SUMMIT ANALYTICAL	**STATISTICAL ANALYSIS PLAN** **CONFIDENTIAL**

OSE IMMUNO THERAPEUTICS

STATISTICAL ANALYSIS PLAN

Protocol Title:	A randomized parallel group phase III trial of OSE2101 as 2nd or 3rd line compared with standard treatment (docetaxel or pemetrexed) in HLA-A2 positive patients with advanced Non-Small-Cell Lung Cancer with progressive disease after last treatment with immune checkpoint inhibitors (ICI) (OSE2101C301)
Protocol Number:	OSE2101C301 (ATALANTE 1)
Phase:	Phase III
Sponsor:	OSE Immunotherapeutics 22 boulevard Benoni Goullin 44200 Nantes France
Author:	Summit Analytical, LLC 8354 Northfield Blvd. Bldg. G Suite 3700 Denver, CO 80238
SAP Date:	2021-07-23
Status:	Draft v2.2_augmented after FDA Type C meeting (1st July 2021)

SIGNATURES

The undersigned have approved this Statistical Analysis Plan for use in this study.

Date

Date

Date

1. TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

2. LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Only abbreviations and terms relevant to the SAP are repeated herein. The reader is referred to the protocol for the complete and comprehensive list of abbreviations and definitions of terms for the study.

Abbreviation/Term	Definition
ADaM	Analysis data Model
AE	Adverse Event
ALK	Anaplastic Lymphoma Kinase
ATC	Anatomic Therapeutic Chemical
BMI	Body Mass Index
CDISC	Clinical Data Interchange Standards Consortium
CR	Complete Response
CSR	Clinical Study Report
dNLR	derived Neutrophils/(Leukocyte minus neutrophils Ratio)
DCR	Disease Control Rate
DR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EGFR	Epidermal Growth Factor
HLA-A2	Human Leukocyte Antigen - A2
ICI	Immune Checkpoint Inhibitor
IDMC	Independent Data Monitoring Committee
irAE	Immune-related Adverse Event
ITT	Intention to Treat
LDH	Lactate Dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention to Treat
mPP	Modified Per Protocol
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PD-L1	Programmed Death-Ligand 1

Abbreviation/Term	Definition
PFS	Progression Free Survival
PoI	Population Of Interest
PP	Per Protocol
PR	Partial Response
PT	Preferred Term
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SDTM	Study Data Tabulation Model
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TFL	Tables, Figures, and Listings
TTD	Time to Deterioration
ULN	Upper Limit of Normal Range
WHO-DD	World Health Organization Drug Dictionary

3. INTRODUCTION

3.1. Preface

This document presents the final statistical analysis plan (SAP) for OSE Immunotherapeutics (*A randomized parallel group Phase III trial of OSE2101 as 2nd or 3rd line compared with standard treatment (docetaxel or pemetrexed) in HLA-A2 positive patients with advanced Non-Small-Cell Lung Cancer with progressive disease after last treatment with immune checkpoint inhibitors (ICI [OSE2101C301 – ATALANTE 1]).* The present SAP is intended to plan the analyses to be presented in the final CSR (update of the step 1 CSR).

Reference materials for this statistical plan include the Protocol (Version 5.0 29MAR2019) and Annotated Case Report Form (Version 18.4 10DEC2019).

The present SAP is an extension of the IDMC SAP (Planned step 1 analysis version 1.0 signed the 5 February 2021 – N=103 patients) and describes the SAP for the discontinued step-2 part of this Phase 3 trial.

On March 25, 2020, the step-2 part was ongoing when the IDMC met to review the final Step-1 results and concluded that the predefined thresholds of survival rate at 12 months in OSE2101 arm was achieved (cut-off date of February 26, 2020 (n=103)). This analysis was blinded from the investigators and the Sponsor. At this time, the IDMC was concerned about the impact of the COVID-19 pandemic on the primary endpoint of the study (Overall Survival) and the risk on data integrity. Despite step-1 positive results, the IDMC and the Steering Committee (SC) recommended to halt patients screening and accrual and to continue the treatment of patients enrolled before April 2020.

So, the trial was stopped in April 2020. At that date, 219 patients post ICI have been enrolled and followed for survival up to 15 January 2021 (cutoff date for final analysis).

The present SAP is describing the final analysis on overall post ICI population (N=219 patients).

3.2. Purpose of Analyses

The purposes of the planned analyses described in this SAP are to assess the safety and efficacy of OSE2101 in comparison with standard treatment (docetaxel or pemetrexed). Results from the analyses will be included in the final clinical study report for OSE2101C301, and may also be utilized for regulatory submissions, manuscripts, or other clinical development activities.

The Step-1 results in 103 patients showed a 12-month OS rate of 46.0% [95%CI 33.4%, 59.1%], and the step-1 objectives on 1-year OS rate (H0: OS rate at 12 months >25% and H1 > 40%) were achieved.

An outstanding beneficial effect of OSE2101 in patients with secondary resistance to ICI given as last treatment was observed in step 1. It is proposed to confirm the efficacy and safety of OSE2101 in this population of interest (PoI) as primary analysis from the 219 randomized patients. Considering that the total number of randomized patients is 219 and that the initial hypotheses were 278 events (deaths) in 363 anticipated patients, statistical hypotheses were revised in this SAP for PoI.

3.3. Summary of Statistical Analysis Changes to the Protocol

The SAP update Version 2.0 reflects a major improvement to the SAP in clarity, detailedness, organization and coherence. The SAP update is completed after last patient last visit of this phase 3 trial was done on January 15th, 2021, and database cleaning up to June 2021. Database lock is anticipated in July 2021, after FDA' meeting occurred July 1st, 2021.

The aim of the update is to provide sufficient details to ensure accurate and correct execution of statistical programming of tables, figures, and listings in implementing this plan, improve clarity and coherence of the document to both internal and external stakeholders, as well as correcting some apparent errors and inconsistencies. Some new analyses were also added to ensure comprehensive and robust data analyses and interpretation. To this end, text was significantly added, revised, consolidated, or reorganized in appropriate sections. However, the specifics of study critical details such as those concerning definitions of primary endpoint, significance levels for testing OS are not changed.

Notable changes are delineated below:

- Primary analysis: The primary population of this analysis will be the PoI. PoI is defined as patients with secondary resistance to ICI (ICI $\geq 2^{nd}$ line administered for > 12 weeks).

- Sensitivity analyses on primary endpoint (OS) (described in an ancillary SAP):

    - Additional subgroups (e.g., ICI as maintenance therapy in 1st line, OS with or without anti-cancer treatment after study treatment discontinuation, post-progression survival with or without anti-cancer treatment, duration of stable disease, duration of OSE2101 treatment…)

    - OS according to Best Response to study treatment

    - OS for patients continuing study treatment beyond progression

- Sequential analysis of secondary criteria: The secondary endpoints will focus on the clinical benefit provided by treatments: they will be Post progression Survival, Time to worsening of ECOG, QLQ-C30 Global health status score change from baseline, QLQ-C30 Functional total scores change from baseline, QLQ-C30 total symptom scores change from baseline, QLQ-LC13 total score change from baseline, DCR at 6 and 12 months, and Progression Free Survival.

Consequently, the order of secondary endpoints was updated (see 4.2.4). The hierarchical method in the ordering sequence of secondary endpoints to control the overall type I error remains unchanged. Once the primary objective is established, each of the secondary efficacy endpoints will be tested at 5% significance level using a further hierarchical order in the ordering sequence of the list described above to control the overall type I error. If the primary objective is not established, these endpoints will still be analyzed but interpreted as exploratory endpoints. Similarly if one secondary endpoint in the sequence is not significant, the lower rank secondary endpoint(s) will be considered exploratory.

- Time to ECOG deterioration will be evaluated. It is an accepted endpoint for clinicians who use it to propose treatment strategy in clinical practice. It could be used later in development as a validated secondary endpoint. Complementary analyses will be described in the ancillary SAP.

- Quality of Life analyses were updated and more detailed in this SAP. A complete analysis according to EORTC guidelines (https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf) and the FDA Guidance "Core Patient-Reported Outcomes in Cancer Clinical Trials" (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials) is considered instead of the previous approach replacing the time to deterioration of the 3 selected symptoms. QLQ-C30 Global health status/QoL score change from baseline, QLQ-C30 Functional scores change from baseline and QLQ-C30 Symptom scores change from baseline, and QLQ-LC13 Symptom scores change from baseline were added in the analysis as secondary endpoints per EORTCguidance. As exploratory, quality of life items will be analyzed per FDA guidance to elaborate an instrument to measure the PRO (described in an ancillary SAP). As an example, summary score for the EORTC QLQ-C30 as described in the paper from Giesinger, Journal of Clinical Epidemiology 2015 "Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust" will be explored.

- Sensitivity analyses on PFS (described in an ancillary SAP) as described in FDA guidance (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-non-small-cell-lung-cancer-drugs-and-biologics)

- COVID-19 Impact: In the OSE2101C301 study, 110 patients out of the 219 (50.2%) were potentially impacted by the pandemic. Out of these 110 patients, 87 (79.1%) were included in France, Italy, Spain, or US. Following FDA Guidance on "*Conduct of Clinical Trials of Medical Products During COVID-19 Public Health Emergency: Guidance for Industry, Investigators, and Institutional Review Boards.*" and "*Statistical Considerations for Clinical Trials During the COVID-19 Public Health*

Emergency: Guidance for Industry", it is proposed to test COVID-19 impact in the ITT population using:

  - Time-dependent OS survival analyses to test the heterogeneity of the periods of randomization with or without COVID-19.

  - OS analysis using different cut-off dates for censored data as previously proposed (www.ema.europa.eu/en/implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical-trials; (Degtyarev et al., 2020; Meyer et al., 2020)

- Baseline data of first line ICI population versus second line ICI and PoI population versus non-PoI population (described in an ancillary SAP).

4. STUDY OBJECTIVES AND ENDPOINTS

Study objectives and endpoints defined in the protocol include safety and efficacy endpoints. Objectives and pre-specified endpoints are as follows.

4.1. Study Objectives

4.1.1. Primary Objective(s)

The primary objective of the discontinued step-2 is to demonstrate that in the population of interest (patients with secondary resistance to ICI), OSE2101 is superior to control treatment with respect to OS in HLA-A2 positive patients with advanced NSCLC after failure to platin-based chemotherapy and ICIs.

4.1.2. Secondary Objectives

The secondary objectives of the discontinued step-2 are as follows:

- To compare secondary measures of clinical benefit including Post progression Survival, Time to worsening of ECOG, QLQ-C30 Global health status score change from baseline, QLQ-C30 Functional scores change from baseline, QLQ-C30 Symptom scores change from baseline, QLQ-LC13 Symptom scores change from baseline, DCR at 6 months, DCR at 12 months, and Progression-Free Survival (PFS) in the PoI
- To assess the safety and tolerability of OSE2101 compared to the control treatment in PoI.

Exploratory objectives are to compare other criteria: ORR, Duration of Response, and time to next lung cancer therapy in both treatment arms in PoI.
Safety, tolerability, and efficacy including OS will also be done in ITT overall population as exploratory.

4.2. Efficacy Endpoints in step-2

4.2.1. Primary Efficacy Endpoint

- Overall survival (OS)

4.2.2. Secondary Efficacy Endpoints

- Post Progression Survival
- Time to worsening of ECOG
- QLQ-C30 Global health status score change from baseline
- QLQ-C30 Functional scores change from baseline
- QLQ-C30 Symptom scores change from baseline
- QLQ-LC13 Symptom scores change from baseline

- Disease Control Rate (DCR) at 6 months and 12 months
- Progression-Free Survival (PFS)

Once the primary objective will be established (null hypothesis rejected at 5% two-sided significance level), each of the secondary efficacy endpoints will be tested at 5% significance level using a further hierarchical order in the ordering sequence of the list described above to control the overall type I error. If the primary objective is not established, these endpoints will still be analyzed but interpreted as exploratory endpoints, similarly if one secondary endpoint in the sequence is not significant, the lower rank secondary endpoint(s) will be considered exploratory.

4.2.3. Exploratory Efficacy Endpoints

- Time to next lung cancer therapy
- Objective Response Rate (ORR)
- Duration of response (DR)
- All the other scores of the quality of life questionnaires

4.2.4. Summary of the primary and secondary endpoints

The list of primary and secondary endpoints can be summarized as follows:

Population	Efficacy endpoint	Status
PoI-ITT	OS	P
PoI-PP	OS	S01
PoI-ITT	Post progression Survival	S02
PoI-ITT	Time to worsening of ECOG	S03
PoI-ITT	QLQ-C30 Global health status change from baseline	S04
PoI-ITT	QLQ-C30 Functional total scores change from baseline (Physical, Role, Emotional, Cognitive, and Social)	S05
PoI-ITT	QLQ-C30 Symptom scores change from baseline (Fatigue, Nausea & Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, and Financial Difficulties)	S06
PoI-ITT	QLQ-LC13 Symptom scores change from baseline change from baseline	S07
PoI-ITT	DCR at 6 months	S08
PoI-ITT	DCR at 12 months	S09
PoI-ITT	Progression Free Survival	S10

P: primary, S: secondary, PoI: Population of Interest

4.3. Safety and Tolerability Endpoints

- The incidence, severity, seriousness, and relationship to study treatment of adverse events (AEs) and immune-related adverse events (irAEs)
- The abnormalities in safety laboratory parameters
- The abnormalities in vital signs
- The abnormalities in 12-lead ECGs

5. STUDY METHODS

5.1. General Study Design and Plan

As background for the statistical methods presented below, this section provides an overview of the study design and plan of study execution.

This is an open-label, multicenter, controlled, randomized, parallel group, two-stage design phase II/III trial of OSE2101 (Arm A – experimental treatment) versus docetaxel or pemetrexed (Arm B – control treatments) as 2nd or 3rd line treatment in HLA-A2 positive patients with advanced NSCLC with progressive disease after last treatment with ICI.

Randomization to Arms A and B was stratified according to histology (squamous vs. non-squamous), initial response to first line treatment (objective response: complete or partial response, vs. no objective response: stabilization or progressive disease), and - from version 4.1 of the protocol - previous treatment with immune checkpoint inhibitors (used in 1st line vs. used in 2nd line). Before this amendment, the 3rd randomization stratification factor was previous treatment with immune checkpoint inhibitor (yes vs. no).

The schedule for assessments and timing of events is presented in Tables 1 and 2 (Arm A and Arm B respectively).

Table 1 Schedule of Time and Events for Arm A

Period	Pre-screening	Screening	Treatment					End of treatment	
Visit		Screening and randomization	Cycle 1 (3 week-cycle)		Cycles 2 to 6 (3 week-cycle)	Cycles 7 to 9 (8 week-cycle)	Cycles ≥ 10 or EOT (12-week cycle)	End of treatment[1]	Post-treatment Follow-up[2]
Cycle day		-35 Day/ -1 Day	Day 1 (±2)	Day 15 (±2)	Day 1 (±2)	Day 1 (±4)	Day 1 (±8)		
Informed consent for HLA-A2 determination and biobanking	X								
HLA-A2 determination[3]	X								
Informed consent		X							
Medical/ Oncological history		X							
Previous medications		X							
EGFR/ALK alterations, PD-L1[4]		X							
Body height		X							
Body weight		X	X		X	X	X	X	
Physical examination		X	X		X	X	X	X	
Vital signs (Body temperature, BP/PR, oxygen saturation)		X	X		X	X	X	X	
12-lead ECG		X			X	X	X	X	
ECOG Performance status		X	X		X	X	X	X	
Hematology/ Biochemistry[5]		X	X	X	X	X	X	X	
Coagulation/ HIV test[6]		X							
Pregnancy test (as appropriate)		X						X	
Tumor assessments (including scans)		X	every 6 (± 1) weeks until documented RECIST 1.1 progression[7]						(X)[8]
Randomization		X							
OSE2101 administration[9]			X		X	X	X		
EORTC QLQ-C30, QLQ-LC13			X		X	X	X	X	
Survival follow-up									X
Concomitant medications		X	X	X	X	X	X	X	
Adverse Events		X	X	X	X	X	X	X	
Serious Adverse Events		X	X	X	X	X	X	X	
Hospitalizations		X	X	X	X	X	X	X	
Informed consent for translational study and biobanking (optional)[10]		X							
Blood samples for biomarkers (optional)[11]		X			every 6 (± 3) weeks (e.g. at D1 of Cycles 3, 5 & 7)	every 12 (± 3) weeks (e.g. at time of imaging) until EOT and/or progression			
Tumor samples for biomarkers (optional)[12]		X			X (e.g. between Cycles 3 & 6)			X (e.g. at tumor progression)	

[1] End of treatment: visit to obtain last assessments after the study drug stop from the 4 previous weeks.
[2] Post-treatment follow-up: data collection on survival status and any subsequent cancer therapy (if any) every 2 months after study drug discontinuation until death.
[3] HLA-A2 testing can be done at any time during patient management before inclusion. A pre-screening consent is available to test patients for instance during 1st or 2nd line therapy.
[4] EGFR/ALK alterations and PD-L1 expression on tumor tissue: if these results are not available, they have to be done in local or central laboratories during the screening period. EGFR/ALK alterations is mandatory before randomization in patients with non-squamous cancer or never/light smokers with squamous cancer. If PD-L1 expression is not documented and no tumor tissue available to perform the testing, then the PD-L1 status can be omitted.
[5] Hematology and biochemistry tests: to be done at Day 1 Cycle 1 only if screening assessment done more than 7 days before. Thereafter at Day 1 of each subsequent cycle, local laboratory tests will be done every 3 weeks (less than 72 hours (3 days) before study treatment administration) until Cycle 6, then every 8 weeks (less than 4 days before study administration) until Cycle 9, then every 12 weeks (less than 8 days before each subsequent administration.
[6] HIV test: to be done for all eligible patients with HIV status unknown.
[7] Tumor assessments (CT scan of thorax, abdomen and brain CT or brain MRI) to be done every 6 weeks (± 1-week allowance) after Day 1 Cycle 1, meaning it could be between 2 treatment visits; CT scan of the pelvis will be performed only upon request according to clinical suspicion. In case of clinical suspicion of bone metastasis out of the area reported by CT scan, imaging test will be performed under discretion of the investigator and will be repeated every 12 weeks. For Germany, with respect to brain metastases, MRI should be the privileged imaging technic and CT should only be considered when there are contra-indications to MRI.
[8] Tumor assessments should be continued until documented RECIST 1.1 disease progression for patients who have not progressed during the treatment period.
[9] OSE2101 treatment will continue until unequivocal RECIST 1.1-defined disease progression as determined by the investigator, unacceptable toxicity, or consent withdrawal. Should pseudo progression or delayed response to treatment suspected in arm A, investigator may continue treatment beyond the time of RECIST-defined progression, if the patient is perceived to be experiencing clinical benefit.
[10] Separate informed consent (optional) will be obtained from the patient who agreed to participate in a translational study and authorizing the bio banking of the remaining samples to be used for future biomarker research and drug development projects. Patient can agree to participate in all and/or some parts of the translational study. Patient can refuse to participate in the translational study without compromising his/her participation in the main study.
[11] Blood samples for biomarkers will be taken at baseline (during the screening period), then every 6 (± 3) weeks until Cycle 7 (e.g., at D1 of Cycles 3, 5 and 7), then every 12 (± 3) weeks (e.g., at time of imaging) until end of treatment and/or progression.
[12] Specimens from archiving tumor biopsies (at first diagnosis and/or at any time during previous treatment) will be collected. As possible, a new tumor biopsy will be collected within 35 days prior to the first study treatment administration (during the screening period). An archiving tumor biopsy collected before the screening period is acceptable, should the biopsy be collected after the end of ICI treatment. A second tumor biopsy (e.g., in the same tumor site) will be undertaken between Cycle 3 and Cycle 6 and/or at time of the disease progression. A blood sample will be collected for Whole-Exome sequencing (WES) at screening.

Table 2 Schedule of Time and Events for Arm B

Period	Pre-screening	Screening	Treatment				End of treatment	
Visit		Screening and randomization	Cycle 1 (3 week-cycle)		Cycle 2 (3 week-cycle)	Cycles ≥3 (3 week-cycle)	End of treatment[1]	Post-treatment Follow-up[2]
Study day		-35 Day/ -1 Day	Day 1 (±2)	Day 15 (±2)	Day 1 (±2)	Day 1 (±2)		
Informed consent for HLA-A2 determination and biobanking	X							
HLA-A2 determination[3]	X							
Informed consent		X						
Medical/ Oncological history		X						
Previous medications		X						
EGFR/ALK alterations, PD-L1[4]		X						
Body height		X						
Body weight		X	X		X	X	X	
Physical examination		X	X		X	X	X	
Vital signs (Body temperature, BP/PR, oxygen saturation)		X	X		X	X	X	
12-lead ECG		X			X	X	X	
ECOG Performance status		X	X		X	X	X	
Hematology/ Biochemistry[5]		X	X	X	X	X	X	
Coagulation/ HIV test[6]		X						
Pregnancy test (as appropriate)		X					X	
Tumor assessments (including scans)		X	every 6 (± 1) weeks until documented RECIST 1.1 progression[7]					(X)[8]
Randomization		X						
Docetaxel or Pemetrexed[9]			X		X	X		
EORTC QLQ-C30, QLQ-LC13			X		X	X	X	
Survival follow-up								X
Concomitant medications		X	X	X	X	X	X	
Adverse Events		X	X	X	X	X	X	
Serious Adverse Events		X	X	X	X	X	X	
Hospitalizations		X	X	X	X	X	X	
Informed consent for translational study and biobanking (optional)[10]		X						
Blood samples for biomarkers (optional)[11]		X	every 6 (± 3) weeks (e.g. at D1 of Cycles 3, 5 & 7), then every 12 (± 3) weeks (e.g. at time of imaging) until EOT and/or progression					
Tumor samples for biomarkers (optional)[12]		X				X (e.g. between Cycles 3 & 6)	X (e.g. at tumor progression)	

[1] End of treatment: visit to obtain last assessments after the study drug stop from the 4 previous weeks.
[2] Post treatment follow-up: data collection on survival status and any subsequent cancer therapy (if any) every 2 months after study drug discontinuation until death.
[3] HLA-A2 testing can be done at any time during patient management before inclusion. A pre-screening consent is available to test patients for instance during 1st or 2nd line therapy.
[4] EGFR/ALK alterations and PD-L1 expression on tumor tissue: if these results are not available, they have to be done in local or central laboratories during the screening period. EGFR/ALK alterations is mandatory before randomization in patients with non-squamous cancer or never/light smokers with squamous cancer. If PD-L1 expression is not documented and no tumor tissue available to perform the testing, then the PD-L1 status can be omitted.
[5] Hematology and biochemistry tests: to be done at Day 1 Cycle 1 only if screening assessment done more than 7 days before. Thereafter at Day 1 of each subsequent cycle, local laboratory tests will be done every 3 weeks (less than 72 hours (3 days) before study treatment administration) until Cycle 6, then every 8 weeks (less than 4 days before study administration) until Cycle 9, then every 12 weeks (less than 8 days before each subsequent administration.
[6] HIV test: to be done for all eligible patients with HIV status unknown.
[7] Tumor assessments (CT scan of thorax, abdomen and brain CT or brain MRI) to be done every 6 weeks (± 1-week allowance) after Day 1 Cycle 1 meaning it could be between 2 treatment visits; CT scan of the pelvis will be performed only upon request according to clinical suspicion. In case of clinical suspicion of bone metastasis out of the area reported by CT scan, imaging test will be performed under discretion of the investigator and will be repeated every 12 weeks. For Germany, with respect to brain metastases, MRI should be the privileged imaging technic and CT should only be considered when there are contra-indications to MRI.
[8] Tumor assessments should be continued until documented RECIST 1.1 disease progression for patients who have not progressed during the treatment period.
[9] Docetaxel and pemetrexed treatment will continue until unequivocal RECIST 1.1-defined disease progression as determined by the investigator, unacceptable toxicity, or consent withdrawal. Premedication is required with docetaxel and pemetrexed. Premedication will occur before Day 1 (i.e., Day 1 of each cycle will be injection day for chemotherapy).
[10] Separate informed consent (optional) will be obtained from the patient who agreed participating in a translational study and authorizing the bio banking of the remaining samples to be used for future biomarker research and drug development projects. Patient can agree to participate in all and/or some parts of the translational study. Patient can refuse to participate in the translational study without compromising his/her participation in the main study.
[11] Blood samples will be taken at baseline (during the screening period), then every 6 (± 3) weeks until Cycle 7 (i.e., at D1 of Cycles 3, 5 and 7), then every 12 (± 3) weeks (e.g., at time of CT imaging) until end of treatment and/or progression.
[12] Specimens from archiving tumor biopsies (at first diagnosis and/or at any time during previous treatment) will be collected. As possible, a new tumor biopsy will be collected within 35 days prior to the first study treatment administration (during the screening period). An archiving tumor biopsy collected before the screening period is acceptable, should the biopsy be collected after the end of ICI treatment. A second tumor biopsy (e.g., in the same tumor site) will be undertaken between Cycle 3 and Cycle 6 and/or at time of the disease progression. A blood sample will be collected for Whole-Exome sequencing (WES) at screening.

5.2. Inclusion – Exclusion Criteria and General Study Population

The inclusion and exclusion criteria defined in the protocol apply to all subjects and are not repeated herein the SAP.

5.3. Randomization and Blinding

All eligible and consented subjects will receive OSE2101 (Arm A) or docetaxel or pemetrexed (according to histology and previous treatment) (Arm B).

The final analysis of step-2 will include a combination of patients recruited before amendment (ratio=1:1; 38 patients) and after amendment in 2017 (ratio=2:1, 181 patients).

5.4. Analysis Variables

Variables to be analyzed include demographics and baseline characteristics, safety variables (adverse events, concomitant medications, clinical laboratory investigations, vital signs, ECOG and ECG), efficacy variables (tumor assessments and survival status), and quality of life variables.

6. SAMPLE SIZE

Step 2 – Phase III sample size for all ICI population

The initial hypotheses in the protocol version 4.0 were: “To reach a power of 80% for the two-sided log rank test at the 5% significance level when assuming a median OS of 7 months for the control and 10 months for the experimental arm, a total of 278 events is required (HR=0.70). If the new patients are uniformly included over an accrual period of 36 months followed by a follow-up of 6 months after the last inclusion, observing 278 events would require a total of 363 new patients.”

The Phase 3 sample size consists in 219 patients. With this final sample size of 219 patients, the power decrease to 62% with the same hypotheses (HR=0.70).

Considering Step 1 results, statistical hypotheses have been revised to address the PoI as follow:

Assuming a HR of 0.55 consistent with that observed in Step-1 (HR stratified =0.67, HR unstratified = 0.52) with a 7 months OS in the SoC group (unchanged as initially planned in the protocol), a total of 90 events in the PoI will provide a power of 80% and a 2-sided logrank test at a 5% two-sided level, if the patients are included over an accrual period of 48 months followed by a follow-up of 6 months after the last inclusion.

7. GENERAL CONSIDERATIONS

7.1. Analysis Sets

Six different analysis sets are defined in the post ICI population (N=219). The 3 analysis sets will be duplicated in the population of interest and in the overall post ICI population.

For safety analyses, baseline is the first dose of study treatment and for « baseline characteristics » section, baseline is defined as last value before randomization.

7.1.1. Safety Set PoI

This analysis set includes all patients included in the PoI who were randomized and received at least one dose of randomized study treatment. Patients will be analyzed in the group according to the treatment they received. All safety data analyses will be analyzed using the safety set.

7.1.2. Intent-to-Treat Set PoI (ITT_PoI)

This analysis set includes all patients included in the PoI and randomized. Patients will be analyzed in the group as randomized. All efficacy data analyses will be analyzed using the ITT_PoI set.

7.1.3. Per-Protocol Set PoI (PP_PoI)

This analysis set includes all patients of the ITT_PoI set who did not violate the protocol in a way that might affect the evaluation of the effect of the study drugs on the overall survival. The patients to be excluded from the PP_PoI set will be identified before performing the analysis. Selected sensitivity efficacy data analyses will be performed using the PP_PoI set.

7.1.4. Safety Set

This analysis set includes all patients included who were randomized and received at least one dose of randomized study treatment. Patients will be analyzed in the group according to the treatment they received. All safety data analyses will be analyzed using the safety set.

7.1.5. Intent-to-Treat Set (ITT)

This analysis set includes all patients included and randomized. Patients will be analyzed in the group as randomized. All efficacy data analyses will be analyzed using the ITT set.

7.1.6. Per-Protocol Set (PP)

This analysis set includes all patients of the ITT set who did not violate the protocol in a way that might affect the evaluation of the effect of the study drugs on the overall survival. The patients to be excluded from the PP set will be identified before performing the analysis. Selected sensitivity efficacy data analyses will be performed using the PP set.

The exclusion of patients from the PP_PoI and PP sets was prepared using the minutes and the deviations report of the blind review meeting that was held on 08 April 2021.

7.2. Subgroups

The following variables will be used to define subgroups:

Stratification criteria
- Histology (squamous *vs.* non-squamous),
- Best response to 1st line treatment (CR/PR *vs.* SD/PD)
- Line rank of ICI therapy (2nd line vs. 1st line)*

Other criteria
- Gender (male *vs.* female)
- Age (< 75 *vs.* ≥ 75)
- Age (< 65 *vs.* ≥ 65)
- Smoking status: never, former, or current smoker
- Baseline ECOG score (0 *vs*. 1)
- Best response to 1st line treatment (CR/PR/SD *vs.* PD)
- Best response to 1st line treatment (CR/PR vs SD *vs.* PD)
- Best response to ICI therapy (CR/PR *vs.* SD vs PD)
- Best response to ICI therapy (CR/PR/SD *vs.* PD)
- Duration of previous ICI therapy (weeks: 0-12, > 12-24, > 24 weeks)*
- Duration of previous ICI therapy (weeks: 0-12, > 12 weeks)*
- Time from end of ICI to study treatment initiation (weeks: 0-6, 6-12, > 12 weeks)
- Time from end of chemotherapy to study treatment initiation (weeks: 0-24, 24-48, > 48 weeks)
- Duration of previous ICI therapy (>12 weeks) and Line rank of ICI therapy (2nd line)*
- Duration of previous ICI therapy (>12 weeks) and Line rank of ICI therapy (1st line)*
- Liver metastases at study entry (presence *vs.* absence)
- Brain metastases at study entry (presence *vs.* absence)
- Pleural metastases at study entry (presence *vs.* absence)
- Number of different metastasis locations (0, 1, 2, ≥ 3)
- Cancer staging at study entry (III *vs.* IV)
- dNLR (derived Neutrophils/(Leukocyte minus neutrophils Ratio)) at baseline (< 3 vs ≥ 3)
- LDH Classification at baseline (<= upper limit of normal range (ULN), > ULN)
- PDL1 (Positive, Negative, Unknown)

- Albumin Classification at baseline (<35g/l, ≥35g/l)

*: only in ITT or PP populations and not in ITT_PoI or PP_PoI

'dNLR (derived Neutrophils/(Leukocyte minus neutrophils Ratio))' must be read as 'dNLR (derived Neutrophils/(Leukocyte minus Neutrophils) Ratio)'

Additional subgroups will be explored in the ancillary SAP.

7.3. Management of Analysis Data

7.3.1. Missing Data

In general, there will be no substitutions made to accommodate missing data points. All data recorded on the electronic case report form (e-CRF) will be included in data listings that will accompany the clinical study report.

7.3.1.1. Handling of Missing Date Values

Partial or Missing Dates

The following conventions will be used to impute missing portions of dates for adverse events and concomitant medications, if warranted. Note that the imputed values outlined here may not always provide the most conservative date. In those circumstances, the imputed value may be replaced by a date that will lead to a more conservative analysis.

A. Start Dates
   1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.
   2) If the month is unknown, then:
      i) If the year matches the first dose date year, then impute the month and day of the first dose date.
      ii) Otherwise, assign 'January.'
   3) If the day is unknown, then:
      i) If the month and year match the first dose date month and year, then impute the day of the first dose date.
      ii) Otherwise, assign the first day of the month.

B. Stop Dates
   1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.
   2) If the month is unknown, then assign 'December.'
   3) If the day is unknown, then assign the last day of the month.

Unless having an end date before the first administration of the study drug, an adverse event with a completely missing start date will be interpreted as started between the first dose date included and 28 days after last administration of the study treatment (included) or the analysis cut-off date included, whichever comes first.

No imputation should be used for the start date of immune-related adverse events recorded in medical history (MH2 VIEDOC dataset).

A medication with a completely missing start date will be interpreted as started before the first dose date.

A medication with a completely missing end date will be interpreted as ongoing at the time of data extraction (and therefore being not stopped before the first dose).

7.3.1.2. Imputation Methods

No safety or efficacy data will be imputed for this study. All data will be observed cases, without imputation, except specified otherwise (as for time to worsening of ECOG or quality of life sensitivity analyses as examples).

7.3.2. Coding Conventions for Events and Medications

All adverse events, and medical history will be mapped to the Medical Dictionary for Regulatory Activities (MedDRA Version 20.1) system for reporting (preferred term and body system).

Prior and Concomitant medications will be coded using WHO-DD (Drug Dictionary) (Version March 2018).

7.3.3. Analysis Software

Data manipulation, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed primarily using SAS (release 9.4 or higher) for Windows. If the use of other software is warranted, the final clinical study report will detail what software was used and for what purposes.

7.3.4. Study Data

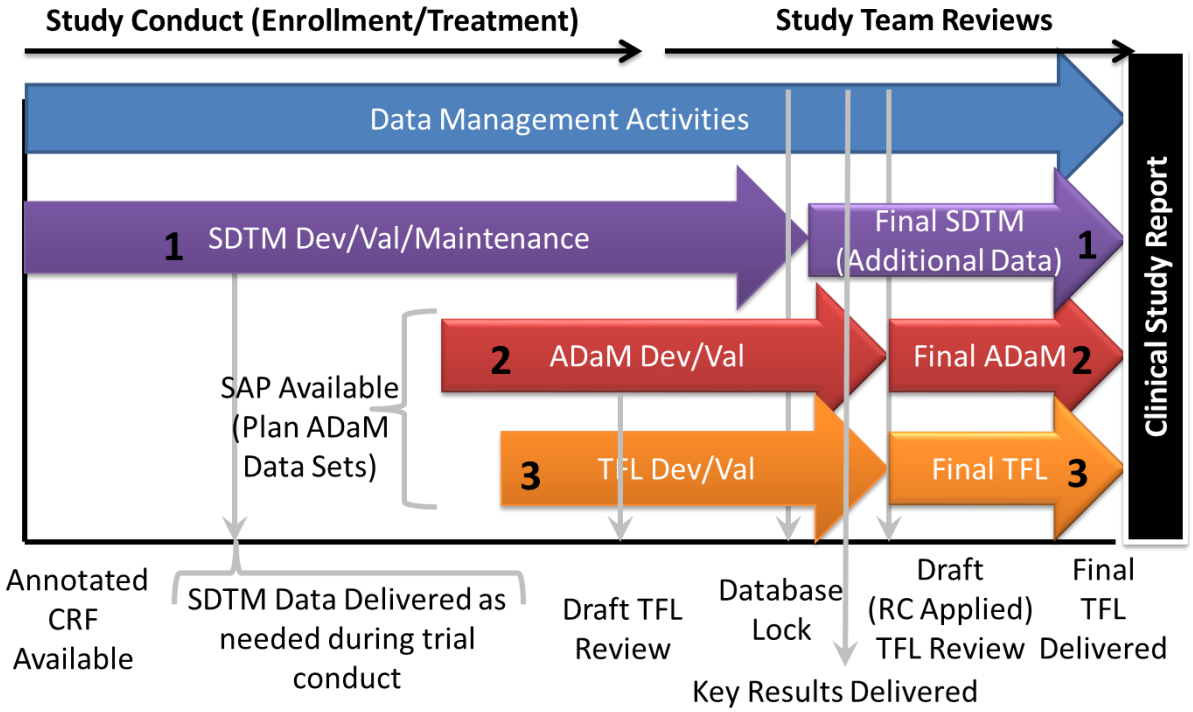
Study data identified in the schedule for time and events in the Protocol are collected, and source verified, on the electronic data capture tool: VIEDOC (PCG Solutions, Uppsala, Sweden).

All study data will be formulated into regulatory compliant data sets to provide transparency, traceability, and integrity of trial analysis results from the collection source. Observed study data will be mapped to the CDISC Study Data Tabulation Model (SDTM) and serve as the

source data from the trial. All study analyses will be completed using analysis data sets that are derived from the SDTM and follow the CDISC Analysis Data Model (ADaM) architecture.

The methods for programming the CDISC SDTM and ADaM data sets are described in Figure 1.

Figure 1 SDTM, ADaM, and TFL Development and Validation



Where:

1. Development, Validation, and Maintenance of SDTM Domains
2. Development and Validation of Analysis Data Sets (ADaM), with input source the appropriate SDTM domains.
3. Development and Validation of Tables, Figures, and Listings (TFL), with input data source the SDTM domains and analysis specific ADaM data sets.

7.4. Planned Study Analyses

7.4.1. Statistical Summaries: Descriptive and Inferential

All statistical tests will be two-sided and a difference resulting in a p-value of less than or equal to 0.05 will be considered statistically significant (although possibly of exploratory nature). All p-values will be rounded to and displayed in four decimals. If a p-value less than 0.0001 occurs it will be shown in tables as <0.0001.

Descriptive summaries of variables will be provided where appropriate. For continuous variables, the number of non-missing values (n), mean, standard deviation, median, minimum, and maximum will be tabulated by treatment. For categorical variables, the counts and proportions of each value will be tabulated by treatment. Expansion of descriptive table categories within each treatment may occur if such elaborations are thought to be useful.

All study related data collected will be presented in listings. Study related data not subject to analysis according to this plan will not appear in any tables or graphs but will be included in the data listings.

7.4.2. Interim Analyses and Data Monitoring

No interim analysis of the primary endpoint was planned.

However, OS rate at 12 months was to be analyzed on Arm A at end of Step 1 before a decision to stop or continue enrollment in Step 2 could be made by the Steering Committee after recommendation by the IDMC. This analysis was considered as a futility analysis. Consequently, no adjustment for multiplicity (as combination test for example) was considered for the primary analysis of step 2.

The enrolment stop because of the foreseen impact of the COVID-19 pandemic on the primary endpoint of the step 2 was not planned.

7.4.3. Final Analysis and Publication of Study Results

This analysis (cut-off date: 15 January 2021 - 219 post ICI patients) is considered as an update of the step 1 analysis. This CSR will be complemented by an addendum with this analysis.

8. SUMMARY OF STUDY DATA

All the analyses described in this section will be duplicate in the population of interest and in the overall population.

8.1. Subject Disposition

A summary of the analysis sets considered in the CSR will include the number and percentage of subjects for the following categories: subjects randomized, subjects in the ITT sets, subjects not treated, subjects in the Safety sets, and subjects having permanently discontinued the study treatment, subjects in the PP sets. All percentages will be based on the number of subjects randomized.

A table will present by treatment arm on the ITT sets the primary reason for not being treated and the primary reason for permanent treatment discontinuation (with sub-counts of patients having entered the follow-up and patients having not entered the follow-up).

A table will present by treatment arm on the ITT sets the number and percentage of patients randomized by month.

8.2. Patients Randomized by Month

The number and percent of subjects randomized will be summarized by month by treatment arm for the ITT sets.

8.3. Protocol Deviations

Protocol deviations on inclusion/exclusion criteria will be summarized by treatment arm on the ITT sets.

Major protocol deviations, as finalized during the blinded review of the data prior to database lock (organized by the Sponsor with the participation of two coordinating investigators and of one independent expert) may result in the removal of subjects from the PP sets. The Sponsor or designee will be responsible for producing the final major deviation file; this file will include a description of the protocol deviation leading to the exclusion from the PP sets. This file will be finalized prior to database lock, and all information will be included in the SDTM.DV domain (deviations domain). A summary table of the major protocol deviations will be generated on the ITT sets.

8.4. Demographics and Baseline Characteristics

Subject demographic data and baseline characteristics will be tabulated and summarized descriptively by treatment arm and overall. The demographic data and baseline characteristics

will be summarized for the ITT sets. Individual subject demographics and baseline characteristics will be provided in listings.

The demographics consist of age, gender, race, ethnicity, smoking status, BMI, BMI classification and country.

The country will be derived from the 2 leading digits of the site as follows:

Leading digits of site identification	Country
11	Czech Republic
12	France
13	Germany
14	Hungary
15	Italy
16	Poland
17	Spain
18	United Kingdom
19	USA
22	Israel

The baseline characteristics consist of:

- Histology (Squamous/Non-Squamous)
- Initial (i.e., best) response to 1st line treatment (CR/PR, SD/PD)
- Previous treatment with ICI
- Line rank of previous treatment with ICI
- Previous Pemetrexed treatment
- Line rank of study treatment
- Previous radiotherapy combined with chemotherapy
- Best response during 1st line therapy
- Time from end of previous line of therapy to study treatment initiation (weeks)
- Best response during ICI therapy
- Time from end of ICI to study treatment initiation (weeks)
- Time from end of chemotherapy to study treatment initiation (weeks)

- Duration of ICI therapy (weeks: 0-12, >12-24, > 24)
- Cancer staging at diagnosis (I/II, IIIA/IIIB/IIIC, IVA/IVB)
- Cancer staging at study entry (III, IV)
- T classification at diagnosis

- N classification at diagnosis
- M classification at diagnosis
- M classification at study entry
- Location of metastases at diagnosis
- Location of metastases at study entry
- Number of different metastasis locations
- ECOG score at baseline
- HLA-A2 (negative, positive) as documented in NSCLC disease history
- EGFR (negative, positive) as documented in NSCLC disease history
- PD-L1 (negative, positive) as coded in the Data Management manually derived variable DIS_HIST.PDL1_POSNEG_D
- Albumin Classification at baseline (< 35 g/L/>= 35 g/L)
- LDH Classification at baseline (<= upper limit of normal range (ULN), > ULN)
- dNLR at baseline (< 3 vs ≥ 3)
- Antibiotic treatment within 30 days before randomization
- Corticosteroid treatment within 30 days before randomization

For the baseline characteristics, the baseline is defined as the last value available up to the date of randomization included.

Previous medications for NSCLC were reviewed before database lock and identified as ICI, chemotherapy other than pemetrexed, pemetrexed, other. To get the drug name of the ICI treatment, a manually prepared Excel file (‘210415 Previous ICI description.xlsx') will also be used. This Excel file provides for each patient the name (variable ICI_DCI) and the class (variable Class) of the ICI treatment and a technical variable (ICI_DCI_link) allowing to match the other ICI information recorded in PRIOTHER.

The line rank of ICI therapy was manually derived by Data Management after medical and clinical review of each patient prior medications. The category '3rd line therapy', recorded for one patient, will be displayed in the description of baseline characteristics but pooled with the category '2nd line therapy' in the overall survival subgroup analyses.

To derive Best response during ICI therapy, Time since end of ICI therapy, Duration of ICI therapy, additional information about the ICI (best response, start date, end date) will be recovered by respectively retaining the best "best response" across records, the earliest start date across records, and the latest end date across records.

An additional table will be provided presenting per arm and overall, the two variables Best response during ICI therapy crossed with duration of ICI therapy (weeks: 0-12, >12-24, > 24).

The dataset DIS_HIST was collecting cancer stage according to TNM V7.0 until June 2017, i.e., during the first part of the study. The study restarted in June 2018. In the meantime, protocol 5.0 dated 15 March 2018 had introduced the use of TNM V8.0. Therefore, the dataset DIS_HIST collected cancer stage according to TNM V8.0 for all patients of the second part of the study. Therefore, cancer stage was reviewed and sometimes updated by clinical expert

before database lock. Reclassified cancer staging value will be analyzed, categorized as follows:

- Cancer stage at diagnosis: (I/II) mapped to 'I/II', (IIIA, IIIB, IIIC) mapped to 'IIIA/IIIB/IIIC', (IV, IVA, IVB) mapped to 'IVA/IVB'
- Cancer stage at study entry: (IIIB, IIIC) mapped to 'III', (IV, IVA, IVB) mapped to 'IV', (IIIA, Other) mapped to 'Other'.

The category 'Other' will be displayed in the description of baseline characteristics but not retained in the overall survival subgroup analyses.

Liver, Bone, and Brain will be counted as a metastasis location if the number of metastases reported is in (1, 2, 3, >3, NS). Other will be counted as a metastasis location if the number of metastases reported is in (1, 2, 3, >3, NS) for at least one of the 'Other' locations. Adrenal will be counted as a metastasis location if reported as ('Unilateral','Bilateral',NS). Pleura will be counted as a metastasis location if reported as 'Present'.

'Number of different metastasis locations' must be understood as 'Number of different metastasis locations' at study entry'. This number is left missing if the M classification at study entry is missing.

Protocol section 4.3.3 requires that patients of arm B take dexamethasone (a corticosteroid) twice daily before, the day of and the day after study drug (docetaxel or pemetrexed). Therefore, comparing the 2 treatment arms on presence/absence of corticosteroids on the day before randomization or on the day of randomization is biased. 'Corticosteroid treatment within 30 days before randomization' is changed to 'Corticosteroid treatment within 2 to 30 days before randomization'.

The Data Management manually derived variables WHODRUG.ANTIBIO_D and WHODRUG.CORTI_D identify systemic antibiotic treatments and corticosteroid treatments without selection on their administration dates. Start and end dates of these treatments will be checked versus the randomization date to identify antibiotic treatment within 30 days before randomization and corticosteroid treatment within 2 to 30 days before randomization.

8.5. Previous ICI Treatment

The number and percentages of patients by previous ICI treatment will be summarized on the ITT sets by treatment arm, Anatomical Therapeutic Chemical (ATC) level 4 (sorted according to the corresponding 5-digit ATC code), and PT. The previous ICI treatment information will be taken from information entered (not necessarily in a consistent way with the actual ICI line rank) for 1st line therapy, 2nd line therapy or 3rd line therapy in prior medications for NSCLC based on the Data Management manually derived variable PRIOTHER.LINE_ICI_D.

A listing describing for the most used ICI treatments as nivolumab, pembrolizumab, atezolizumab, avelumab, durvalumab with a complete description will be provided (dose, treatment duration, response, …).

The description of the previous ICI treatment by class will be prepared from the Excel file ('200720 Previous ICI description.xlsx') as described in section 8.4 under 'Addition or change after 04 June 2020'.

8.6. History of Immune-Related Adverse Events

The number and percent of subjects with a history of irAEs will be summarized by irAE category and preferred term by treatment arm for the ITT sets. The same presentation will be used to summarize the irAEs ongoing at screening.

8.7. Randomization Stratification Factors

Randomization stratification factors will be tabulated and summarized descriptively by treatment arm and overall, for the ITT sets.

The following strata will be summarized:
- Squamous/Non-squamous
- Objective response/No objective response
- First line therapy/Second line therapy.

The values reported for the stratification factors will be those possibly corrected by the site after time of randomization and therefore may be different of the values used for the randomization. For analyses, corrected from eCRF (and not IWRS) values will be used.

For the 38 ICI patients randomized before the amendment (i.e., at a time where the stratification factor First line therapy/Second line therapy was not yet applicable), the Data Management manually derived IC.RANDIMM2_1_D will be used in place of IC.RANDIMM2_1 (the data field IC.RANDIMM2_1was not existing in the database at the time the "Old" ICI were randomized).

Randomization stratification factors as entered in the IVRS/IWRS at time of randomization will also be tabulated and summarized descriptively by treatment arm and overall, for the ITT set, separately for patients randomized before the amendment and patients after the amendment. Within each subset of patients (randomized before/after amendment), the randomization stratification factors as entered in the IVRS/IWRS for each randomized patient will be derived from the leading digit of the corresponding randomization number.

For patients randomized before the amendment the correspondence is as follows:

Leading digit	Histology	Initial response to 1st line treatment	Previous treatment with ICI
1	Squamous	Objective response	Previous ICI
2	Squamous	Objective response	No previous ICI
3	Squamous	No objective response	Previous ICI
4	Squamous	No objective response	No previous ICI
5	Non-Squamous	Objective response	Previous ICI
6	Non-Squamous	Objective response	No previous ICI
7	Non-Squamous	No objective response	Previous ICI
8	Non-Squamous	No objective response	No previous ICI

For patients randomized before the amendment the correspondence is as follows:

Leading digit	randhist_1r	randrsp_1r	randimm2_1r
1	Squamous	Objective response	1st line ICI
2	Squamous	Objective response	2nd line ICI
3	Squamous	No objective response	1st line ICI
4	Squamous	No objective response	2nd line ICI
5	Non-Squamous	Objective response	1st line ICI
6	Non-Squamous	Objective response	2nd line ICI
7	Non-Squamous	No objective response	1st line ICI
8	Non-Squamous	No objective response	2nd line ICI

A patient listing will display the patients with stratification factors discrepancies between the values entered in the IVRS/IWRS and values corrected by the site after time of randomization.

The second table in this section applies to patients randomized after the amendment (not before). In this table 'randhist_1r', randrsp_1r, and randimm2_1r respectively correspond to Histology, Initial response to 1st line treatment, and line rank of ICI therapy.

8.8. Medical History

The number and percent of subjects with individual concurrent illness and medical histories will be summarized by treatment arm. Individual subject listings will also be provided for concurrent illness and medical history.

Medical history will be coded using MedDRA Version 20.1.

The number and percentage of subjects with any medical history will be summarized overall and for each system organ class (SOC) and preferred term (PT). Percentages will be calculated based on number of subjects in the ITT sets.

Subject medical history data including specific details will be presented in a listing.

8.9. Concomitant Medications

A concomitant medication is defined as any medication taken on or after the day of first exposure to study drug up until 28 days after last dose of study drug.

Prior medications are defined as any medication that is taken prior to the day of first exposure to any study drug.

The number and percentages of all concomitant medications will be summarized on the ITT sets and by treatment arm, Anatomical Therapeutic Chemical (ATC) level 1 (sorted according to the corresponding first letter of the 5-digit ATC code), and Preferred Term. The number and percentages of subjects with at least one concomitant medication will be summarized by treatment arm.

Medications received for the disease after stopping study drug will summarized on the ITT sets by treatment arm using the same presentation.

The total number of concomitant medications present at baseline (day of first exposure to study) and the corresponding number and percentages of subjects will also be summarized by treatment arm.

Since treatment received for the disease after stopping drug is not limited to medications (it includes radiotherapy), the text ‘Medications received for the disease after stopping study drug will summarized on the ITT sets by treatment arm using the same presentation.’ should be changed to ‘Treatments received for the disease after stopping study drug will be summarized on the Safety set by treatment arm using the presentation (using ATC level 1 and PT for medications and SOC and PT for non-drug treatments)’.

8.10. Treatment Compliance

The number of days between first and last exposure to treatment up to the analysis cut-off date, the total administered dose up to the analysis cut-off date and the corresponding dose intensity – calculated as total administered dose as a proportion of total planned dose – will be summarized by treatment arm for the Safety sets. The total dose will be reported separately for each of the two treatments in arm B (docetaxel, pemetrexed). The descriptive statistics of the number of cycles where the study drug was administered received up to the analysis cut-off date will be summarized by treatment arm.

8.11. Time Intervals

8.11.1. Time Intervals between Cycles

Descriptive statistics by treatment arm of the number of weeks between Cycle 1 Day 1 and the randomization and then between each of the other Cycle Day 1 visits and the preceding Cycle Day 1 visit will be presented by treatment arm on the ITT sets. Cycles performed after the analysis cut-off date will be ignored. Patients with a cycle delayed more than 7 days after the theoretical date defined for the treatment arm will be listed.

Cycles taken into account will be those with a study drug administration recorded.

8.11.2. Survival Data Collected on the Censored Patients and Duration of Follow-up

Descriptive statistics by treatment arm of the weeks since randomization and days between last survival information and date of database extraction will be prepared on the censored patients (excluding withdrawals of consent) for the ITT sets. For patients with survival information collected after the analysis cut-off date witnessing that they were alive at the time of the cut-off date, weeks since randomization are counted to the cut-off date and days between last survival information and date of database extraction is counted as 0 days.

Duration of Follow-up will be presented using reverse Kaplan Meier methodology. This method is calculated in the same way as the Kaplan-Meier estimate of the survival function, but with the meaning of the status indicator reversed so that our event of interest becomes the censor, and the censor becomes event.

8.11.3. Time Intervals between RECIST Assessments

Descriptive statistics by treatment arm of the number of weeks between the first post-baseline RECIST assessment and the randomization and then between each of the other post-baseline RECIST assessments and the preceding assessment will be presented on the ITT set. Assessments performed after the analysis cut-off date will be ignored.

8.11.4. Time Intervals between Quality of Life Assessments

Descriptive statistics by treatment arm of the number of weeks between the first post-baseline QOL questionnaire and the randomization and then between each of the other post-baseline QOL questionnaires and the preceding QOL questionnaire will be presented on the ITT set. Questionnaires completed after the analysis cut-off date will be ignored.

9. EFFICACY ANALYSES

Unless otherwise noted, all efficacy analyses will be completed using the ITT sets and in PP sets only for primary endpoint analyses. All efficacy analyses will be completed using the treatment arms as assigned by the randomization.

9.1. Overall Survival (OS)

For the ITT sets, OS is defined as time from randomization to death, expressed in months, will be censored at the analysis cut-off date (15JAN2021). For instance, a patient with last news = alive after the cut-off date will be censored at the cut-off date, a patient with death after the cut-off date be counted as alive (but censored) at the cut-off date. [Note to programmers: the OS time and censoring flag derived without using the analysis cut-off date will also be derived and kept as supplementary records in the ADTTE ADaM dataset]

OS will be summarized using the Kaplan-Meier method and displayed graphically. The median and Q1, Q3 event times for each treatment arm and the corresponding 2-sided 95% confidence interval will be provided.

A 2-sided log-rank test stratified for the randomization stratification factors will be used to compare OS between the two treatment arms. The Cox regression model, stratified for the same stratification factors, will be fitted, and the estimated hazard ratio and 2-sided 95% confidence interval will be provided. The hazard ratio, as estimated in this Cox regression model, will be provided with its 2-sided 95% CI on the Kaplan-Meier plot.

A two-sided unstratified log-rank test and an unstratified Cox regression model will also be used as additional secondary analyses for OS.

The six-month survival rate will be estimated using the Kaplan-Meier method and a 2-sided 95% confidence interval for the log [-log(6-months survival probability)] will be calculated using a normal approximation and then back transformed to give a confidence interval for the 6-months survival rate itself. The 12-month survival rate will be estimated similarly.

These analyses will be implemented by the following example SAS® code:

Log-rank test stratified:

```
proc lifetest;
 strata factor1 factor2 factor3 / group= trt;
 time time*cnsr(1);
run;
```

Log-rank test unstratified:

```
proc lifetest;
```

```
 strata trt;
 time time*cnsr(1);
run;
```

Cox model stratified:

```
proc phreg;
 class trt(ref="1");*1 = control group;
 model time*cnsr(1) = trt / ties=exact rl;
 strata factor1 factor2 factor3;
run;
```

Cox model unstratified:

```
proc phreg;
 class trt(ref="1");*1 = control group;
 model time*cnsr(1) = trt / ties=exact rl;
run;
```

NB: ties=exact may be replaced by ties=efron in case of calculation issues.

All survival analyses described above will be repeated on the PP set (OS).

Subgroups analyses of OS will be performed on the ITT sets to determine whether treatment effect is consistent among subgroups. More specifically, results of the subgroup analyses will be displayed in a forest plot including:

- For each subgroup level, the number of events, the sample size, the median survival time (months) and the estimate of the treatment effect (HR) with its 95% confidence interval as obtained for a separate unstratified and stratified Cox model within each subgroup level.
- The logarithmic scale used to graphically represent the treatment effect estimates and their 95% confidence interval will be restricted to 0.20 - 5.00 (a few confidence intervals may overflow these limits).
- A p-value for the interaction test obtained from the unstratified Cox model including terms for “subgroup”, “treatment”, “treatment by subgroup” interaction.
- A vertical reference line displayed at the level of the overall treatment effect.
- The subgroup analysis will not be performed in subgroups with less than 10 patients (for variables with more than 2 levels, the level with the sample size below 10 may be collapsed with one of the adjacent levels, if appropriate).

Additionally, multivariate Cox regression models, stratified for the same stratification factors as used for the randomization, will be used on the ITT sets to explore the potential joint influences of the other factors (those defining subgroups in the text above) on the primary OS endpoint. If deemed appropriate backward selection of variables will be used.

The results of the subgroup analyses will be detailed in a summary table and summarized in a forest plot. The forest plot will present for each subgroup variable the p-value of the interaction test and for within each subgroup level the sample size by treatment arm, the median survival

time (months) by treatment arm with is 95% confidence interval, the HR with is 95% confidence interval. Same information will be provided in the Figures presenting the Kaplan-Meier curves.

9.2. Progression-free Survival (PFS)

Progression-free survival (PFS) will be defined as time (in months) from randomization to the first documented disease progression (as defined using the RECIST 1.1 criteria and FDA Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics, Guidance for Industry, April 2015) or death whichever occurs first.

The PFS derivation rule is as follows (ignoring information collected after the analysis cut-off date):

- Check whether a candidate time point (for event or censoring) is more than 14 weeks (2 times 6+1 weeks) after the last available tumor assessment before this time point; apply Table C1 of FDA guidance:
  - No baseline tumor assessment: censored at randomization (0 days after randomization)
  - Progression documented in tumor assessment or in Survival Follow-up form or in End of Treatment form:
    - If previous tumor assessment without progression ≤ 14 weeks before: PD
    - If previous tumor assessment without progression > 14 weeks before: censored at previous tumor assessment without progression
  - Death without post-baseline tumor assessment
    - ≤ 14 weeks after randomization: event at date of death
    - > 14 weeks after randomization: censored at randomization
  - Death with at least 1 post-baseline tumor assessment
    - if previous tumor assessment without progression ≤ 14 weeks before: event at date of death
    - if previous tumor assessment without progression > 14 weeks before: censored at previous tumor assessment without progression ("more than one missed visit")
  - new anticancer treatment started without documented progression recorded: censored at previous tumor assessment without progression
  - no documented progression, no death, no new anticancer treatment started: censored at previous tumor assessment without progression

Since the protocol plans for a tumor assessment every 6 weeks (± 1 week) until documented RECIST 1.1 progression, the 14 weeks (98 days) mentioned above correspond to 'more than one missed visit' used in Table C1 in Appendix C of the FDA guidance 'FDA Clinical Trial Endpoints for the Approval of Non-Small Cell Lung Cancer Drugs and Biologics, Guidance for Industry, April 2015'. Other censoring strategies described in the guidance will be explored in the SAP appendix.

As described in Appendix 3 of the FDA guidance 'Clinical Endpoints for the Approval of Cancer Drugs and Biologics, May 2007', clinical progression is not considered a progression endpoint.

In case of conflicting information between RECIST records indicating no progression of disease and EOT form or survival form reporting a 'documented clinical progression', the worst value (progression) will be retained in the PFS derivation.

PFS will be summarized using the Kaplan-Meier method and displayed graphically. The median event time for each treatment arm and the corresponding 2-sided 95% confidence interval will be provided.

A 2-sided log-rank test stratified for the randomization stratification factors will be used to compare PFS between the two treatment arms. The Cox regression model, stratified for the same stratification factors, will be fitted, and the estimated hazard ratio and 2-sided 95% confidence interval will be provided. The hazard ratio, as estimated in this Cox regression model, will be provided with its 2-sided 95% confidence interval on the Kaplan-Meier plot.

The six-month PFS rate will be estimated using the Kaplan-Meier method and a 2-sided 95% confidence interval for the log [-log(6-months survival probability)] will be calculated using a normal approximation and then back transformed to give a confidence interval for the 6-months PFS rate itself. The 12-month PFS rate will be estimated similarly.

As sensitivity analysis, Clinical Progression-free survival (PFS) analysis will be performed with following definition of progression event: earliest date either of a RECIST tumor assessment with PD status (without consideration for the number of weeks since previous tumor assessment without progression) or of the mention of a documented progression in the Survival Follow-up form or in the End of Treatment form) or death whichever occurs first.

9.3. Overall response Rate (ORR), Disease Control rate (DCR) and Duration of response (DR)

Assessment of response will be made using RECIST version 1.1 (the protocol does not require confirmation at a subsequent time point for CR and PR) and ignoring assessments performed after the analysis cut-off date.

Patients will be considered as without measurable lesions at baseline as soon as (at least) one post-baseline RECIST e-CRF form indicates 'No Target Lesions at Baseline'. The remaining patients will be considered as with measurable lesions at baseline.

Based on the randomized patients, the last available response in patients still alive at the analysis cut-off date will be displayed in the summary of overall survival time (see DMC Table 2.1.1).

Based on the randomized patients with measurable lesions at baseline, the best response (complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD), Not evaluable (NE) or no post-baseline assessment) and the ORR (counting as objective response the patients with CR and PR) will be determined for each patient and will be summarized by treatment arm (see DMC Table 2.2). The corresponding exact 2-sided 95% confidence interval of the ORR will be calculated for each treatment arm. The Mantel-Haenszel test will be used to compare ORR between the treatment arms, using the

randomization strata. The corresponding odds-ratio will be provided with its 95% confidence interval. In addition, the difference in ORR between the two treatments will be provided and its 95% confidence interval will be calculated based on the normal approximation.

DCR at 6 months will be calculated in each treatment arm as the number of patients remaining within CR, PR or SD responses within the considered time period divided by the total number of randomized patients with measurable lesions at baseline. For patients alive at 6 months the last RECIST assessment available up to 183 days (included) after randomization will be retained; a progression declared as 'documented progression' by the site will also be retained and counted as PD. Patients with death within 6 months from randomization will be counted as PD. Patients with consent withdrawal before 6 months from randomization will be excluded from this analysis. DCR will then be analyzed as described for ORR. DCR at 12 months will be calculated and analyzed as described for DCR at 6 months.

Duration of DCR will be summarized for the subgroup of patients who have achieved CR, PR, or SD responses. The duration of DCR will be defined as the time (in months) from the date of the first initial occurrence of a CR or PR or SD to the PFS date (corresponding to a PFS event or to a censoring). Duration of DCR will be analyzed as described for the PFS. The rate à 6 months and 12 months will be also estimated as sensitivity analysis.

DR will be summarized for the subgroup of patients who have achieved objective response. The DR will be defined as the time (in months) from the date of the first initial occurrence of a CR or PR to the PFS date (corresponding to a PFS event or to a censoring). DR will be analyzed as described for the PFS.

Depending on the availability of the data waterfall plot (to described quantitative data on target lesion) or spider plots (to describe synthetically each patient) will be proposed.

9.4. Time to next lung cancer therapy

The date of initiation of a lung cancer therapy will be taken as the date of contact of the first survival follow-up form mentioning that the patient has received a treatment for the disease after stopping the study drug. This date will be censored at the analysis cut-off date for patients having received a lung cancer therapy. This date will be censored at the cut-off date or at the date of death, whichever comes first, for patients having not initiated a lung cancer therapy.

The time (in months) from the randomization to this date will be analyzed as described for the PFS. As a sensitivity analysis, same analysis will be performed with the time (in months) from the earliest date either of a RECIST tumor assessment with PD status or of the mention of a documented progression in the Survival Follow-up form or in the End of Treatment form to this date.

The text ‘This date will be censored at the cut-off date or at the date of death, whichever comes first, for patients having not initiated a lung cancer therapy.’ should be changed to ‘This date will be censored at the cut-off date or at the date of death (or at the date of last survival information for patients who did not experience death), whichever comes first, for patients having not initiated a lung cancer therapy’.

Sensitivity analysis using the date of progression (in months) to this date will be analyzed as described for the PFS. The date of progression will be defined as the earliest date either of a RECIST tumor assessment with PD status (without consideration for the number of weeks since previous tumor assessment without progression) or of the mention of a documented progression in the Survival Follow-up form or in the End of Treatment form.

9.5. Post-progression Survival

The date of progression will be defined as the earliest date either of a RECIST tumor assessment with PD status (without consideration for the number of weeks since previous tumor assessment without progression) or of the mention of a documented progression in the Survival Follow-up form or in the End of Treatment form.

Post-progression survival of patients having experienced a progression will be summarized by treatment arm using the Kaplan-Meier method and displayed graphically. The median event time for each treatment arm and the corresponding 2-sided 95% confidence interval will be provided.

An additional analysis will be performed to explore impact on OS to continue study treatment beyond progression. The number of patients continuing study treatment 28 days or more after progression defined as the earliest date either of a RECIST tumor assessment with PD status or of the mention of a documented progression in the Survival Follow-up form or in the End of Treatment will be presented in each group. Post-progression survival of patients having continuing study treatment 28 days or more after progression versus patients not having continued study treatment 28 days or more after progression (other treatment after progression) will be summarized by treatment arm (4 curves) using the Kaplan-Meier method and displayed graphically. The median event time for each treatment arm and each subgroup and the corresponding 2-sided 95% confidence interval will be provided. Overall survival of patients having continued study treatment 28 days or more after progression versus patients not having continued study treatment 28 days or more after progression (other treatment after progression) will be summarized by treatment arm (4 curves) using the Kaplan-Meier method and displayed graphically. The median event time for each treatment arm and each subgroup and the corresponding 2-sided 95% confidence interval will be provided.

A listing of patients continuing treatment after PFS will be provided with following information: date of RECIST progression and PFS in months, date of clinical progression, treatment start date, treatment end date, treatment duration, reason for treatment discontinuation, duration between date of RECIST progression and date of treatment discontinuation, date of death and OS in months. A summary table of number of patients continuing treatment after PFS in each arm (>=28 days) and overall will be provided. A table of study treatment exposure per cycle beyond 28 days post PFS will be provided in each arm and overall.

.

9.6. Quality of life analyses

The scores will be derived from the EORTC QLQ-C30 and EORTC QLQ-LC13 questionnaires items according to the respective scoring manuals, ignoring assessments performed after the analysis cut-off date or after treatment discontinuation.

A summary of each total score and sub-score for each dimension, for the EORTC QLQ-C30 will be described by cycle and treatment arm.

- **Global health status / QoL**
- **Functional scales**
  - Physical functioning
  - Role functioning
  - Emotional functioning
  - Cognitive functioning
  - Social functioning
- **Symptom scales / items**
  - Fatigue
  - Nausea and vomiting
  - Pain
  - Dyspnea
  - Insomnia
  - Appetite loss
  - Constipation
  - Diarrhea
  - Financial difficulties

A summary of each total score and sub-score for each dimension, for the EORTC QLQ-LC13 will be described by cycle and treatment arm.

- **Symptom scales / items**
  - Dyspnea
  - Coughing
  - Hemoptysis
  - Sore mouth
  - Dysphagia
  - Peripheral neuropathy
  - Alopecia
  - Pain in chest
  - Pain in arm or shoulder
  - Pain in other parts

After scoring process, domains score will be from 0 to 100. The higher the value, the higher the quality of life. After scoring process, symptoms score will be from 0 to 100. The lower the value, the poorest the quality of life.

Questionnaires collected as "end of treatment" visit will be mapped to next protocol visit if they are not duplicate compared to the last non-missing visit.

Absolute changes from baseline will be summarized as well. Plots representing means and 95%CI for raw data and absolute changes from baseline will be associated to graphically visualized this data.

Missing data fractions for quality of life will be reported under different scenarios:

- rate of patients completing baseline assessments and the assessments at designated time points over the total number of patients eligible and entered the trial.
- rate of patients completing assessments at designated time points while on study treatment over the total number completing assessment at baseline.
- rate of patients completing assessments at designated time points over the number of patients still on study, who were expected to complete questionnaires at each of those time points (excluding those who progressed or were dead at that time point).

For each domain or symptom, mean changes within arms from baseline to the last cycle with more than 25% of the total number of included patients (Cycle 8) will be reported to provide a complete picture of QoL behavior.

Changes in scores from baseline until treatment discontinuation will be assessed using mixed-effects model for repeated measures (MMRM) analysis. Missing data are not imputed. The p value will be determined using MMRM analysis, with patient, treatment, visit, and treatment by visit interaction as explanatory variables and baseline score as covariate. Patient will be fitted as a random effect, and compound score covariance structure will be used for all models. The Least Squares Mean differences for each symptom will be reported and plotted with corresponding 95% CIs.

Two sensitivity analyses for consistency with those from the main analysis will be conducted only on the quality of life secondary endpoints (QLQ-C30 Global health status, QLQ-C30 Physical and Role functioning scores):

- using only patients with complete data from cycle 1 to the median number of cycles
- imputing the worst value to next protocol visit after the last non-missing visit of patients discontinued treatment for death

Best QoL response from baseline will be derived for each domain or symptom as follows: a change score of at least 10 points from baseline will be defined as clinically relevant. Patients will be considered improved if they reported a score of 10 points or more better than baseline at any time and will be considered worsened if they reported a score 10 or more points worse than baseline (without improvement). Patients whose scores change less than 10 points from baseline will be considered stable. Only patients who had completed the baseline and at least one follow-up questionnaire will be included. An exact linear rank test will be used to test whether the two study arms had the same underlying multinomial distribution of the ordered QoL response

9.7. ECOG

The time to worsening of ECOG will be defined as the time from randomization to the earliest time when the ECOG becomes greater than 1 ignoring all assessments performed after the analysis cut-off date. Patients who die without having recorded an ECOG >1 will be assigned an ECOG = 5 (Dead) at the time of death. Patients without worsening will be censored at the last time when an ECOG value was recorded. The endpoint will be analyzed as described for the PFS.

By protocol, ECOG is not collected anymore when a subject discontinues study treatment while the time between study treatment discontinuation and death can be quite long. The analysis described above assigns an ECOG 5 at the time of death. This imputation may seem exaggerated as, for most of the imputed patients, a condition corresponding to a (unrecorded) Grade 2 or worse is likely to have occurred markedly earlier than the date of death imputed as Grade 5. In that situation the imputation artificially extends the TTD-free period.

For this reason, an additional analysis without imputation for death will be performed. This analysis will exclude patients without any post-baseline ECOG assessment and censor patients without observed ECOG worsening at their last ECOG assessment.

10. SAFETY ANALYSES

All Safety analyses will be conducted using the Safety sets. An overall treatment summary will be included. All safety analyses will be completed using the actual treatment a subject received.

For laboratory safety parameters, vital signs and ECGs, the baseline will be the last available value up to the date of first study drug administration included.

10.1. Adverse Events

All AEs and SAEs will be coded using the MedDRA dictionary for System Organ Class (SOC) and Preferred Term (PT), with the creation of an ad hoc SOC 'Injection site reactions' grouping all PTs with High Level Term 'Administration site reactions'.

Level of intensity was graded according to the CTCAE v5.0 grading.

Treatment-emergent adverse events (TEAEs) are defined as any adverse event or worsening of an existing condition after the date of the first administration of the study treatment (included) and up to 28 days after the date of the last administration of the study treatment (included) or up to the analysis cut-off date (whichever comes first).

Immune-related TEAEs (irAEs) are defined as any TEAE that has been specifically flagged as immune-related on the e-CRF.

Serious adverse events were recorded from the date of informed consent, throughout the clinical trial, and for up to 28 days after the final administration of study drug.

The number of emergent events and the number and percentage of subjects with TEAEs will be summarized by SOC and PT by treatment arm for different categories of TEAEs. At each level of tabulation (e.g., at the PT level) patients will be counted only once if they had more than one such event reported during the AE collection period.

The following summary tables will be presented for TEAE data:

- Treatment-Emergent Adverse Events Overview
- Treatment-Emergent Adverse Events, by Severity and Grade
- Treatment-Emergent Adverse Events, by SOC and PT
- Treatment-Emergent Severe Adverse Events, by SOC and PT
- Treatment-Emergent Serious Adverse Events, by SOC and PT
- Treatment-Emergent Adverse Events with Fatal Outcome, by SOC and PT
- Treatment-Emergent Adverse Events leading to Withdrawal from the Study, by SOC and PT
- Treatment-Emergent Drug Related Adverse Events, by SOC and PT
- Treatment-Emergent Severe Drug Related Adverse Events, by SOC and PT
- Treatment-Emergent Serious Drug Related Adverse Events, by SOC and PT
- Treatment-Emergent Drug Related Adverse Events with Fatal Outcome, by SOC and PT
- Treatment-Emergent Drug Related Adverse Events leading to Withdrawal from the Study, by SOC and PT

The Treatment-Emergent Adverse Events Overview table will display by treatment arm the number of emergent events and the number and percentage of subjects with TEAEs for each of the following TEAE categories:

- Treatment-Emergent Adverse Events
- Treatment-Emergent Drug Related Adverse Events
- Treatment-Emergent Severe Adverse Events
- Treatment-Emergent Severe Drug Related Adverse Events
- Treatment-Emergent Serious Adverse Events
- Treatment-Emergent Serious Drug Related Adverse Events
- Treatment-Emergent Adverse Events with Fatal Outcome
- Treatment-Emergent Drug Related Adverse Events with Fatal Outcome
- Treatment-Emergent Adverse Events leading to Withdrawal from the Study
- Treatment-Emergent Drug Related Adverse Events leading to Withdrawal from the Study

For each row of the overview table, the chi-square test comparing the two treatment arms on the proportion of patients with the considered TEAE category will be used as a flagging device.

The same presentation will be repeated for the irAEs.

All the tables listed above will be repeated on the following subgroups:

- Age (< 65 years vs. ≥ 65 years)
- Sex (Male, Female
- Age and sex (Male < 65 years, Male. ≥ 65 years, Female < 65 years, Female ≥ 65 years)
- Region (US, non-US).

Race is not retained for defining subgroups since the number of patients of non-white patients is very limited (1 Asian and 1 Black or African American within the treated patients).

'Severe' will be defined as Grade 3 or above. A missing severity will be considered Grade 3.

'Drug Related' will be defined as 'Definitely Related', 'Probably related' or 'Possibly Related', excluding 'Unlikely Related' and 'Not Related'. A missing relationship will be considered as 'Definitely Related'.

TEAEs 'leading to Withdrawal from the Study' are defined as TEAEs where the AE data 'Withdrawn from Study?' is ticked Yes.

The list of summary tables for TEAE data will include a table for 'Treatment-Emergent Drug Related Adverse Events by Severity and Grade'.

The irAEs will be described by 2 tables: 'Treatment-Emergent Adverse Events Overview' and 'Treatment-Emergent Adverse Events by SOC and PT'. No subgroup tables will be prepared for the irAEs.

Since the number of patients in US is limited (7 within the treated ICI patients randomized before 26 February 2019, 16 within the ICI patients randomized and treated as of 26 February 2020), the subgroup variable Region (US, non US) will not be retained for the CSR.

The ad hoc SOC 'Administration site reactions' will be created by grouping all PTs with High Level Group 'Administration site reactions'.

AEs not coded for patients 1514021 and 2202009 (respectively: 'Right arm inflammation' and 'Increase of cholestasis') will be displayed in the summary tables by their reported term (in place of PT) within the ad hoc SOC 'Not coded'.

10.2. Deaths, Serious Adverse Events and Other Significant Adverse Events

10.2.1. Deaths

All deaths occurring up to the analysis cut-off date, regardless of causality, will be summarized by cause of death.

10.2.2. Serious Adverse Events

A listing of Serious Adverse Events (SAEs) occurring up to the analysis cut-off date will be provided.

SAEs occurring before study drug initiation will be listed separately.

10.2.3. Adverse Events Leading to Discontinuation of Study Drug or Withdrawal from Study

A listing of all AEs occurring up to the analysis cut-off date and leading to discontinuation of study drug or withdrawal from study will be presented.

10.3. Clinical Laboratory Evaluations

Clinical laboratory values will be analyzed after conversion into standard international units. International units will be used in all listings and tables.

The results will be graded according to the CTCAE v5.0, when applicable. The number and percentage of patients with laboratory abnormalities (all grades and by grade) using the worst grade between the date of the first administration of the study treatment (excluded) and up to 28 days after the date of the last administration of the study treatment (included) or up to the analysis cut-off date (whichever comes first) will be displayed by treatment arm in a shift table stratifying the patients according to the grade observed at baseline.

For safety laboratory parameters where the CTCAE v5.0 grading scale is not applicable, the number and percentage of patients with a value below the normal range and the number and percentage of patients with a value above the normal range will be displayed by treatment arm in a shift table stratifying the patients according to their situation at baseline (below the lower limit of normal range, within normal range, above the upper limit of normal range). A same patient may be present in both counts.

For each parameter, and among each grade observed at baseline the percentages will be based on the number of subjects having at least one post-baseline value no later than the analysis cut-off date.

10.4. Vital Signs

The number and percentage of patients with marked abnormalities between the date of the first administration of the study treatment (excluded) and up to 28 days after the date of the last administration of the study treatment (included) or up to the analysis cut-off date (whichever comes first) will be displayed for each type of abnormality by treatment arm in a shift table stratifying the patients according to their situation at baseline (with/without the considered marked abnormality at baseline). Marked abnormalities are classified as high or low based on values occurring above the higher limit or below the lower limit, respectively. It is possible

that, for a given parameter, the same subject is counted as a high marked abnormality for an observed value and as a low marked abnormality for a different observed value.

Variable	Abnormality
Weight (kg)	decrease from baseline > 10% increase from baseline > 10%
Body temperature (°C)	< 36 > 39
Systolic blood pressure (mm Hg)	< 90 > 140
Diastolic blood pressure (mm Hg)	< 60 > 90
Heart Rate (bpm)	< 50 > 100
Respiratory rate (breaths/minute)	< 10 > 20
Oxygen saturation (%)	<95%

For each parameter, the percentages will be based on the number of subjects having at least one post-baseline value no later than the analysis cut-off date.

---- Addition or change after 04 June 2020 --------

For the weight, the percentages will be based on the number of subjects having a baseline value and at least one post-baseline value no later than the analysis cut-off date.

[Reference: e-mail of 18 August 2020].

10.5. 12-Lead ECG

The number and percentage of patients with clinically significant abnormal ECG between the date of the first administration of the study treatment (excluded) and up to 28 days after the date of the last administration of the study treatment (included) or up to the analysis cut-off date (whichever comes first) will be displayed by treatment arm in a shift table stratifying the patients according to their situation at baseline (with/without clinically significant abnormal ECG at baseline). The percentages will be based on the number of subjects having at least one post-baseline ECG no later than the analysis cut-off date.

11. ADDITIONAL SAFETY ANALYSES

The population of all ICI treated patients randomized until the analysis cut-off date will be described by treatment arm for the following items:

- Demographics
- Baseline characteristics

- Previous ICI treatment
- History of immune-related AEs
- Medical history
- Concomitant medications
- Medications received for the disease after stopping study drug

All safety analyses described in the previous section will be repeated on this population.

12. COVID-19 IMPACT

According to the following guidance's

- European Medicines Agency (2020a), "Guidance on the Management of Clinical Trials During the COVID-19 (Coronavirus) Pandemic," Version 3.

- European Medicines Agency (2020b), "PoInts to Consider on Implications of Coronavirus disease (COVID-19) on Methodological Aspects of Ongoing Clinical Trials."

- "FDA Guidance on Conduct of Clinical Trials of Medical Products During COVID-19 Public Health Emergency: Guidance for Industry, Investigators, and Institutional Review Boards."

- Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency: Guidance for Industry

and considering that 110 patients out of the 219 are potentially impacted by the pandemic on the 01JAN2020:

- 60 patients untreated in follow-up for survival

- 24 patients treated

- 26 patients randomized between the 01JAN2020 and 30APR2020

During the data review committee, protocol deviations related to COVID-19 (visits or infusions cancelled or delay, COVID-19 events, death related to COVID, follow-up visits delayed, …) will be reviewed. If some major deviations are decided related to COVID-19 patients will be excluded from PP population.

Number of deviations will be presented overall and per category of deviation (DVCAT in the ADDV database) and protocol deviation coded term (DV.DVDCOD) using 2 different periods of randomization with or without COVID-19 (Period without COVID-19: First inclusion to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion) using following table

	Arm A (N=XX)		Arm B (N=XX)		Total (N=XX)	
	Number of deviations	Number of % patients	Number of deviations	Number of % patients	Number of deviations	Number of % patients
At least one deviation						
DVCAT: EXCLUSION DVCAT:CRITERION						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:NOT MET						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:INCLUSION DVCAT:CRITERION NOT MET						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:NFORMED CONSENT						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:PROCEDURES/ TESTS						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:RANDOMIZATION						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:SERIOUS ADVERSE EVENTS						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:STUDY DRUG						
protocol deviation coded term1						
protocol deviation coded term2						
…						
DVCAT:VISIT SCHEDULE						
protocol deviation coded term1						
protocol deviation coded term2						
…						

Same table will be presented for patients excluded from PP.

Analysis described in 8.11.1 Time Intervals between Cycles, 8.11.3 Time Intervals between RECIST Assessments, and 8.11.4Time Intervals between Quality of Life Assessments will be

repeated for the 2 different periods of randomization with or without COVID-19 (Period without COVID-19: First inclusion to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion). This analysis will be performed at a patient level and at visit level and will be added in the ancillary SAP.

Analysis described in 8.11.1 Time Intervals between Cycles, 8.11.3 Time Intervals between RECIST Assessments, and 8.11.4Time Intervals between Quality of Life Assessments will be repeated for the 2 different periods of randomization with or without COVID-19 (Period without COVID-19: First inclusion to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion). This analysis will be performed at a patient level and at visit level and will be added in the ancillary SAP.

The number of deaths any cause will be summarized by treatment arm for the 2 different periods of randomization with or without COVID-19 (Period without COVID-19: First inclusion to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion). This analysis will be performed at a patient level and using the patient year unit during the 2 periods.

These analyses will be added in the ancillary SAP.

The number of emergent events and the number and percentage of subjects with TEAEs will be summarized by SOC and PT by treatment arm for different categories of TEAEs for the 2 different periods of randomization with or without COVID-19 (Period without COVID-19: First inclusion to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion). This analysis will be performed at event level and at patient level. This analysis will be repeated using the patient year unit during the 2 periods for the denominator.

These analyses will be added in the ancillary SAP.

- Death any cause
- Treatment-Emergent Adverse Events Overview
- Treatment-Emergent Adverse Events, by SOC and PT
- Treatment-Emergent Drug Related Adverse Events, by SOC and PT
- Treatment-Emergent Severe Drug Related Adverse Events, by SOC and PT

Same analyses will be performed using the same 2 periods (without COVID-19, with COVID-19) but considering start date of AE for the tables. In this analysis patients with AEs in the 2 periods will be counted in the 2 periods. For patient without any AE during study, date of randomization will be considered for the denominator of the percentages.

- Period1: AE which start between first inclusion and 31DEC2019

- Period2: AE which start between 01JAN2020 and last inclusion

12.1. Disposition

Status of the population before and after pandemic will be summarized using following table:

N (%)	Arm A (n=XX)	Arm B (n=XX)	Total (n=219)
Randomized until 31DEC19	Na	Nb	Na+Nb
Alive on 01JAN20 or randomized after the 31DEC19			
In Follow-up after the 01JAN20			
Treated after the 01JAN20			
Included after the 01JAN20			

12.2. Sensitivity analysis on censor date

Same analysis for OS as described in 9.1 will be repeated with a censor date at 3 following cut-offs: 31DEC2019, 31JAN2020 and 26FEB2020 (cut off for Step 1 analysis).

12.3. Time-dependent covariate analyses

Time-dependent covariates are those that may change in value over the course of observation. For a model with one time-invariant covariate and one time-dependent covariate, we have
$\log h(t)=\alpha(t)+\beta 1X1 + \beta 2X2(t)$
For example, if we want a model in which the hazard of OS depends on COVID-19 status, we could specify COVID as whether the person is currently COVID exposed in week t.

Cox models taking into account COVID time dependent will be presented using

- 3 different periods of randomization with or without COVID-19 (Period 1: Step1: First inclusion to 26FEB2019, then post step 1: 26FEB2019 to 31DEC2019, then COVID-19 period: 01JAN2020 to last inclusion).

- 2 different periods of randomization with or without COVID-19 (Step1: First inclusion to 26FEB2019, then post step 1: 26FEB2019 to last inclusion).

- 2 different periods of randomization with or without COVID-19 (Step1: First inclusion 31DEC2019, then: 01JAN2020 to last inclusion).

Using AIC criterion, the best model will be chosen.

- One simple model with overall population to test COVID impact

```
proc phreg data = XX;
class COVIDIMPACT /ref = last;
model (before, after) * censored(1) = COVIDIMPACT /ties = EFRON
rl;run;
```

- One simple model for each treatment arm
- One model with overall population to test COVID impact with treatment arm in the model with or without interaction

```
proc phreg data = XX;
class COVIDIMPACT ARMGR/ref = last;
model (before, after) * censored(1) = COVIDIMPACT ARMGR /ties =
EFRON rl;run;
```