011;20(10):1727-1736.

King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. Quality of Life Research 1996;5:555-67.

Lan, K. K., & DeMetes, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika*, 70(3): 659-63.

Liang K, Zeger, S (2000). Longitudinal data analysis of continuous and discrete responses for pre-post designs. Sankhyā: The Indian Journal of Statistics, 62 (Series B), 134-148.

Maringwa JT, Quinten C, King M, et al. Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer.* 2011;19(11):1753-1760.

Maurer, W., & Bretz, F. (2013). Multiple Testing in Group Sequential Trials using Graphical Approaches. *Statistics in Biopharmaceutical Research*, 5(4): 311-320.

Miettinen, O., & Nurminen, M. (1985). Comparative analysis of two rates. *Statistics in Medicine*, 4:213-226.

O'Brien, P. C., & Fleming, T. R. (1979). A Multiple Testing Procedure for Clinical Trials. *Biometrics*, 35(3): 549-556.

Osoba, D., Rodrigues, G., Myles, J., Zee, B., & Pater, J. (1998). Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*, 16:139-44.

Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health and Quality of Life Outcomes.* 2007;5:70.

Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., Graeff , A. d., Groenvold, M., . . . Sprangers, M. A. (2008). *EORTC QLQ-C30 Reference Values.* Retrieved from https://www.eortc.org/app/uploads/sites/2/2018/02/reference_values_manual2008.pdf