Official Protocol Title:	A Phase 3 Randomized, Open-label Clinical Study to Evaluate the Pharmacokinetics and Safety of Subcutaneous Pembrolizumab Coformulated With Hyaluronidase (MK-3475A) Versus Intravenous Pembrolizumab, Administered With Chemotherapy, in the First-line Treatment of Participants With Metastatic Non-small Cell Lung Cancer
NCT number:	NCT06212752
Document Date:	23-Aug-2024

Supplemental Statistical Analysis Plan (sSAP) - Japan

TABLE OF CONTENTS

TABLE OF CONTENTS 2

LIST OF TABLES 4

1 INTRODUCTION 5

2 SUMMARY OF CHANGES 5

3 STATISTICAL ANALYSIS PLAN 7

3.1 STATISTICAL ANALYSIS PLAN SUMMARY 7
3.2 RESPONSIBILITY FOR ANALYSES/IN-HOUSE BLINDING 9
3.3 HYPOTHESES/ESTIMATION 9
3.4 ANALYSIS ENDPOINTS 11
3.4.1 Efficacy Endpoints 11
3.4.2 Pharmacokinetics Endpoints 12
3.4.3 Safety Endpoints 13
3.4.4 Immunogenicity Endpoint 13
3.5 ANALYSIS POPULATIONS 13
3.5.1 Efficacy Analysis Populations 13
3.5.2 Pharmacokinetics Analysis Populations 14
3.5.3 Safety Analysis Populations 16
3.6 STATISTICAL METHODS 16
3.6.1 Statistical Methods for Efficacy Analyses 16
3.6.1.1 Objective Response Rate 17
3.6.1.2 Progression-free Survival 18
3.6.1.3 Overall Survival 19
3.6.1.4 Duration of Response 19
3.6.1.5 Analysis Strategy for Key Efficacy Variables 20
3.6.2 Statistical Methods for Pharmacokinetics Analyses 21
3.6.3 Statistical Methods for Safety Analyses 22
3.6.3.1 Overall Safety Assessment 22
3.6.3.2 Assessment of Safety Topics of Special Interest 23
3.6.4 Demographic and Baseline Characteristics 24
3.7 INTERIM ANALYSIS 24
3.7.1 Efficacy Interim Analysis 24
3.7.2 Safety Interim Analysis 25
3.8 MULTIPLICITY 25
3.8.1 ORR 25
3.9 SAMPLE SIZE AND POWER CALCULATIONS 26
3.10 SUBGROUP ANALYSES 27
3.11 COMPLIANCE (MEDICATION ADHERENCE) 27
3.12 EXTENT OF EXPOSURE 28

4 APPENDICES 28

4.1 APPENDIX 1: SYNTHESIS METHOD FOR ORR NON-INFERIORITY 28
4.2 APPENDIX 2: APPROVAL INFORMATION 31

5 REFERENCES 32

LIST OF TABLES

Table 1 Censoring Rules for Primary and Sensitivity Analysis of PFS 19

Table 2 Censoring Rules for DOR 20

Table 3 Analysis Strategy for Key Efficacy Variables 21

Table 4 Analysis Strategy for Safety Endpoints 23

Table 5 Summary of Analysis Strategy 25

Table 6 Boundaries and Properties for ORR Analysis 26

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.

To meet local regulatory requirements, different primary objective/hypothesis testing from the protocol is specified for Japan and is detailed in this sSAP.

2 SUMMARY OF CHANGES

This sSAP aligns with protocol amendment 03. Changes from the previous Japan-specific sSAP and the rationale are summarized below.

Section Number and Name	Description of Change	Brief Rationale
Section 3.1, Statistical Analysis Plan Summary	Interim analyses 1 and 2 were removed. The timing for the interim safety analysis was updated.	The statistical analysis strategy has been updated. This change accounts for rapid enrollment in the study. The timing of final analysis is updated to be performed when a minimum of 27 weeks follow-up is achieved after last participant has been randomized to allow for sufficient follow up for efficacy data. The futility analysis is not needed given internal data from earlier studies.
	Sample size and power for PK endpoints were updated.	The prevalences of histology and PDL1 status assumed in the ORR noninferiority test are different in the current study compared to the historical data.
	Removed PRO related information	PRO will not be analyzed in Japan population.
	Added description of Extension Study in Japan.	To extend the enrollment period beyond the global study to achieve required number of participants in Japan to investigate efficacy safety, and PK.
Section 3.2, Responsibility for Analyses/In-house Blinding	Removed reference to interim analyses.	Refer to Section 3.1 rationale.
	Added description of Extension Study in Japan.	Refer to Section 3.1 rationale for Extension study in Japan.
Section 3.3 Hypotheses/Estimation	Removed PRO endpoints	Refer to Section 3.1 rationale for removing PRO.
	Removed the exploratory pharmacoeconomic endpoint.	The Time & Motion pharmacoeconomic study will be conducted in parallel, but is not an endpoint of this study.

Section Number and Name	Description of Change	Brief Rationale
Section 3.4.5, PRO Endpoints	Removed PRO endpoints	Refer to Section 3.1 rationale for removing PRO.
Section 3.5.1, Efficacy Analysis Populations	Specified that for the participant population, the subset of participants being described pertains to the global study and description of Extension Study in Japan.	Refer to Section 3.2 rationale for Extension study in Japan.
	Added the population for DoR analysis.	To specify the details for population used for DoR analysis.
Section 3.5.2, Pharmacokinetics Analysis Populations	Specified that for the participant population, the subset of participants being described pertains to the global study and description of Extension Study in Japan.	Refer to Section 3.1 rationale for Extension study in Japan.
Section 3.5.3, Safety Analysis Populations	Specified that for the participant population, the subset of participants being described pertains to the global study and description of Extension Study in Japan.	Refer to Section 3.2 rationale for Extension study in Japan.
Section 3.5.4, PRO Analysis Population	Removed	Refer to Section 3.1 rationale for removing PRO.
Section 3.6.1 Objective Response Rate	Added the stratum pooling strategy for ORR endpoint in ITT population and overall ITT population.	To pool small strata to ensure sufficient number of objective responses (>=5) in each stratum for ORR stratified analyses.
	Specified unstratified analysis will be performed in Japan ITT population	To clarify the unstratified analysis for efficacy endpoints will be used for Japan ITT population.
Section 3.6.4, Statistical Methods for Patient Reported Outcome Analyses	Removed	Refer to Section 3.1 rationale for removing PRO.
Section 3.7, Interim Analysis	Removed reference to interim analyses and updated the timing of PK endpoints and interim safety analysis.	Refer to Section 3.1 rationale for removing interim analyses
Section 3.7.1, Efficacy Interim Analysis	Removed reference to interim PK analysis.	Refer to Section 3.1 rationale for removing interim analyses
	Added description of efficacy analyses strategies in different populations.	To clarify that efficacy endpoints will be summarized in a descriptive way with longer follow up time among different populations.
Section 3.7.2, Safety Interim Analysis	Updated interim safety analysis timing.	Refer to Section 3.1 rationale for removing interim analyses
Section 3.8.1, ORR	Removed reference to interim analyses and updated boundaries and properties tables.	Refer to Section 3.1 rationale.
	Removed the footnote regarding *p* (1-sided) in the table.	The statement is not applicable.
Section 3.8.2, Cycle 3 C_{trough}	Removed the paragraph.	Refer to Section 3.1 rationale that no futility IA will be performed based on Cycle 3 C_{trough}.

Section Number and Name	Description of Change	Brief Rationale
Section 3.8.3, Safety Analyses	Removed the paragraph.	Prior to FA, DMC has not requested any ad-hoc PK analyses and thus no adjustment has bee implemented in the testing strategy.
Section 3.9, Sample Size and Power Calculations	Updated the number of evaluable participants, removed reference to interim analyses, updated assumptions, and added description of Extension Study in Japan.	Refer to Section 3.1 rationale.
Section 3.10, Subgroup Analyses	Added between-group treatment difference in ORR will be estimated in subgroup analyses.	Clarified ORR difference will be summarized.
Section 4, Appendix 1: Synthesis Method for ORR Noninferiority	The assumptions in ORR noninferiority test were updated.	Refer to Section 3.1 rationale for sample size and power updates.

3 STATISTICAL ANALYSIS PLAN

3.1 Statistical Analysis Plan Summary

Key elements of the statistical analysis plan are summarized below; the comprehensive plan is provided in Sections 3.2-3.12.

Study Design Overview	A Phase 3 Randomized, Open-label Clinical Study to Evaluate the Pharmacokinetics and Safety of Subcutaneous Pembrolizumab Coformulated With Hyaluronidase (MK-3475A) Versus Intravenous Pembrolizumab, Administered With Chemotherapy, in the First-line Treatment of Participants With Metastatic Non-small Cell Lung Cancer
Treatment Assignment	Approximately 339 participants will be randomized in a 2:1 ratio between 2 treatment groups: (1) MK-3475A SC Q6W in combination with platinum doublet chemotherapy and (2) pembrolizumab IV Q6W in combination with platinum doublet chemotherapy. Randomization stratification factors are: 1) ECOG (0 versus 1), 2) Histology (squamous versus nonsquamous), 3) PD-L1 TPS (<50% versus ≥50%; PD-L1 nonevaluable participants will be included with the TPS <50% group), and 4) Region (East Asia versus North America/Western Europe/Australia/New Zealand versus Rest of the World). This is an open-label study.
Analysis Populations	• Efficacy (primary): ITT • PK: Per-protocol Set • Safety: APaT
Primary Endpoint	• ORR

Secondary Endpoints	• Cycle 1 $AUC_{0\text{-}6\ wks}$
	• Steady-state (Cycle 3) C_{trough} (the primary analysis will be performed on the model-based values of C_{trough})
	• For descriptive comparison to pembrolizumab IV Q6W:
	  ○ Cycle 1: C_{max}, C_{trough}
	  ○ Cycle 3: $AUC_{0\text{-}6\ wks}$, C_{max}
	• For descriptive comparison to pembrolizumab IV Q3W:
	  ○ Model-based C_{trough} at Cycle 1 and steady state
	• ADA
	• PFS
	• OS
	• DOR
	• Safety and tolerability

Statistical Methods for Key Efficacy Analyses	For primary hypothesis of ORR, synthesis method (as described in Appendix 1) will be used to test non-inferiority. Stratified Miettinen and Nurminen method as well as stratified Cochran-Mantel-Haenszel method will also be used for comparison of the ORR between 2 treatment groups.
Statistical Methods for Key Immunogenicity/ Pharmacokinetic Analyses	For secondary endpoints of Cycle 1 $AUC_{0\text{-}6\ wks}$ and Cycle 3 C_{trough}, GMR will be evaluated between MK3475-A SC and pembrolizumab IV. Computation of the CIs of GMR will be calculated using Welch's t-test statistics (which does not rely on the assumption of equal variances for SC and IV) with the log-transformed AUC and C_{trough}.
Statistical Methods for Key Safety Analyses	For analyses in which 95% CIs will be provided for between-treatment differences in the percentage of participants with events, these analyses will be performed using the Miettinen and Nurminen's method [Miettinen, O. and Nurminen, M. 1985].
Interim Analyses	**<u>Efficacy and Pharmacokinetics</u>** There are no planned IAs for efficacy or pharmacokinetics analysis in this study. One final analysis (FA) is planned to be performed when a minimum of 27 weeks follow-up after the last participant is randomized in the global study. The purpose of FA is for the testing of non-inferiority of ORR. **<u>Safety</u>** • The study plans 1 interim safety analysis, which will be performed approximately 7 months after the first participant is randomized. Details will be specified in the DMC charter.
Multiplicity	The overall Type I error for the primary endpoint of ORR is strongly controlled at 0.025 (1-sided).

Sample Size and Power	The planned sample size is approximately 378 participants. For the primary endpoint of ORR, based on the overall sample size of 378 participants, the study can have approximately 89% power to reject the null hypothesis, ie, log(ORR ratio of Arm 1 versus Arm 2) ≤50% of -log(ORR ratio of pembrolizumab IV in combination with chemotherapy versus chemotherapy), under a true ORR of 51.1% for both arms at an overall α level of 0.025 (1 sided) using the synthesis method for noninferiority [U.S. Food and Drug Administration 2016]. Details are provided in Appendix 1.
Japan Extension Study	Japan participants randomized during the global study phase will be included in all global study analyses (efficacy, PK and safety). Japan participants randomized during the Japan extension phase will be excluded from all global study analyses. Japan participants randomized during global and extension phases will both be included in the Japan-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 randomized, open-label study, ie, participants, investigators, and Sponsor personnel will be aware of participant treatment assignments after each participant is enrolled and treatment is assigned.

The Sponsor will generate the randomized allocation schedule(s) for study intervention assignment, and the randomization will be implemented in an interactive voice response system by a study vendor.

Although the study is open-label, analyses or summaries generated by randomized intervention assignment, or actual intervention received will be limited and documented.

Extension Study In Japan

For all participants in Japan, including participants randomized in the global study and the extension study, analyses or summaries generated by randomized intervention assignment, or actual intervention received will be limited and be documented by the statistician(s)/programmer(s) responsible for the analysis of the Extension Study in Japan. The extent to which individuals are unblinded to the results will be limited, and blinded and unblinded members will be clearly documented.

3.3 Hypotheses/Estimation

Objectives and hypotheses of the study are stated below.

Objectives	Endpoints
Primary	
To evaluate MK-3475A SC and pembrolizumab IV with respect to ORR per RECIST 1.1 as assessed by BICR Hypothesis: MK-3475A SC in combination with chemotherapy retains at least 50% of the treatment effect of IV pembro combo over chemotherapy	Objective response: CR or PR
Secondary	
To compare MK-3475A SC to pembrolizumab IV with respect to Cycle 1 AUC	Cycle 1 $AUC_{0\text{-}6wks}$
To compare MK-3475A SC to pembrolizumab IV with respect to steady-state (Cycle 3) C_{trough}	Steady-state (Cycle 3) C_{trough} The primary analysis will be performed on the model-based values of C_{trough}
To evaluate pembrolizumab exposure for MK-3475A SC relative to pembrolizumab IV Q6W	Cycle 1: C_{max} and C_{trough} Steady state (Cycle 3): $AUC_{0\text{-}6wks}$ and C_{max}
To evaluate the development of circulating anti-pembrolizumab antibodies for MK-3475A SC and pembrolizumab IV	Anti-pembrolizumab antibodies
To evaluate pembrolizumab C_{trough} for MK-3475A SC relative to pembrolizumab IV Q3W	Cycle 1: Model-based C_{trough} Steady state: Model-based C_{trough}
To evaluate MK-3475A SC and pembrolizumab IV with respect PFS per RECIST 1.1 as assessed by BICR	PFS: The time from randomization to the first documented disease progression or death due to any cause, whichever occurs first
To evaluate MK-3475A SC and pembrolizumab IV with respect to OS	OS: The time from randomization to death due to any cause

Objectives	Endpoints
To evaluate MK-3475A SC and pembrolizumab IV with respect DOR per RECIST 1.1 as assessed by BICR	DOR: The time from the first documented evidence of CR or PR until disease progression or death due to any cause, whichever occurs first
To evaluate the safety and tolerability of MK-3475A SC and pembrolizumab IV	AE Discontinuation of study intervention due to AEs
CCI [redacted]	

3.4 Analysis Endpoints

Efficacy, pharmacokinetics, and safetyendpoints that will be evaluated for within- and/or between-treatment differences are listed below.

3.4.1 Efficacy Endpoints

Primary

- **Objective Response Rate (ORR)**

The ORR is defined as the percentage of participants who achieve a confirmed CR or PR per RECIST 1.1 as assessed by BICR.

Secondary

- **Progression-free survival (PFS)**

PFS is defined as the time from randomization to the first documented disease progression per RECIST 1.1 by BICR or death due to any cause, whichever occurs first.

- **Overall Survival (OS)**

OS is defined as the time from randomization to death due to any cause.

- **Duration of Response (DOR)**

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

3.4.2 Pharmacokinetics Endpoints

The methodological details for assessment of model-based PK exposures ($AUC_{0\text{-}6\ wks}$ and C_{trough}) will be summarized in a separate MAP.

Secondary

- **Cycle 1 $AUC_{0\text{-}6\ wks}$**

Cycle 1 $AUC_{0\text{-}6\ wks}$ is defined as the model-based area under curve exposure over a 6-week dosing interval in Cycle 1.

- **Cycle 3 C_{trough}**

Cycle 3 C_{trough} is defined as the trough concentration at the end of the dosing interval in Cycle 3, representing steady-state.

Two assessments of this endpoint will be made: model-based C_{trough}, which is the value predicted by the PK model, and observed C_{trough}, which is the measured value. The primary analysis for Cycle 3 C_{trough} will be based on the model-based value. A sensitivity analysis of this endpoint will be performed on the observed value.

For descriptive comparison with pembrolizumab IV Q6W:

- **Cycle 1 C_{trough}**

Cycle 1 C_{trough} is defined as the trough concentration at the end of the dosing interval in Cycle 1. Two assessments of this endpoint will be made: observed C_{trough}, which is the measured value, and model-based C_{trough}, which is the value predicted by the PK model.

- **Cycle 3 $AUC_{0\text{-}6\ wks}$**

Cycle 3 $AUC_{0\text{-}6\ wks}$ is defined as the model-based area under curve exposure over a 6-week dosing interval in Cycle 3, representing steady-state.

- **Cycle 1 C_{max}**

Cycle 1 C_{max} is defined as the peak concentration over the dosing interval in Cycle 1. Two assessments of this endpoint will be made: observed C_{max}, which is the measured value, and model-based C_{max}, which is the value predicted by the PK model.

- **Cycle 3 C_{max}**

Cycle 3 C_{max} is defined as the peak concentration over the dosing interval in Cycle 3, representing steady-state. Two assessments of this endpoint will be made: observed C_{max}, which is the measured value, and model-based C_{max}, which is the value predicted by the PK model.

For descriptive comparison with pembrolizumab IV Q3W:

- **Model-based Cycle 1 C_{trough}**

Model-based Cycle 1 C_{trough} is defined as the value of trough concentration at the end of the dosing interval in Cycle 1, as predicted by the PK model.

- **Model-based steady-state C_{trough}**

Model-based steady-state C_{trough} is defined as the value of trough concentration at the end of the dosing interval at steady-state, as predicted by the PK model. This corresponds to the model-predicted C_{trough} value at Cycle 3 for MK-3475A Q6W and at Cycle 6 for pembrolizumab 200 mg IV Q3W.

3.4.3 Safety Endpoints

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

3.4.4 Immunogenicity Endpoint

Immunogenicity (ADA incidence) will be assessed by analyzing the development of ADAs following administration of MK-3475A SC and pembrolizumab IV.

3.5 Analysis Populations

3.5.1 Efficacy Analysis Populations

The ITT population will serve as the population for efficacy analysis. All randomized participants in the global study will be included in this population. Participants will be included in the treatment group to which they are randomized. The analysis population for DOR consists of participants in the analysis population of OR who demonstrate confirmed CR or PR.

Extension Study in Japan

After enrollment of the global study is completed, the study will continue to randomize participants in Japan until the sample size for randomized participants in Japan reaches approximately 39. The efficacy endpoints will also be analyzed in the overall ITT population, defined as all randomized participants in the global study and Japan extension study, and the Japan ITT population, defined as all randomized participants in Japan in the global study and the extension study.

3.5.2 Pharmacokinetics Analysis Populations

The PP population in the global study will be the primary population used for the analysis of PK data in this study.

- **Cycle 1 $AUC_{0\text{-}6\ wks}$**

The PP population for the secondary PK endpoint of Cycle 1 $AUC_{0\text{-}6\ wks}$ consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose PK sample in Cycle 1 and for whom a model-based assessment of $AUC_{0\text{-}6\ wks}$ can be made.

- **Cycle 3 C_{trough}**

The primary analysis for the secondary PK endpoint of Cycle 3 C_{trough} will be performed on the model-based Cycle 3 C_{trough}. An additional sensitivity analysis of this endpoint will be performed using observed Cycle 3 C_{trough}.

- The PP population for model-based Cycle 3 C_{trough} consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose sample and for whom a model-based assessment of Cycle 3 C_{trough} can be made.

- The PP population for the observed Cycle 3 C_{trough} consists of the subset of participants in the global study who received all 3 doses from Cycle 1 to 3 within the permissible dosing window as per the SoA and have a valid PK sample on Cycle 3 Day 42 with no documented assay or bioanalytical error and within permissible window as per the PK SoA (ie, within Days 41 to 43 of the dosing day in Cycle 3).

Any participants or data values excluded from the primary analyses for the secondary PK endpoints will be identified, along with the reasons for exclusion, in the CSR.

For descriptive comparison with pembrolizumab IV Q6W

- **Cycle 1 C_{trough}**

The PP population for observed Cycle 1 C_{trough} consists of the subset of participants in the global study who received the Cycle 1 dose and have a valid PK sample on Cycle 1 Day 42 with no documented assay or bioanalytical error and within permissible window as per the PK SoA (ie, within Days 41 to 43 of the dosing day in Cycle 1).

The PP population for model-based Cycle 1 C_{trough} consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose sample in Cycle 1 and for whom a model-based assessment of C_{trough} can be made.

- **Cycle 3 $AUC_{0\text{-}6\ wks}$**

The PP population for Cycle 3 $AUC_{0\text{-}6\ wks}$ consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose sample and for whom a model-based assessment of Cycle 3 $AUC_{0\text{-}6\ wks}$ can be made.

- **Cycle 1 C_{max} and Cycle 3 C_{max}**

The PP population for the observed Cycle 1 C_{max} consists of the subset of participants in the global study who received the Cycle 1 dose within the permissible dosing window as per the SoA, and have a valid PK sample on Cycle 1 Day 1 end-of-infusion for the IV arm or at least 1 valid PK sample on Cycle 1 Days 5 to 10 for the SC arm with no documented assay or bioanalytical error and within permissible window as per the PK SoA.

The PP population for the observed Cycle 3 C_{max} consists of the subset of participants in the global study who received all 3 doses from Cycle 1 to 3 within the permissible dosing window as per the SoA, and have a valid PK sample on Cycle 3 Day 1 end-of-infusion for the IV arm or at least 1 valid PK sample on Cycle 3 Days 5 to 10 for the SC arm with no documented assay or bioanalytical error and within permissible window as per the PK SoA.

The PP population for model-based Cycle 1 C_{max} and model-based Cycle 3 C_{max} consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose sample in Cycle 1 and for whom a model-based assessment of Cycle 1 C_{max} and Cycle 3 C_{max}, respectively, can be made.

For descriptive comparison with pembrolizumab IV Q3W

- **Model-based Cycle 1 C_{trough} and Steady-state C_{trough}**

The PP population for model-based Cycle 1 C_{trough} and model-based steady-state C_{trough} consists of the subset of participants in the global study who received the Cycle 1 dose, with at least 1 valid postdose sample in Cycle 1 and for whom a model-based assessment of Cycle 1 and steady-state C_{trough}, respectively, can be made.

Extension Study in Japan

The participants in Japan randomized and treated in the extension study after completion of the global enrollment will not be included in the PP population for the global study. The Japan PP population for PK endpoints will be supportive. The Japan PP population for Cycle 1 $AUC_{0\text{-}6\ wks}$ consists of the subset of participants in Japan in the global study and extension study who received the Cycle 1 dose, with at least 1 valid postdose PK sample in Cycle 1 and for whom a model-based assessment of $AUC_{0\text{-}6\ wks}$ can be made. The Japan PP population for Cycle 3 C_{trough} consists of participants in Japan in the global study and extension study who received the Cycle 1 dose, with at least 1 valid postdose sample and for whom a model-based assessment of Cycle 3 C_{trough} can be made.

3.5.3 Safety Analysis Populations

Safety Analyses will be conducted in the APaT population, which consists of all randomized participants in the global study who received at least 1 dose of study treatment. Participants will be included in the treatment group corresponding to the study treatment 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 treatment for the entire treatment period; such participants will be included in the treatment group corresponding to the study treatment actually received.

At least 1 laboratory, vital sign, or ECG measurement obtained subsequent to at least 1 dose of study treatment 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 Japan

The participants in Japan randomized and treated in the Japan extension study after completion of the global enrollment will not be included in the safety analysis population for the global study. The Japan APaT population, including all participants in Japan randomized in the global study and the extension study who received at least 1 dose of study treatment, will be analyzed separately.

3.6 Statistical Methods

Statistical testing and inference for safety analyses are described in Section 3.6.2. 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 may be computed for other efficacy analyses and pharmacokinetics analyses, but should be interpreted with caution due to potential issues of multiplicity, sample size, etc.

3.6.1 Statistical Methods for Efficacy Analyses

This section describes the statistical methods that address the efficacy primary and secondary objectives.

The stratification factors used for randomization (see Section 6.3.2 of the protocol) will be applied to all stratified analyses, in particular, stratified Miettinen and Nurminen method [Miettinen, O. and Nurminen, M. 1985], stratified Cochran-Mantel-Haenszel method, and stratified Cox model. Participants were randomized into 24 strata defined by stratification factors: ECOG (0 vs. 1), Histology (squamous vs. nonsquamous), PD-L1 status (TPS<50% vs. TPS≥50%), and geographic region (East Asia vs. North America/Western Europe/Australia/New Zealand vs. Rest of the World). Based on a blinded review of objective response counts per RECIST 1.1 by BICR by stratum prior to final analysis in the global study, since there are <5 response counts in one or more strata, the stratification factors are to be combined for analyses to ensure sufficient number of responses in each strata based on the order of clinical importance of stratificantion factors (PD-L1 status > Histology > ECOG status > geographic region). Specifically, the following 13 strata will be

used for stratified ORR analyses in the global ITT population and overall ITT population (including Japan extension population):

- ECOG 0, squamous, TPS<50%, Rest of World/North America/Western Europe/Australia/New Zealand
- ECOG 0, Squamous, TPS≥50%, East Asia/Rest of World; or ECOG 1, Squamous, TPS≥50%, East Asia/North America/Western Europe/Australia/New Zealand
- ECOG 0, squamous, TPS<50%, East Asia
- ECOG 1, nonsquamous, TPS≥50%, East Asia
- ECOG 1, squamous, TPS≥50%, Rest of World
- ECOG 0, nonsquamous, TPS<50%, East Asia/North America/Western Europe/Australia/New Zealand
- ECOG 1, nonsquamous, TPS<50%, East Asia/North America/Western Europe/Australia/New Zealand
- ECOG 0, nonsquamous, TPS≥50%, East Asia/Rest of World
- ECOG 1, squamous, TPS<50%, East Asia
- ECOG 1, squamous, TPS<50%, Rest of World/North America/Western Europe/Australia/New Zealand
- ECOG 1, nonsquamous, TPS≥50%, Rest of World
- ECOG 0, nonsquamous, TPS<50%, Rest of World
- ECOG 1, nonsquamous, TPS<50%, Rest of World

The efficacy analyses in Japan ITT population will be conducted using the unstratified analysis.

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

3.6.1.1 Objective Response Rate

The primary objective is to determine that MK-3475A SC in combination with chemotherapy (thereafter referred as SC MK-3475A combo) retains at least 50% of the treatment effect of IV pembro over chemotherapy, as per the FDA NI guidance
[U.S. Food and Drug Administration 2016].

This objective will be assessed via the following non-inferiority hypothesis:

H_0: {log(ORR ratio of SC MK-3475A combo versus IV pembro combo)}≤(-1/2){log(ORR ratio of IV pembro combo versus chemotherapy)}

Synthesis method (as described in Appendix 1) will be used to test the above hypothesis. The ORR ratio of SC MK-3475A combo versus IV pembro combo will be computed using the stratified Cochran-Mantel-Haenszel method.

The stratified Miettinen and Nurminen's method will be used for comparison of the ORR between 2 treatment groups. The difference in ORR and its 95% CI from the stratified Miettinen and Nurminen's method with strata weighting by sample size will be reported. Furthermore, the stratified Cochran-Mantel-Haenszel method for ORR ratio will also be used for comparison of the ORR between 2 treatment groups. The strata pooling strategy described above in Section 3.6.1 will be applied to the stratified analysis specified above. The ratio of ORR and its 95% CI derived using logarithmic scale based on the normal approximation will be reported as applicable.

3.6.1.2 Progression-free Survival

The nonparametric Kaplan-Meier method will be used to estimate the PFS curve in each treatment group. 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, 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 (see Section 6.3.2 of the protocol) will be applied to the stratified Cox model.

Since disease progression 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. The true date of disease progression will be approximated by the earlier of the date of the first assessment at which PD is objectively documented per RECIST 1.1 by BICR and the date of death.

For the primary analysis, any participant who experiences an event (PD or death) immediately after 2 or more missed disease assessments will be censored at the last disease assessment prior to the missed visits. In addition, any participant who initiates new anticancer therapy will be censored at the last disease assessment prior to the initiation of new anticancer therapy. Participants who do not start new anticancer therapy and who do not experience an event will be censored at the last disease assessment. If a participant meets multiple criteria for censoring, the censoring criterion that occurs earliest will be applied. Sensitivity analyses will be performed for comparison of PFS based on investigator's assessment.

In order to evaluate the robustness of the PFS endpoint per RECIST 1.1 by BICR, an additional sensitivity analysis with different sets of censoring rules will be performed. The sensitivity analysis follows the intention-to-treat principle. That is, PDs/deaths are counted as events regardless of missed study visits or initiation of new anticancer therapy. If a participant meets multiple criteria for censoring, the censoring criterion that occurs earliest

will be applied. The censoring rules for the primary and sensitivity analysis are summarized in Table 1.

Table 1 Censoring Rules for Primary and Sensitivity Analysis of PFS

Situation	Primary Analysis	Sensitivity Analysis
PD or death documented after ≤1 missed disease assessment, and before new anticancer therapy, if any	Progressed at date of documented PD or death	Progressed at date of documented PD or death
Death or progression immediately after ≥2 consecutive missed disease assessments, or after new anticancer therapy	Censored at last disease assessment prior to the earlier date of ≥2 consecutive missed disease assessment and new anticancer therapy, if any	Progressed at date of documented PD or death
No PD and no death; and new anticancer treatment is not initiated	Censored at last disease assessment	Censored at last disease assessment
No PD and no death; new anticancer treatment is initiated	Censored at last disease assessment before new anticancer treatment	Censored at last disease assessment

PD = progressive disease; PFS = progression-free survival

3.6.1.3 Overall Survival

The nonparametric Kaplan-Meier method will be used to estimate the survival curves. 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 (see Section 6.3.2 of the protocol) will be applied to the stratified Cox model. Participants without documented death at the time of analysis will be censored at the date the participant was last known to be alive.

3.6.1.4 Duration of Response

If sample size permits, DOR will be summarized descriptively using Kaplan-Meier medians and quartiles. Only the subset of participants who show a confirmed complete response or partial response will be included in this analysis. Censoring rules for DOR are summarized in Table 2.

For each DOR analysis, a corresponding summary of the reasons responding participants are censored will also be provided. Responding participants who are alive, have not progressed, have not initiated new anticancer treatment, have not been determined to be lost to follow-up, and have had a disease assessment within ~ 5 months of the data cutoff date are considered ongoing responders at the time of analysis. If a participant meets multiple criteria for censoring, the censoring criterion that occurs earliest will be applied.

Table 2 Censoring Rules for DOR

Situation	Date of Progression or Censoring	Outcome
No progression nor death, no new anticancer therapy initiated	Last adequate disease assessment	Censor (nonevent)
No progression nor death, new anticancer therapy initiated	Last adequate disease assessment before new anticancer therapy initiated	Censor (nonevent)
Death or progression immediately after ≥2 consecutive missed disease assessments or after new anticancer therapy, if any	Earlier date of last adequate disease assessment prior to ≥2 missed adequate disease assessments and new anticancer therapy, if any	Censor (nonevent)
Death or progression after ≤1 missed disease assessments and before new anticancer therapy, if any	PD or death	End of response (event)
DOR = duration of response; PD = progressive disease A missed disease assessment includes any assessment that is not obtained or is considered inadequate for evaluation of response.		

3.6.1.5 Analysis Strategy for Key Efficacy Variables

A summary of the primary analysis strategy for the key efficacy endpoints is provided in Table 3.

Table 3 Analysis Strategy for Key Efficacy Variables

Endpoint/ Variable	Statistical Method	Analysis Population	Missing Data Approach
ORR per RECIST 1.1 by BICR	Testing: synthesis method for non-inferiority Estimation: stratified Miettinen and Nurminen method, stratified CMH method	ITT	Participants with missing data are considered nonresponders
PFS per RECIST 1.1 by BICR	Estimation: stratified Cox model with Efron's tie handling method	ITT	Censored according to rules in Table 1
OS	Estimation: stratified Cox model with Efron's tie handling method	ITT	Censored at participant's last known alive date
BICR = blinded independent central review; CMH = Cochran-Mantel-Haenszel; ITT = intent-to-treat; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; RECIST 1.1 = Response Evaluation Criteria in Solid Tumors.			

3.6.2 Statistical Methods for Pharmacokinetics Analyses

This section describes the statistical methods that address the pharmacokinetics objectives (which are secondary study objectives for Japan). For Japan, pharmacokinetics analyses are descriptive only. No type I error control is applied to pharmacokinetics analyses, so p-values, if provided, are nominal only and provided for descriptive purposes.

The secondary objectives are to compare GMR between MK-3475A SC and pembrolizumab IV based on Cycle 1 $AUC_{0\text{-}6\ wks}$ and Cycle 3 C_{trough}. For Cycle 3 C_{trough}, the primary analysis will be based on Cycle 3 model-based C_{trough}. A sensitivity analysis will be performed for Cycle 3 observed C_{trough}. The CIs for GMR will be calculated using Welch's t-test statistics (which does not rely on the assumption of equal variances for SC and IV) with the log transformed AUC and C_{trough}.

For each PK exposure (C_{trough}, $AUC_{0\text{-}6\ wks}$ and C_{max}, for Cycles 1 and 3) by treatment, the following descriptive statistics will be provided: N (number of participants with nonmissing data), median, minimum, maximum, GM, and geometric percent CV (calculated as 100 x sqrt(exp(s^2) – 1), where s^2 is the observed variance on the natural log-scale).

Based on PK data obtained in this study as well as historical PK data, an integrated population PK analysis will be performed to characterize the PK profile of pembrolizumab following SC and IV administrations and provide individual model-based PK exposure measures. Details are provided in the MAP.

3.6.3 Statistical Methods for Safety Analyses

The primary safety analyses will include only events that occur prior to Second Course treatment.

3.6.3.1 Overall Safety Assessment

The overall safety evaluation will include a summary by treatment group of the number and percentage of participants with at least 1 AE, drug-related AE, serious AE, serious drug-related AE, Grade 3 to 5 AE, a discontinuation from study treatment due to an AE, an AE that led to treatment interruption, and an AE resulting in death. Only point estimates by treatment group are provided. The number and percentage for injection-site reactions will be provided for the MK-3475A SC arm.

Point estimate and 95% CIs for the difference between treatment groups in the percentage of participants with specific AEs will be provided if at least 10% of participants in any treatment group exhibit the event. The threshold of at least 10% of participants was chosen because the population enrolled in this study is in critical condition and usually experiences 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, difference in the percentage of participants with specific Grade 3 to 5 AEs (≥5% of participants in 1 of the treatment groups) and SAEs (≥2% of participants in 1 of the treatment groups) will also be summarized by point estimate and 95% CIs.

CIs for between treatment group differences will be provided using the Miettinen and Nurminen's method [Miettinen, O. and Nurminen, M. 1985]. 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 as a formal method for assessing the statistical significance of the between-group differences.

Table 4 summarizes the analysis strategy for safety endpoints in this study.

Table 4 Analysis Strategy for Safety Endpoints

Analysis Part	Safety Endpoint	Descriptive Statistics	95% between-group CI
Overall Safety Assessment	Specific AEs (incidence ≥10% of participants in 1 of the treatment groups)	X	X
	Specific Grade 3-5 AE (incidence ≥5% of participants in 1 of the treatment groups)	X	X
	Specific serious AE (incidence ≥2% of participants in 1 of the treatment groups)	X	X
	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 from Study Treatment due to AE	X	
	AE that Resulted in Death	X	
	AE that Led to Treatment Interruption	X	
	Injection-site Reaction	X	
	Specific AEs, SOCs (incidence >0% of participants in any treatment group)	X	
	Change from Baseline Results (lab toxicity shift)	X	
Assessment of safety topics of special interest	Pembrolizumab AEOSI	X	

AE = adverse event; AEOSI = adverse event of special interest; CI = confidence interval; SOC = system organ class

3.6.3.2 Assessment of Safety Topics of Special Interest

AEs that are immune-mediated or potentially immune-mediated will be evaluated separately. These events have been characterized consistently throughout the pembrolizumab clinical development program. Point estimates and 95% CIs for between-group difference is not expected to add value to the safety evaluation, and hence only number and percentage of participants with such pembrolizumab AEOSI will be provided, as well as the number and percentage of participants with corticosteroids administration to treat an AEOSI. Summary statistics will be provided for the analysis of time from first dose to the onset of an AEOSI.

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 reasons 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 Analysis

There are no planned IAs for efficacy or pharmacokinetics endpoints for this study. The study plans 1 interim safety analysis.

An eDMC will serve as the primary reviewer of the results of the IAs of the study and will make recommendations for discontinuation of the study or protocol modifications to an executive committee of the Sponsor. If the eDMC recommends modifications to the design of the protocol or discontinuation of the study, this executive committee (and potentially other limited Sponsor personnel) may be unblinded to results at the treatment level in order to act on these recommendations. The extent to which individuals are unblinded with respect to results of IAs will be documented by the unblinded statistician. Additional logistical details will be provided in the eDMC Charter.

Treatment-level results from the IA will be provided to the eDMC by the unblinded statistician. Prior to final study unblinding, 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 IAs.

Although the study is open label, analyses or summaries generated by randomized treatment assignment, or actual treatment received, will be limited and documented. In addition, the independent radiologist(s) will perform the central imaging review without knowledge of treatment group assignment.

3.7.1 Efficacy Interim Analysis

There is no planned IAs for efficacy/pharmacokinetics analysis in this study. The analyses planned, endpoints evaluated, and drivers of timing are summarized in Table 5.

Table 5 Summary of Analysis Strategy

Key Endpoints	Timing	Estimated Time after First Participant Randomized	Primary Purpose of Analysis
ORR	~A minimum of 27 weeks follow-up after last participant randomized in the global study	~ 16.2 months	• noninferiority of ORR
By the timing of analysis, ~378 participants in the global study are expected to be randomized and followed up for at least 27 weeks.			

The noninferiority test of ORR will be conducted when a minimum of ~27 weeks after last participant randomized in the global study as indicated in Table 5. After that, the global population and Japan extension population will be followed up until a minimum of ~27 weeks after the last participant randomized in the Japan extension study. By then, the efficacy analyses including ORR, DOR, PFS and OS will be conducted and summarized descriptively in the global ITT population, Japan ITT population, and overall ITT population.

3.7.2 Safety Interim Analysis

The eDMC will be responsible for periodic interim safety reviews as specified in the DMC charter. The study plans 1 interim safety analysis, which will be performed approximately 7 months after the first participant is randomized. Details will be specified in the DMC charter.

3.8 Multiplicity

The Type I error rate for testing of the primary endpoint of ORR will be strongly controlled at an overall α level of 0.025 (1-sided).

3.8.1 ORR

The study will test ORR only once in ITT population in which the participants randomized in the extension study after the completion of the global enrollment will not be included. Table 6 shows the boundary properties for ORR analysis.

Table 6 Boundaries and Properties for ORR Analysis

Non-inferiority Analysis Based on ORR	**Value**	**α=0.025**
n*: 378 Month: 20.0	*Z*	1.9600
	p (1-sided)	0.025
	ORR ratio at bound[a]	0.8923
	P(Cross) if ORR ratio=1[b]	0.8925

The number of participants with evaluable data and timings are estimated approximately.

* n is the number of participants with evaluable ORR data (i.e., at least 27 weeks of follow-up) at analysis timing.

[a] ORR ratio at bound is the approximate ORR ratio of MK-3475A SC combo vs permbro IV combo required to reach an efficacy bound, assuming observed pembro IV combo ORR is similar to historical ORR.

[b] P(Cross) if ORR ratio=1 is the probability of crossing an efficacy bound under the alternative hypothesis.

Additional assumptions used for the calculation are specified in Section 3.9.

3.9 Sample Size and Power Calculations

The study will randomize approximately 378 participants in a 2:1 ratio into the MK-3475A SC and pembrolizumab IV arms. ORR is the primary endpoint for the study in Japan, with Cycle 1 $AUC_{0\text{-}6\ wks}$ and Cycle 3 C_{trough} are secondary endpoints.

The projected enrollment period is approximately 10 months. ORR is the primary endpoint and based on the overall sample size of 378 participants, the study can have approximately 89% power to reject the null hypothesis, ie, log(ORR ratio of Arm 1 versus Arm 2) ≤50% of -log(ORR ratio of pembrolizumab IV in combination with chemotherapy versus chemotherapy), under a true ORR of 51.1% for both arms at an overall α level of 0.025 (1 sided) using the synthesis method for noninferiority
[U.S. Food and Drug Administration 2016]. Details are provided in Appendix 1.

Based on the historical data from KEYNOTE-189 and KEYNOTE-407, the power calculation for ORR assumes the following:

- In the current study, the projected prevalence of PDL1 TPS ≥50% is 19% and PDL1 TPS<50% is 81%.

CCI

- ORR ratio=1 (MK-3475A SC vs pembrolizumab IV) under the alternative hypothesis.

- The log (ORR ratio) and the standard error of log (ORR ratio) from the historical studies (in this case, KEYNOTE-189 and KEYNOTE-407) are assumed to CCI [redacted] respectively.

The sample size and power calculations were performed using R.

Extension Study in Japan

After the enrollment for the global study has completed, the study will continue to randomize participants in a 2:1 ratio into the MK-3475A SC and pembrolizumab IV arms in Japan until the sample size for the Japanese participants reaches approximately 39.

3.10 Subgroup Analyses

To determine whether the treatment effect is consistent across various subgroups, the between-group difference in ORR (with a nominal 95% CI) will be estimated and plotted within each category of the following classification variables:

- Age category (≤65, >65 years)
- Sex (female, male)
- Race (white, non-white)
- Smoking status (never, former/ current)
- ECOG (0, 1)
- Histology (squamous, nonsquamous)
- Geographic region (East Asia, North America/Western Europe/Australia/New Zealand, Rest of the World)
- PD-L1 expression (unknown, TPS <1%, or TPS ≥1%)
- PD-L1 expression (unknown, TPS <50%, or TPS ≥50%)
- PD-L1 expression (unknown, TPS <1%, 1%≤TPS≤49%, or TPS ≥50%)

For subgroups, the derived strata based on eCRF collected information will be used. For ORR, the unstratified Miettinen and Nurminen method will be used. The consistency of the treatment effect will be assessed descriptively via summary statistics by category for the classification variables listed above. If any level of a subgroup variable has fewer than 10% of the ITT population, above analysis will not be performed for this level of the subgroup variable. If a subgroup variable has two levels and one level of the subgroup variable has fewer than 10% of the ITT population, then this subgroup will not be displayed in the forest plot.

3.11 Compliance (Medication Adherence)

Drug accountability data for study treatment 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 the number of cycles in which the participant receives the study intervention. Summary statistics will be provided on the extent of exposure for the APaT population.

4 APPENDICES

4.1 Appendix 1: Synthesis Method for ORR Non-inferiority

Following the FDA guidance on Noninferiority Clinical Trials to Establish Effectiveness [U.S. Food and Drug Administration 2016], the synthesis method is proposed to be used to evaluate noninferiority of ORR in this study. Synthesis method combines or synthesizes the data from the historical studies and the current study. KEYNOTE-189 and KEYNOTE-407 are proposed as the historical studies included in the synthesis method, because:

- The current study plans to enroll participants for the first-line treatment of metastatic nonsquamous or squamous NSCLC, a patient population consistent with KEYNOTE-189 (nonsquamous) and KEYNOTE-407 (squamous)
- The control arm treatment in the current study with IV pembrolizumab in combination with chemotherapy (thereafter referred as IV pembro combo) is the same as the experimental arm studied in KEYNOTE-189 and KEYNOTE-407
- KEYNOTE-189 and KEYNOTE-407 studies are randomized, double-blind studies, which provide the most reliable estimate for the treatment effect of IV pembro combo over the prior standard of care of chemotherapy
- Based on the common global footprint in site selection and similar patient characteristics between the current study and KEYNOTE-189/KEYNOTE-407, the assumption about a similar effect of the IV pembro combo between the current study and the historical studies (ie, constancy assumption) is expected to be valid.

It is proposed that SC MK-3475A in combination with chemotherapy (thereafter referred as SC MK-3475A combo) needs to retain at least 50% of the treatment effect of IV pembro combo over chemotherapy. So, the null hypothesis for the synthesis method is specified as:

H_0: {log(ORR ratio of SC MK-3475A combo versus IV pembro combo)}≤(-1/2){log(ORR ratio of IV pembro combo versus chemotherapy)}

The log ORR ratio of IV pembro combo versus chemotherapy and its standard error need to be calculated to carry out the synthesis method. The current study plans to enroll both the nonsquamous and squamous populations, with a population prevalence of *c* and *1-c,* respectively. So, the ORR in the study population can be considered as a weighted average from the nonsquamous and squamous populations. To compute the log ORR ratio of IV pembro combo versus chemotherapy and its standard error, the KEYNOTE-189 and KEYNOTE-407 data is weighted by the population prevalence of *c* and *1-c*, respectively.

Let X_{11} and X_{12} be the observed number of responders for PDL1 TPS≥50% and TPS<50% subgroups in Pembro Combo arm, Y_{11} and Y_{12} be the observed number of responders for PDL1 TPS≥50% and TPS<50% subgroups in Control arm in KEYNOTE-189. Similarly, let X_{21} and X_{22} be the observed number of responders for PDL1 TPS≥50% and TPS<50% in Pembro Combo arm, Y_{21} and Y_{22} be the observed number of responders for PDL1 TPS≥50% and TPS<50% in Control arm in KEYNOTE-407. Thus,

$$X_{11} \sim Bin(n_{11}, P_{X11})$$
$$X_{12} \sim Bin(n_{12}, P_{X12})$$
$$Y_{11} \sim Bin(m_{11}, P_{Y11})$$
$$Y_{12} \sim Bin(m_{12}, P_{Y12})$$
$$X_{21} \sim Bin(n_{21}, P_{X21})$$
$$X_{22} \sim Bin(n_{22}, P_{X22})$$
$$Y_{21} \sim Bin(m_{21}, P_{Y21})$$
$$Y_{22} \sim Bin(m_{22}, P_{Y22}),$$

where, n_{11} and n_{12} are the sample sizes, and P_{X11} and P_{X12} are the underlying ORRs for PDL1 TPS≥50% and TPS<50% in Pembro Combo arm, respectively, in KEYNOTE-189; m_{11} and m_{12} are the sample sizes, and P_{Y11}, and P_{Y12} are the underlying ORRs for PDL1 TPS≥50% and TPS<50% in Control arm, respectively, in KEYNOTE-189; similarly, n_{21}, n_{22}, m_{21} and m_{22} are the sample sizes, and P_{X21}, P_{X22}, P_{Y21}, and P_{Y22} are the underlying ORRs for PDL1 TPS≥50% and TPS<50% in Pembro Combo and PDL1 TPS≥50% and TPS<50% in Control arm, respectively, in KEYNOTE-407. Let $\hat{P}_{11}$ and $\hat{P}_{10}$ be the estimates of weighted ORRs by PDL1 status for Pembro Combo and Control arms, respectively, from KEYNOTE-189, and $\hat{P}_{21}$ and $\hat{P}_{20}$ be the estimates of weighted ORRs by PDL1 status for Pembro Combo and Control arms, respectively, from KEYNOTE-407. Then,

$$\hat{P}_{11} = w\hat{P}_{X11} + (1-w)\hat{P}_{X12}$$
$$\hat{P}_{10} = w\hat{P}_{Y11} + (1-w)\hat{P}_{Y12}$$
$$\hat{P}_{21} = w\hat{P}_{X21} + (1-w)\hat{P}_{X22}$$
$$\hat{P}_{20} = w\hat{P}_{Y21} + (1-w)\hat{P}_{Y22}$$

where w is the prevalence of PDL1 TPS≥50%, and

$$\hat{P}_{X11} = \frac{X_{11}}{n_{11}}, \hat{P}_{X12} = \frac{X_{12}}{n_{12}}, \hat{P}_{Y11} = \frac{Y_{11}}{m_{11}}, \hat{P}_{Y12} = \frac{Y_{12}}{m_{12}}, \hat{P}_{X21} = \frac{X_{21}}{n_{21}}, \hat{P}_{X22} = \frac{X_{22}}{n_{22}}, \hat{P}_{Y21} = \frac{Y_{21}}{m_{21}}, \hat{P}_{Y22} = \frac{Y_{22}}{m_{22}}$$

So, the ORR ratio in the combined historical studies can be estimated as:

$$\hat{T} = \frac{c\hat{P}_{11} + (1-c)\hat{P}_{21}}{c\hat{P}_{10} + (1-c)\hat{P}_{20}}$$

with its logarithm being,

$$\log(\hat{T}) = \log\left(c\hat{P}_{11} + (1-c)\hat{P}_{21}\right) - \log\left(c\hat{P}_{10} + (1-c)\hat{P}_{20}\right)$$

Using the delta method, the standard error of log ($\hat{T}$) can be estimated to be:

$$SE(\log(\hat{T})) \approx \sqrt{\frac{c^2\widehat{\sigma_{11}^2} + (1-c)^2\widehat{\sigma_{21}^2}}{(c\hat{P}_{11} + (1-c)\hat{P}_{21})^2} + \frac{c^2\widehat{\sigma_{10}^2} + (1-c)^2\widehat{\sigma_{20}^2}}{(c\hat{P}_{10} + (1-c)\hat{P}_{20})^2}}$$

, where

$$\widehat{\sigma_{11}^2} = w^2\frac{\hat{P}_{X11}(1-\hat{P}_{X11})}{n_{11}} + (1-w)^2\frac{\hat{P}_{X12}(1-\hat{P}_{X12})}{n_{12}}$$

$$\widehat{\sigma_{10}^2} = w^2\frac{\hat{P}_{Y11}(1-\hat{P}_{Y11})}{m_{11}} + (1-w)^2\frac{\hat{P}_{Y12}(1-\hat{P}_{Y12})}{m_{12}}$$

$$\widehat{\sigma_{21}^2} = w^2\frac{\hat{P}_{X21}(1-\hat{P}_{X21})}{n_{21}} + (1-w)^2\frac{\hat{P}_{X22}(1-\hat{P}_{X22})}{n_{22}}$$

$$\widehat{\sigma_{20}^2} = w^2\frac{\hat{P}_{Y21}(1-\hat{P}_{Y21})}{m_{21}} + (1-w)^2\frac{\hat{P}_{Y22}(1-\hat{P}_{Y22})}{m_{22}},$$

$\hat{P}_{X11}$ = CCI $\hat{P}_{X12}$ = CCI, $\hat{P}_{Y11}$ = CCI $\hat{P}_{Y12}$ = CCI $\hat{P}_{X21}$ = CCI $\hat{P}_{X22}$ = CCI $\hat{P}_{Y21}$ = CCI $\hat{P}_{Y22}$ = CCI, and

n_{11} = CCI n_{12} = CCI m_{11} = CCI m_{12} = CCI m_{21} = CCI n_{22} = CCI m_{21} = CCI m_{22} = CCI

Using the data from studies KEYNOTE-189 and KEYNOTE-407, the point estimate and the standard error (SE) for log of ORR ratio from historical studies can be estimated.

Hence, from the synthesis method, the test statistic for noninferiority of ORR in this study comes out to be

$$S = \frac{\log(\hat{\pi}) + 0.5\log(\hat{T})}{\sqrt{SE(\log(\hat{\pi}))^2 + \left(0.5SE(\log(\hat{T}))\right)^2}}$$

, where $\widehat{\pi}$ is the ORR ratio of SC MK-3475A combo versus IV pembro combo in the current study.

The test statistic S approximately follows the standard normal distribution and hence can be compared to upper bound of the 95% CI of the standard normal distribution to assess noninferiority.

4.2 Appendix 2: Approval Information

The sSAP Amendment 01 of Protocol MK-3475A-D77-03 was approved by the BARDS TA head (or designee).

Name: PPD Date: 23-AUG-2024

5 REFERENCES

[Miettinen, O. and Nurminen, M. 1985]	Miettinen O and Nurminen M. Comparative analysis of two rates. Stat Med 1985;4:213-26.	[00VMQY]
[U.S. Food and Drug Administration 2016]	U.S. Food and Drug Administration. Non-Inferiority clinical trials to establish effectiveness guidance for industry [Internet]. Silver Spring, MD: U.S. Department of Health and Human Services; 2016. Available from: https://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf.	[04P3GH]