STATISTICAL ANALYSIS PLAN

Protocol title:	**Open-label, Phase 2 study, evaluating the efficacy and safety of tusamitamab ravtansine in non-squamous non-small-cell lung cancer (NSQ NSCLC) participants with negative or moderate CEACAM5 expression tumors and high circulating CEA**
Protocol number:	**ACT17241**
Compound number (INN/Trademark):	**SAR408701** **tusamitamab ravtansine/Not applicable**
Study phase:	**Phase 2**
Short Title:	**Tusamitamab ravtansine in NSQ NSCLC participants with negative or moderate CEACAM5 expression tumors and high circulating CEA** **CARMEN-LC06**
Statistician:	[redacted]
Statistical project leader:	[redacted]
Date of issue:	**12-Jan-2024**
Regulatory agency identifier number(s):	
IND:	144484
EudraCT:	2021-004423-32
NCT:	Not applicable
WHO:	U1111-1264-2828
EUDAMED:	Not applicable
Other:	Not applicable

Date: 12-Jan-2024 Total number of pages: 46

TABLE OF CONTENTS

LIST OF TABLES

VERSION HISTORY

This statistical analysis plan (SAP) for study ACT17241 is based on the protocol dated 27-Sep-2021. There are no major changes to the statistical analysis features in this SAP.

The first participant was enrolled on 03-Aug-2022. This SAP is approved before the interim analysis is conducted.

Major changes in statistical analysis plan

SAP Version	Approval Date	Changes	Rationale
1	12-Jan-2024	Not Applicable	Original version

1 INTRODUCTION

There are no major changes to the analyses described in the protocol.

1.1 STUDY DESIGN

This is a single group, Phase 2, open-label, multi-center study assessing efficacy (anti-tumor activity), safety, and PK of tusamitamab ravtansine single agent in NSQ NSCLC participants with negative (ie, intensity 1+ whatever the percentage of stained tumor cells or <1% of tumor cells) or moderate (ie, intensity ≥2+ in ≥1% and <50% of tumor cells) CEACAM5 expression tumors and high circulating CEA (≥100 ng/mL) at baseline.

During the pre-screening phase, participants who were pre-screened failed in the EFC15858 Phase 3 trial with available CEACAM5 status (central assessment by immunohistochemistry, IHC) as negative or moderate expression on archival tumor tissue can be pre-screened for ACT17241 study. Participants with high circulating CEA (≥100 ng/mL) will be screened and will go through protocol screening procedures.

This study is designed to obtain preliminary efficacy, safety and PK data on tusamitamab ravtansine administered as monotherapy at 100 mg/m² Q2W to participants with NSQ NSCLC tumors with CEACAM5 moderate-negative tumor expression and high circulating CEA levels.

Approximately 285 participants will be pre-screened to achieve approximately 38 treated participants, based on an estimated pre-screening failure rate of 84% and an estimated study screening failure rate of 15%.

The cycle duration is 2 weeks. The median expected duration of study per participant is estimated as 40 weeks: up to 4 weeks for screening, a median of 24 weeks for treatment, and a median of 12 weeks for end of treatment assessments and the safety follow-up visit. Enrolled participants may receive study intervention until disease progression, unacceptable AE, initiation of a new anticancer therapy, or the participant's or Investigator's decision to stop the treatment, whichever comes first.

1.2 OBJECTIVES AND ENDPOINTS

Table 1 - Objectives and endpoints

Objectives	Endpoints
Primary	
• To assess the anti-tumor activity of tusamitamab ravtansine when given every 2 weeks (Q2W) in non-squamous non-small-cell-lung-cancer (NSQ NSCLC) participants with negative or moderate CEACAM5 expression tumors and high circulating carcinoembryonic antigen (CEA) levels	• Objective Response Rate (ORR) of tusamitamab ravtansine, defined as the proportion of participants who have a confirmed complete response (CR) or partial response (PR) as best overall response (BOR) per Response Evaluation Criteria In Solid Tumors (RECIST) v1.1

Objectives	Endpoints
Secondary	
• To assess the safety and tolerability of tusamitamab ravtansine	• Incidence of participants with treatment-emergent adverse events (TEAEs), serious adverse events (SAEs) and laboratory abnormalities according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v5.0
• To assess other efficacy parameters of tusamitamab ravtansine	• Progression-free survival (PFS), defined as the time from the date of first tusamitamab ravtansine administration to the date of the first documented disease progression or death due to any cause, whichever comes first. • Disease control rate (DCR), defined as the percentage of participants who have achieved confirmed CR or PR, or stable disease as BOR per RECIST v1.1 • Duration of response (DOR), defined as the time from first documented evidence of CR or PR until progressive disease (PD) determined per RECIST v1.1 or death from any cause, whichever occurs first
• To evaluate the immunogenicity of tusamitamab ravtansine	• Incidence of participants with anti-therapeutic antibodies (ATAs) against tusamitamab ravtansine
• To document the pharmacokinetics (PK) of tusamitamab ravtansine	• Tusamitamab ravtansine concentrations
Tertiary	
• To evaluate patient-reported outcomes (PROs)	• Time to deterioration in the overall severity of disease-related symptoms (cough, pain, dyspnea, fatigue, appetite) as measured by the NSCLC-Symptom Assessment Questionnaire (SAQ) • Patient global impression of severity (PGIS) and patient global impression of change (PGIC) in disease-related symptoms will be measured by generic PGIS lung cancer symptom (PGIS-LCS) and PGIC lung cancer symptom (PGIC-LCS) scales • Change from baseline in overall side effect impact as measured by the Functional Assessments of Cancer Therapy Item GP-5 (FACT-GP5)
• To explore modulations of circulating CEA as a potential pharmacodynamic biomarker of response to tusamitamab ravtansine treatment and to evaluate circulating CEA levels at pre-screening and its correlation with CEACAM5 tumor expression on fresh biopsy (if available)	• Circulating CEA at pre-screening, screening, during the treatment period and during the follow-up period; tumor CEACAM5 expression (on fresh biopsy tumor sample, if available)
• To explore the relationship between CEACAM5 expression in archived and fresh biopsy at screening (if available)	• CEACAM5 expression on archived and fresh biopsy at screening
• To explore the relationship between the tumor mutation profiles detected in the circulating free DNA (cfDNA) at baseline with efficacy outcome	• Mutation analysis for tumor cfDNA at baseline

1.2.1 Estimands

Primary estimand defined for primary efficacy endpoint is summarized in below Table 2. More details are provided in Section 3.

For this estimand, the study intervention of interest will be SAR408701.

Table 2 - Summary of primary estimand for main endpoints

Endpoint Category (estimand)	Estimands			
	Endpoint	**Population**	**Intercurrent event(s) handling strategy**	**Population-level summary (Analysis and missing data handling)**
Primary objective: To assess the anti-tumor activity of tusamitamab ravtansine Q2W in NSCLC participants with negative or moderate CEACAM5 expression tumors and high circulating CEA levels				
Primary endpoint (primary estimand)	Confirmed objective response (confirmed CR or PR as BOR), determined according to RECIST v1.1	All-treated population	Regardless of investigational medicinal product IMP discontinuation (treatment policy strategy) Based on tumor assessments done before initiation of further anticancer therapy ("while not initiating anticancer therapy" strategy)	Objective response rate defined as rate of participants with confirmed objective response and two-sided 95% confidence interval (CI) using the Clopper-Pearson method. In the absence of confirmed objective response, participants will be considered as non-responders, whatever the reason (including participants with non-evaluable BOR).

2 ANALYSIS POPULATIONS

The following populations for analyses are defined:

Table 3 - Population for analyses

Population	Description
Pre-screened	All participants who signed the pre-screening informed consent.
Screened	All participants who signed the screening informed consent for study participation.
Enrolled	All participants from screened population who have been allocated to intervention regardless of whether the intervention was received or not.
All-treated	All enrolled participants exposed to the study treatment, regardless of the amount of treatment administered. All safety analyses will be performed on this population, which is also the primary population for analysis of all efficacy parameters.
Activity	All-treated participants who have measurable disease at study entry and at least 1 postbaseline evaluable tumor assessment. Participants with no postbaseline evaluable tumor assessment but with an early clinical progression or who died from disease progression will also be included in this set. This population is the secondary population for analysis of efficacy parameters.
PK	All participants from the all-treated population with at least 1 postbaseline PK concentration with adequate documentation of dosing and sampling dates and times.
ATA	All participants from the all-treated population with at least 1 postbaseline ATA result (negative, positive, or inconclusive).
Population without trial impact (disruption) due to COVID-19	All exposed participants: • without any critical or major deviation related to COVID-19 • and who did not permanently discontinue treatment due to COVID-19 • and who did not permanently discontinue study due to COVID-19.

Abbreviations: ATA = anti-therapeutic antibody; PK = pharmacokinetic.
Note: In practice, a participant will be included in the enrolled population if the question “Will the subject continue in the treatment phase?” has been answered as “Yes” in the “Completion of screening phase” electronic case report form page.

In practice, a participant will be included in the enrolled population if the question “Will the subject continue in the treatment phase?” has been answered as “Yes” in the “Completion of screening phase” electronic case report form page.

Participants exposed to study intervention before or without being enrolled will not be considered enrolled and will not be included in any analysis population. The safety experience of these participants will be reported separately.

Enrolled participants for whom it is unclear whether they took the study intervention will be considered as exposed and will be included in the “all-treated” population.

For any participant enrolled more than once, only the data associated with the first enrollment will be used in any analysis population. The safety experience associated with any later enrollment will be reported separately.

Participants will be analyzed according to the intervention they actually received.

In practice, participants whose BOR is non-evaluable due to minimum criteria for SD duration not met (ie, overall response of SD and the minimum duration, defined as 0.75 x duration between the first IMP intake and the first planned tumor assessment is not met) and there is no subsequent evaluable tumor assessment, or documented PD after two or more non-evaluable tumor assessments (ie, overall response of PD and the time between the date of first IMP intake and the documentation of PD is greater than the theoretical planned date of the second tumor assessment) will be considered as not evaluable for the activity population.

3 STATISTICAL ANALYSES

3.1 GENERAL CONSIDERATIONS

This study is not intended to explicitly test a hypothesis. For the primary and secondary efficacy endpoints, 95% CIs will be provided.

All efficacy endpoints based on radiological assessments of tumor burden (ie, ORR, DOR, PFS, and DCR) will be derived using the local radiologist's/Investigator's assessment.

In general, continuous data will be summarized using the number of observations available, mean, standard deviation, median, Q1, Q3, minimum, and maximum. Categorical and ordinal data will be summarized using the count and percentage of participants.

The baseline value is defined as the last available value before the first dose of IMP.

The study cut-off for analysis of the primary endpoint, ORR, corresponds to the date on which all evaluable treated participants have had at least 2 postbaseline tumor assessments, experienced confirmed objective response, or have discontinued the study treatment for any reason. For participants with 2 postbaseline tumor assessments and occurrence of response at the second postbaseline tumor assessment, it will also include the confirmatory assessment. This study cut-off can be up to approximately 20 weeks (16 weeks for 2 tumor assessments and 4 weeks for confirmation of response, if needed) after the last participant's first IMP administration. Of note, DCR will also be assessed at this cut-off.

The final study cut-off for analysis of the secondary efficacy endpoints, which include DOR and PFS, will be 4 months after the cut-off date for the primary analysis. At that time, the primary analysis of ORR and DCR will also be updated.

Observation period

The observation period will be divided into 4 segments:

- The **pre-treatment period** is defined as the period up to first IMP administration.
    - The **pre-screening period** is defined as the period from the pre-screening informed consent to the day before the screening informed consent.
    - The **screening period** is defined as the period from the screening informed consent up to the first IMP administration.
- The **on-treatment period** (ie, treatment-emergent period) is defined as the period from the first IMP administration to 30 days after the last IMP administration.
- The **post-treatment period** is defined as the period from the end of the on-treatment period.

3.2 PRIMARY ENDPOINT(S) ANALYSIS

3.2.1 Definition of endpoint(s)

The primary endpoint is confirmed objective response determined according to RECIST v1.1. A confirmed objective response is defined as a confirmed CR or PR as best overall response (BOR). The BOR will be derived according to RECIST v1.1 definitions (1, 2) based on the Investigator's assessment. The BOR is the best overall response observed from the date of the first administration of IMP until documented disease progression, death, start of an anticancer therapy, or analysis cut-off date, whichever occurs first.

3.2.2 Main analytical approach

The primary analysis will be based on a primary estimand introduced in Section 1.2.1. It is defined according to the following attributes:

- The primary endpoint is confirmed objective response (confirmed CR or PR as BOR) as per RECIST v1.1 .
- The treatment condition of interest is tusamitamab ravtansine.
- The analysis population is the all-treated population (defined in Section 2).
- Intercurrent events:
  - The study intervention discontinuation intercurrent event will be handled with the treatment policy strategy. Confirmed objective response will be assessed based on tumor assessments regardless of study intervention discontinuation.
  - The further anticancer therapy (including further systemic anticancer therapies) intercurrent event will be handled with the "while not initiating further anticancer therapy" strategy. Confirmed objective response will be assessed based on tumor assessments done up to the time of initiation of further anticancer therapy.
- The population-level summary will be the ORR, defined as the rate of participants with confirmed objective response and two-sided 95% CI using the Clopper-Pearson method.

In the absence of confirmed objective response before the analysis cut-off date (taking into account the intercurrent event handling strategies), participants will be considered as non-responders, whatever the reason (including participants with non-evaluable BOR).

3.2.3 Sensitivity analysis

No sensitivity analysis is planned in this study.

3.2.4 Supplementary analyses

The number (%) of participants within each BOR category, including not evaluable as per RECIST 1.1 and reason for being not evaluable will be provided for the all-treated population.

As a supplementary analysis, ORR as per RECIST v1.1 will also be summarized on the activity population (defined in Section 2). The same analytical approach as for the primary analysis defined in Section 3.2.2 will be used.

In addition, the best relative tumor change from baseline, defined as the smallest relative tumor change from baseline (where tumor change at time t is the difference between sum of the longest diameters of the target lesions at time t and baseline), will be summarized using a waterfall plot on the activity population (defined in Section 2).

3.3 SECONDARY ENDPOINT(S) ANALYSIS

The secondary endpoints detailed in this section are progression-free survival, disease control rate and duration of response. Other secondary endpoints analyses are defined in Section 3.6.2 (AE, SAE), Section 3.6.3 (laboratory abnormalities), Section 3.7.1.1 (PK) and Section 3.7.1.2 (immunogenicity).

3.3.1 Key/Confirmatory secondary endpoint(s)

There are no key/confirmatory secondary endpoints.

3.3.2 Supportive secondary endpoint(s)

3.3.2.1 Definition of endpoints

3.3.2.1.1 Progression-free survival

PFS is defined as the time from the date of the first administration of IMP to the date of the first documentation of objective PD according to RECIST v1.1 (1, 2) or death due to any cause, whichever comes first.

The analysis of PFS will be based on the following censoring rules:

- If documented disease progression or death is not observed before the analysis cut-off date, then PFS will be censored at the date of the last evaluable tumor assessment performed before the analysis cut-off date.
- A participant without PFS event (documented disease progression or death) and without any evaluable post-baseline tumor assessment will be censored at the date of the first administration of IMP (Day 1).

3.3.2.1.2 Disease control rate

The disease control rate is estimated by dividing the number of participants with disease control (confirmed CR, PR or SD as BOR according to RECIST v1.1) by the number of participants from the analysis population.

3.3.2.1.3 Duration of response

The duration of response is defined as the time from the date of first initial occurrence of the confirmed CR or PR, to the date of first documentation of objective PD according to RECIST 1.1 (1, 2) definitions or death due to any cause, whichever occurs first.

For participants with ongoing response at the time of the analysis cut-off date or at the time of initiation of a new anticancer therapy, DOR will be censored at the date of the last evaluable tumor assessment (not showing documented progression) performed before or on the day of initiation of a new anticancer therapy (if any) or the analysis cut-off date.

3.3.2.2 Main analytical approach

3.3.2.2.1 Progression-free survival

Analysis of PFS will be based on an estimand defined according to the following attributes:

- The endpoint is PFS.
- The treatment condition of interest is tusamitamab ravtansine.
- The analysis population is the all-treated population (defined in Section 2).
- Intercurrent events:
  - The study intervention discontinuation intercurrent event will be handled with the treatment policy strategy. PFS will be assessed based on tumor assessments regardless of study intervention discontinuation.
  - The further anticancer therapy (including further systemic anticancer therapies) intercurrent event will be handled with the hypothetical strategy. PFS will be assessed based on tumor assessments had a further anticancer therapy not being taken. PFS will be assessed based on tumor assessments up to the time of initiation of further anticancer therapy.
  - Two or more consecutive missing/unevaluable tumor assessments immediately before documented PD or death will be handled with the hypothetical strategy. PFS will be assessed based on tumor assessments had 2 consecutive tumor assessments not been missed immediately before documented PD or death. PFS will be assessed based on tumor assessments up to the last evaluable tumor assessment documenting no progression.

- The population-level summary will include:
  - Kaplan-Meier estimates of the 25th, 50th, and 75th percentiles and their associated 95% CIs. The method of Brookmeyer and Crowley and a log log transformation of the survival function will be used to construct 95% CIs.
  - Number (%) of participants at risk as well as the probabilities of being event-free at least at 2, 4, 6, 8, and 10 months with 95% CIs using the Kaplan-Meier method and a log log approach based on a normal approximation following Greenwood's formula.
  - Kaplan-Meier curves including the number of participants at risk at key time points.

  In addition, the number (%) of participants with an event and the type of event (documented disease progression or death without documented disease progression) and the number (%) of censored participants and reason for censoring (no baseline tumor assessment, no evaluable postbaseline tumor assessment, alive without documented disease progression, event occurred after 2 or more non-evaluable tumor assessments, or initiation of further anticancer therapy) will be analyzed.

  In the absence of documented disease progression or death before the analysis cut-off date (taking into account the intercurrent event handling strategies), PFS will be censored at the date of the last evaluable tumor assessment (not showing documented disease progression) performed before the analysis cut-off date, or at the date of the first administration of IMP (Day 1) if no baseline tumor assessment or no evaluable postbaseline tumor assessment.

3.3.2.2.2 Disease control rate

Analysis of the DCR will be based on an estimand defined according to the following attributes:

- The endpoint is disease control response as per RECIST v1.1.
- The treatment condition of interest is tusamitamab ravtansine.
- The analysis population is the all-treated population (defined in Section 2).
- Intercurrent events:
  - The study intervention discontinuation intercurrent event will be handled with the treatment policy strategy. Disease control response will be assessed based on tumor assessments regardless of study intervention discontinuation,
  - The further anticancer therapy (including further systemic anticancer therapies) intercurrent event will be handled with the "while not initiating further anticancer therapy" strategy. Disease control response will be assessed based on tumor assessments done up to the initiation of further anticancer therapy.
- The population-level summary will be the DCR, defined as the rate of participants with disease control response and two-sided 95% CIs using the Clopper-Pearson method.

In the absence of disease control response before the analysis cut-off date (taking into account the intercurrent event handling strategies), participants will be considered as non-responders, whatever the reason (including participants with non-evaluable BOR).

As a supplementary analysis, DCR as per RECIST v1.1 will also be summarized on the activity population (defined in Section 2). The same analytical approach as described above will be used.

3.3.2.2.3 Duration of response

Analysis of the DOR will be based on an estimand defined according to the following attributes:

- The endpoint is DOR.
- The treatment condition of interest is tusamitamab ravtansine.
- The analysis population is the subgroup of participants from the all-treated population (defined in Section 2) who achieved a confirmed objective response.
- Intercurrent events:
  - The study intervention discontinuation intercurrent event will be handled with the treatment policy strategy. DOR will be assessed based on tumor assessments regardless of study intervention discontinuation,
  - The further anticancer therapy (including further systemic anticancer therapies) intercurrent event will be handled with the hypothetical strategy. DOR will be assessed based on tumor assessments had a further anticancer therapy not being taken. DOR will be assessed based on tumor assessments up to the time of initiation of further anticancer therapy,
  - Two or more consecutive missing/unevaluable tumor assessments immediately before documented PD or death will be handled with the hypothetical strategy. DOR will be assessed based on tumor assessments had 2 consecutive tumor assessments not been missed immediately before documented PD or death. DOR will be assessed based on tumor assessments up to the last evaluable tumor assessment documenting no progression.
- The population-level summary will include the median DOR and associated 95% CI using Kaplan-Meier methods.

In the absence of documented disease progression or death before the analysis cut-off date (taking into account the intercurrent event handling strategies), DOR will be censored at the date of the last evaluable tumor assessment (not showing documented disease progression) performed before the analysis cut-off date. In the absence of confirmed objective response before the analysis cut-off date (taking into account the intercurrent event handling strategies), DOR will not be derived.

3.4 TERTIARY/EXPLORATORY ENDPOINT(S) ANALYSIS

Tertiary endpoints analyses are defined in Section 3.7.2 (biomarker).

3.4.1 Quality of life endpoints

Patient-reported outcomes (PRO) will be analyzed as tertiary endpoints in this study as described in Table 1. Four PRO scale will be used:

- The Non-Small-Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ) to evaluate time to deterioration in the overall severity of disease-related symptoms (cough, pain, dyspnea, fatigue, appetite),
- The patient global impression of severity- lung cancer symptom (PGIS-LCS),
- The patient global impression of change- lung cancer symptom (PGIC-LCS),
- The Functional Assessments of Cancer Therapy Item GP-5 (FACT-GP5) to evaluate change from baseline in overall side effect impact.

NSCLC-SAQ

The 7 items of the NSCLC-SAQ assess 5 different cardinal NSCLC related symptom concepts over a 7 days recall period: cough (1 item), pain (2 items), dyspnea (1 item), fatigue (2 items), and appetite (1 item). The measure uses a 5-point Likert-type, verbal rating scales ranging from 0: “No at All” to 4: “Very severe” for cough and pain items and from 0: “Never” to 4: “Always” for dyspnea, fatigue, and appetite items. Domain scores are calculated based on the scale developer’s procedures to form a total score, where higher scores indicate higher NSCLC symptom severity (3, 4). Two pain items form a single “Pain” domain representing the highest severity of the two items, and two fatigue items form a single “Fatigue” domain by calculating their mean. The NSCLC-SAQ total score is computed as the sum of the 5 domains and ranges between 0 and 20 (5). Missing data will follow missing data/handling procedures recommended in the NSCLC-SAQ user manual. At form-level, if a respondent does not complete the NSCLC-SAQ, such as due to attrition in longitudinal studies or due to forgetting to complete an individual assessment, his or her NSCLC-SAQ score should not be computed for that time point. At item-level, if a respondent is missing any of the five domain scores, his or her NSCLC-SAQ score should not be computed. It is possible for respondents to miss one of the Pain items or one of the Fatigue items and still have a NSCLC-SAQ total score calculated. No additional imputations will be conducted for missing data.

PGIS-LCS and PGIC-LCS

The PGIS-LCS (1 item) measure severity in lung cancer symptoms at the time of assessment; while the PGIC-LCS (1 item) measure change in lung cancer symptoms over time. PGIS-LCS measure use a unidirectional verbal rating scale, with 5 options ranging from “None” to “Very severe”. PGIC-LCS measure use a bi-directional response scale, with 5 options ranging from “Much better” to “Much worse” with “No change” at the midpoint.

FACT-GP5

The FACT-GP5 (“I am bothered by side effects of treatment”) is a single item from the Functional Assessment of Cancer Therapy General (FACT-G) scale assesses the overall impact of treatment side effects (6). Responses are given on a 5-point Likert-type scale recalling the past 7 days. Higher scores indicate a higher degree of side effect bother (7).

The baseline value is defined as the last questionnaire assessment before or on the day of first IMP. PRO post-baseline assessments will consider the on-treatment period defined as the period from the first IMP administration to 30 days after the last IMP administration. No time window will be defined and visit descriptive statistics will be based on visits as reported in the database.

3.4.1.1 Completion rate

Instrument completion rate by visit will be reported for each score:

- Unadjusted completion rate by visit will be calculated as the number of participants meeting at least the minimum requirements for scoring of the instrument divided by the number of participants in the all-treated population
- Adjusted completion rate by visit will be calculated as the number of participants meeting at least the minimum requirements for scoring of the instrument among those who were expected to complete the questionnaire. To be noted that a participant is expected to complete the PRO assessment if the participant is alive and still on-treatment.

3.4.1.2 Evolution over time

Descriptive statistics will be provided for each score: NSCLC-SAQ total score as continuous data, then PGIS-LCS, PGIC-LCS and FACT-GP5 as ordinal data. Change from baseline will be presented for FACT-GP5 score. Graphical visualizations such as line graph with mean values and corresponding 95% confidence intervals might be considered.

3.4.1.3 Time to deterioration

The time to deterioration (TTD), in disease related symptoms (cough, pain, dyspnea, fatigue, appetite) from NSCLC-SAQ, is defined as the time from baseline until the first ≥2-point change from baseline (considered as clinically meaningful change) up to the end of treatment assessment before the analysis cutoff date (5). In case of missing score, no imputation will be done, and the assessment will be done according to the scales available.

Participants with a non-missing baseline assessment will be censored at the last on-treatment assessment before the start of further systemic anticancer therapy or before the analysis cutoff date (whichever is earlier), provided their symptoms scale had not deteriorated up to that point. Participant without baseline or post-baseline electronic patient-reported outcomes (ePROs) questionnaire or whose baseline scores do not allow further deterioration will be censored at the first IMP day.

TTD will only be analysed for the SAQ as there are no interpretation thresholds for the other measures.

TTD data will be summarized using the Kaplan-Meier method:

- Kaplan-Meier estimates of the 25th, 50th, and 75th percentiles and their associated 95% CIs will be provided. The 95% CIs will be constructed using a log-log transformation of the survival function and the method of Brookmeyer and Crowley.
- Number of participants at risks as well as the probabilities of being event-free at for example least 2, 4, 6, 8 ,10 months with 95% CIs will be estimated using the Kaplan-Meier method.
- The number of censored participants, the reasons and timing for their censoring (censored at first IMP administration, censored at last non-missing assessment before the initiation of a further systemic anti-cancer therapy, censored at last non-missing assessment before the cut-off date, censored at last non-missing on-treatment assessment). Kaplan-Meier curves will be plotted. These plots will include the number of participants at risk at key time points.

3.5 MULTIPLICITY ISSUES

No multiplicity issues are anticipated in this study.

3.6 SAFETY ANALYSES

All safety analyses will be performed on the all-treated population as defined in Section 2, unless otherwise specified, using the following common rules:

- The analysis of the safety variables will be essentially descriptive, and no testing is planned.
- Safety data in participants who do not belong to the all-treated population (eg, treated but not enrolled) will be provided.

3.6.1 Extent of exposure

The dose information will be assessed by the following:

- Number of cycles started per participant.
- Duration of exposure (in weeks) is defined by (date of last administration of the IMP + 14 days – date of first administration of the IMP)/7.
- Actual dose (in mg/m²). In case of dose interruption, actual dose will be the sum of the actual doses administered before and after the dose interruption.
- Cumulative dose (in mg/m²): the cumulative dose is the sum of all actual doses of tusamitamab ravtansine, given from first to last administration
- Actual dose intensity (ADI in mg/m²/week): defined as the cumulative dose divided by the duration of tusamitamab ravtansine exposure (in weeks)

- Planned dose intensity (PDI in mg/m²/week): corresponds to the planned dose at C1D1 + planned dose at subsequent cycles multiplied by the theoretical total number of doses started - 1 and divided by the theoretical cycle duration expressed in weeks (ie, 2 weeks per cycle started)
- Relative dose intensity (RDI, in %):100 × ADI ([mg/m²/week])/PDI ([mg/m²/week])

The total number of cycles started, number of cycles started by participant will be summarized as a quantitative variable and by category (number [%] of participants receiving at least 1 cycle, at least 2 cycles, etc). Duration of tusamitamab ravtansine exposure, cumulative dose, ADI and RDI will be summarized quantitatively.

The following variables will be derived to describe dose modifications and dose interruptions:

- Dose reduction: The first administration (loading dose) and second administration (first 100 mg/m^2 infusion) will not be counted as a dose reduction. For the third and subsequent tusamitamab ravtansine administrations, dose reduction will be determined using the dose level intervals provided in Table 6, by comparing the current dose level to the previous dose level. If the current dose level is below the dose level interval of the previous dose administration, then the current dose Table 6 level is considered reduced.

Table 6 - Tusamitamab ravtansine dose reduction criteria

Actual dose level	Dose level interval
Starting dose (100 mg/m² Q2W)	>90 mg/m²
Dose level -1 (80 mg/m² Q2W)	>72.5 mg/m² and ≤90 mg/m²
Low dose	>0 mg/m² and ≤72.5 mg/m²

- Delay: A cycle will be considered as delayed if the tusamitamab ravtansine administration date of the current cycle - 14 days - tusamitamab ravtansine administration date of the previous cycle is >2 days. Cycle delay is not defined for the first cycle.
- Dose interruption: A dose will be considered as interrupted if the tusamitamab ravtansine administration is stopped during an infusion regardless of whether it is restarted or not.

Dose modifications and dose interruptions will be analyzed by participant and cycle as follows:

- **Participant** (number of participants treated will be used as denominator)
  - Number (%) of participants with at least 1 dose modification
    - Number (%) of participants with at least 1 dose delayed
    - Number (%) of participants with at least 1 dose reduction
  - Number (%) of participants with a least 1 dose interruption

- **Cycle** (number of cycles started will be used as denominator)
  - Number (%) of cycles with at least 1 dose modification
    - Number (%) of cycles with at least 1 dose delayed
    - Number (%) of cycles with at least 1 dose reduction
  - Number (%) of cycles with at least 1 dose interruption

3.6.2 Adverse events

General common rules for adverse events

All AEs will be graded according to the National Cancer Institute Common Terminology for Adverse Events (NCI-CTCAE version 5.0) and coded to a lower-level term (LLT), preferred term (PT), high-level term (HLT), high-level group term (HLGT), and associated primary system organ class (SOC) using the Medical Dictionary for Regulatory Activities (MedDRA) version in effect at Sanofi at the time of database lock.

The AEs will be analyzed in the following 3 categories:

- Pre-treatment AEs: AEs that developed, worsened or became serious during the pre-treatment period.
- Treatment-emergent adverse events (TEAE)s: AEs that developed, worsened or became serious during the treatment-emergent period.
- Post-treatment AEs: AEs that developed, worsened or became serious during the post-treatment period.

Similarly, deaths will be analyzed in the pre-treatment, treatment-emergent and post-treatment periods.

The primary AE analyses will be on TEAEs. Pre-treatment and post-treatment AEs will be described separately.

An AE with incomplete or missing date/time of onset (occurrence, worsening, or becoming serious) will be classified as a TEAE unless there is definitive information to determine it is a pre-treatment or a post-treatment AE.

If the assessment of the relationship to IMP is missing for an AE, this AE will be assumed as related to IMP. Missing grade will be left as missing.

Multiple occurrences of the same event in the same participant will be counted only once in the tables within a treatment phase, using the maximum (worst) grade by treatment phase. Summaries will be provided for all grades combined and for grade ≥3 (including grade 5). Missing grades, if any, will be included in the “All Grades” category.

The AE tables will be sorted as indicated in Table 4.

Table 4 - Sorting of AE tables

AE presentation	Sorting rules
SOC, HLGT, HLT and PT	By the internationally agreed SOC order and by alphabetic order of HLGTs, HLTs and PTs.
SOC and PT	By the internationally agreed SOC order and decreasing frequency of PTs[a, b]
[SMQ/CMQ] and PT	By decreasing frequency of SMQs/CMQs and PTs[a]
PT	By decreasing frequency of PTs[a]

a Sorting will be based on the SAR408701 group incidence.
b The table of all TEAEs presented by primary SOC and PT will define the presentation order for all other tables (eg, treatment-emergent SAEs) presented by SOC and PT, unless otherwise specified.

Analysis of all adverse events

The overview of TEAEs with the details below will be generated:

- Any TEAE
- Any grade ≥3 TEAE
- Any treatment-emergent SAE
- Grade 5 TEAE (any TEAE with a fatal outcome during the treatment-emergent period)
- Any treatment-emergent AESI
- Any TEAE leading to permanent intervention discontinuation
- Any TEAE related to IMP
- Any Grade ≥3 TEAE related to IMP
- Any treatment-emergent SAE related to IMP
- Any treatment-emergent corneal event (CMQ "Corneal events compound level")
- Any treatment-emergent peripheral neuropathy event (SMQ "Peripheral neuropathy" [Narrow and Broad])
- Any ocular/visual symptoms TEAE (CMQ "Eye disorders exclude corneal disorders" ie, all PTs included in "Eye disorders" [SOC] excluding PTs in SMQ "Corneal disorders" [Narrow]")

The AE summaries of Table 5 will be generated with number (%) of participants experiencing at least one event. The analyses will be performed for all grades combined and for grades ≥3.

Table 5 - Analyses of adverse events

Type of AE	MedDRA levels
All TEAE	Primary SOC, HLGT, HLT and PT
	Primary SOC and PT
	PT
TEAE related to IMP as per Investigator's judgment	Primary SOC and PT
Treatment emergent SAE	Primary SOC, HLGT, HLT and PT
	Primary SOC and PT
Treatment emergent SAE related to IMP as per Investigator's judgment	Primary SOC and PT
TEAE leading to permanent intervention discontinuation	Primary SOC and PT
TEAE leading to death[b]	Primary SOC and PT
AE leading to death[b]	Primary SOC and PT
- In context of disease progression[c]	
- In context other than disease progression[d]	
	Primary SOC and PT
Pretreatment AE	Overview[a]
	Primary SOC and PT
Pretreatment SAE	Primary SOC and PT
Post-treatment AE	Overview[a]
	Primary SOC and PT
Post-treatment SAE	Primary SOC and PT
TEAE leading to dose modification (including dose delay and dose reduction)	Primary SOC and PT
TEAE leading to dose interruption	Primary SOC and PT

a Will include the following AE categories: any AEs, any serious AEs, any AEs leading to death, any AEs leading to permanent intervention discontinuation
b Death as an outcome of the AE as reported by the Investigator in the AE page
c Death within 30 days from last IMP administration and the cause of death is disease progression
d Death within 30 days from last IMP administration and for whom cause of death is not disease progression or the death occurred more than 30 days from last IMP administration and the cause of death is AE

Analysis of deaths

In addition to the analyses of deaths included in Table 4, the number (%) of participants in the following categories will be provided:

- Deaths during the treatment-emergent and post-treatment periods by study period and main reason for death

- An overview of Grade 5 AEs will be provided with the following categories:
  - Grade 5 AE (TEAE and post-treatment).
  - Fatal TEAE (regardless of date of death/period).
    - Grade 5 TEAE with a fatal outcome during the treatment-emergent period,
    - Any Grade TEAE with a fatal outcome during the post-treatment period.
  - Post-treatment Grade 5 AE (excluding a TEAE that worsened to Grade 5 during the post-treatment period).

In addition, a listing of deaths in non-enrolled participants or enrolled but not treated participants will be provided on the prescreened population.

Analysis of adverse events of special interest (AESIs) and other AEs of interest

Adverse events of special interest (AESIs) and other AEs of interest will be selected for analyses as indicated in Table 6. Number (%) of participants experiencing at least one event will be provided, by SOC and PT if applicable. Tables will be sorted as indicated in Table 4.

Table 6 - Selections for AESIs and other AEs of interest

AESIs and other AEs of interest	Selection
AESIs	
Grade ≥3 keratopathy/keratitis	e-CRF AESI specific tick box on the AE page. It could include Grade ≥3 events with PTs from CMQ of corneal events
Bundle branch blocks or any conduction defects.	e-CRF AESI specific tick box on the AE page. It could include events with PTs from SMQ "Conduction defects"
Grade ≥3 liver enzyme increased (symptomatic or asymptomatic)	e-CRF AESI specific tick box on the AE page. It could include Grade ≥3 events with the following PTs: "Alanine aminotransferase increased", "Aspartate aminotransferase increased", "Transaminases increased", "Hepatic function abnormal", "Hepatic enzyme increased", "Liver function test increased", "Transaminases abnormal", "Hepatic enzyme abnormal", or "Liver function test abnormal"
Symptomatic overdose (serious or nonserious) with IMP	e-CRF symptomatic overdose specific tick box on the Overdose page
Pregnancy of a female participant entered in a study as well as pregnancy occurring in a female partner of a male participant entered in a study with IMP	eCRF specific tick box for adverse event linked to pregnancy on the Pregnancy page
Other AEs of interest	
Corneal events	CMQ "Corneal events compound level" containing the PTs included in both SOC "Eye disorders" and SMQ "Corneal disorders" (Narrow)
Ocular/visual adverse events (excluding corneal disorders)	CMQ "Eye disorders exclude corneal disorders" containing PTs included in SOC "Eyes disorders" and excluding PTs in SMQ "Corneal disorders" (Narrow)
Cardiac conduction defects	SMQ "Conduction defects" (Narrow)

AESIs and other AEs of interest	Selection
Peripheral neuropathy events	SMQ "Peripheral neuropathy" (Narrow and Broad)
Colitis events (excluding infective)	CMQ "Colitis (excluding infective)" containing PTs included in HLT "Colitis (excl infective)"
Hypersensitivity events	SMQ "Hypersensitivity" (Narrow) and defined as events occurring on the day or the day after the infusion
Hepatic disorders adverse events	SMQ "Hepatic Disorders" (Narrow and Broad)
Hematological adverse events	SMQ "Haematopoietic cytopenias" (Narrow and Broad)
AE related to COVID-19 illness	SMQ "COVID-19" (Narrow)

An overview of corneal TEAE will be provided with the following AE categories: any corneal TEAE, Grade ≥3 corneal TEAE, treatment-emergent corneal SAE, corneal TEAE related to IMP, Grade ≥3 corneal TEAE related to IMP, corneal TEAE leading to treatment discontinuation, and corneal TEAE leading to dose modification (cycle delay or dose reduction). In addition, a summary table of corneal events will be displayed by grade. A summary of treatment-emergent corneal events will be provided.

- Cycle of first onset of corneal event regardless of the grade
- Cycle of first onset of corneal event with the worst grade
- Number (%) of participants by worst grade
- Relationship to the study intervention: in case of multiple events with different relationships, if any event is related, then the relationship will be considered as related
- Action taken with the study intervention: in case of multiple events with different actions, the most severe action taken will be tabulated and selected according to the following order of severity: drug withdrawn, dose reduced and delayed, dose reduced, dose delayed, drug interrupted, dose not changed
- Outcome: in case of multiple events with different outcomes, the most severe outcome will be tabulated and selected according to the following order of severity: fatal, not recovered or resolved, recovering or resolving, recovered or resolved with sequelae, recovered or resolved, unknown

In addition, analyses on occurrence and recurrence of corneal events will be provided.
An occurrence of corneal event is defined as one or a group of concomitant corneal events.
A recurrence is defined as any new occurrence of corneal event starting after a previous resolved occurrence.

- The number of occurrences by participant
- The time to first onset of corneal event will be described using Kaplan-Meier curves. Time to first onset is defined as the time from the date of first IMP administration to the date of the first occurrence of the event. In the absence of an event before the analysis cut-off date, it will be censored at the end date of the treatment-emergent period, analysis cut-off date or date of death, whichever occurs first.

- The time to recovery will be summarized using descriptive statistics in participants who had had at least one recovered or resolved occurrence of corneal event (with or without sequelae), considering the longest duration among all occurrences by participant.
- The time to recurrence will be summarized using descriptive statistics in participants who had had at least one recurrence, considering the shortest time among all recurrences by participant.

Besides the AE categorized as ocular/visual adverse events, all ocular/visual symptoms (coded term) recorded in e-CRF as associated symptom of the reported ocular treatment-emergent adverse events will be reported as descriptive analysis. Same analysis will be done separately for ocular/visual symptoms associated to corneal TEAEs (CMQ)."

3.6.3 Additional safety assessments

3.6.3.1 Laboratory variables, vital signs and electrocardiograms (ECGs)

The following laboratory variables, vital signs and electrocardiogram (ECG) variables will be analyzed. They will be converted into standard international units.

- Hematology and coagulation:
  - Red blood cells and platelets and coagulation: hemoglobin, hematocrit, red blood cell count, platelet count, prothrombin time (expressed as international normalized ratio)
  - White blood cells: white blood cell count, neutrophils, lymphocytes, monocytes, basophils, eosinophils
- Clinical chemistry:
  - Metabolism: glucose, total protein, albumin
  - Electrolytes: sodium, potassium, chloride, corrected calcium, phosphate
  - Corrected calcium (mmol/L) will be derived using the following formula: measured total calcium (mmol/L) + 0.8 x 0.25 × (4.0 – [serum albumin (g/L) × 0.1]), where 4.0 represents the average albumin level.
  - Renal function: creatinine, estimated glomerular filtration rate (eGFR) by the MDRD method, blood urea nitrogen, urea
  - Liver function: alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, lactate dehydrogenase, total and direct bilirubin
- Vital signs: pulse rate, systolic and diastolic blood pressure, weight, temperature, ECOG Performance status
- ECG variables: heart rate, PR, QRS, QT, and corrected QTc (according to Fredericia)

Data below the lower limit of quantitation/detection limit (LLOQ) will be replaced by half of the LLOQ, data above the upper limit of quantification will be replaced by ULOQ value.

For hematological and coagulation parameters and some selected biochemistry parameters, Sanofi sponsor generic ranges (LLN, ULN) are defined and will be used for grading (see list of parameters in Section 5.5). For other biochemistry parameters, grading will be derived using local laboratory normal ranges.

Quantitative analyses

When relevant, for vital signs and ECG variables, descriptive statistics for results and changes from baseline will be provided for each planned visit and at the worst value (minimum and/or maximum value depending on the parameter) during the on-treatment period. These analyses will be performed using local measurements.

For QRS, QT and QTcF variables, blood pressure and heart rate, mean changes from baseline with the corresponding standard error (SE) will be plotted over time.

Analyses according to Potentially clinically significant abnormality (PCSA) and NCI grading

For laboratory variables, analyses according to NCI grading will be made based on NCI-CTCAE version 5.0. In addition, for laboratory variables for which NCI-CTCAE scale is not applicable, vital signs and ECG variables, PCSA analyses will be performed based on the PCSA list currently in effect at Sanofi at the time of the database lock.

Analyses according to PCSA and NCI grading will be performed based on the worst value during the treatment-emergent period, using all measurements (either local or central, either scheduled, nonscheduled or repeated).

QTcF prolongation will be graded according to NCI-CTCAE version 5.0. The frequency of participants in each grade of QTcF prolongation during the on-treatment period will be summarized. For participants with multiple occurrences of QTcF prolongation during the treatment, the maximum grade by participant will be used.

For laboratory variables, vital signs and ECG variables, the incidence of participants with at least one PCSA during the treatment-emergent period will be summarized regardless of the baseline level and according to the following baseline status categories:

- Normal/missing.
- Abnormal according to PCSA criterion or criteria.

The number (%) of participants with QTcF abnormality worsening during the on-treatment period (worst value per participant) according to baseline value will be displayed.

For laboratory variables graded by NCI-CTCAE version 5.0.

- The number (%) of participants with abnormal laboratory tests at baseline will be presented by grade.
- The number (%) of participants with abnormal laboratory tests during the treatment-emergent period will be summarized by grade. When appropriate, the number (%) of participants with abnormality of any grade and with Grade 3-4 abnormalities will be provided.

For laboratory variables not graded by NCI-CTCAE nor by PCSA, the number (%) of participants with laboratory test outside normal ranges will be provided.

Additional analyses for drug-induced liver injury

The following additional analyses will be performed for drug-induced liver injury:

- For each liver function test (eg, ALT), participants having experienced a PCSA will be summarized considering worst on-treatment value.
- A graph of the distribution of peak values of ALT and AST versus peak values of total bilirubin during the treatment-emergent period will be provided.
- For each liver function test (eg, ALT), participants having experienced a PCSA (eg, ALT >5 ULN) will be summarized using the following categories: Returned to baseline PCSA status (or returned to value ≤ ULN in case of missing baseline) before last IMP dose, Returned to baseline PCSA status after last IMP dose, Never returned to baseline PCSA status, No assessment after elevation. This summary will be performed by categories of elevation (ALT >3, >5, >10, >20 ULN).
- A listing of possible Hy's law cases identified (eg, participants with any elevated AST or ALT of >3 ULN and elevated Total bilirubin >2 ULN, 2 days apart), will be provided, displaying ALT, AST, Total bilirubin and ALP values.

3.6.3.2 Ocular examinations

Schirmer test

Participants reported Schirmer's test will be classified into three classes at baseline: normal ≥10 mm, moderate >5 and <10 mm, severe ≤5 mm (8). The worst classification between the laterality will be considered. A frequency table of the baseline status with and without anaesthetics will be provided by treatment group for the participants with corneal events (CMQ) during the treatment-emergent period and for the participants without event.

Visual acuity test

Analyses are done on the best corrected visual acuity assessed using Snellen Chart method.

Frequency tables on best corrected visual acuity (BCVA) measured during on-treatment period will be performed separately on all participants and on participants who experienced at least one treatment-emergent corneal event:

- on the worst change from baseline: No worsening versus baseline, Worsening versus baseline (1 to 3 lines decrease versus baseline, or >3 lines decrease versus baseline)
- on the worst absolute value: No change versus baseline, Worsening versus baseline: BCVA equals to 20/40 or better, BCVA worse than 20/40 up to 20/200, BCVA 20/200 or worse.

- on the CTCAE worst vision decrease: No worsening versus baseline, BCVA equals to 20/40 or better or 1 to 3 lines decrease versus baseline, BCVA worse than 20/40 up to 20/200 or >3 lines decrease versus baseline, BCVA 20/200 or worse.

- For participants who had worsening versus baseline on BCVA, the outcome on the last BCVA value will be displayed: Recovered to baseline, Not recovered to baseline, Lost to follow-up or dead with ongoing corneal events

A participant will be considered as lost to follow-up if the participant discontinued the study and did not perform any follow-up visit.

A shift table of the category of the last BCVA value (Normal (20/20 or better), worse than 20/20 up to 20/40, worse than 20/40 up to 20/200 or 20/200 or worse) versus the category of the worst BCVA value (20/40 or better, worse than 20/40 up to 20/200, 20/200 or worse) will be performed for participants who experienced at least one treatment-emergent corneal event.

For the summary table on participants who experienced at least one treatment-emergent corneal event, the worst value and the worst change from baseline are measured during any of the treatment-emergent corneal events experienced by the participant (between start date and end date of the corneal events). For the summary table on all participants, all on-treatment BCVA values are considered.

The worst classification between the laterality will be considered for the worst value of participant either left or right eye then the change from baseline and the last value will be measured from same eye with worst value. If the worst absolute value is the same for both eyes, then the eye with the worst change from baseline is considered, and if identical, then the eye with the worst last value is considered.

Participants whose baseline visual acuity had been reported in naked eye instead of BCVA (ie, whenever baseline value was worse than the values reported during the during the treatment period/treatment-emergent corneal events) will be excluded from this analysis.

Slit lamp examination

Descriptive statistics of slit lamp examination will be provided separately at baseline and at the time of first abnormal slit lamp after occurrence of a treatment-emergent corneal event and at the time of the worst BCVA value during a treatment-emergent corneal event for participants experiencing treatment-emergent corneal events (CMQ). The outcome (normal, abnormal), and for abnormal findings, the type of lesions and the distribution will be described by laterality (unilateral, bilateral, all).

3.6.3.3 ECOG PS

A shift table of baseline ECOG PS versus last and worst ECOG PS on treatment will be provided.

3.7 OTHER ANALYSES

3.7.1 Other variables and/or parameters

3.7.1.1 PK analyses

All concentration values below the LLOQ will be treated as zero in individual listings and respective descriptive statistics. The geometric mean will not be computed in case at least one concentration is below LLOQ.

Individual observed predose concentrations (C_{trough}) and concentrations observed at end of infusion (C_{eoi}, C_{eoi}+1h if any) of tusamitamab ravtansine will be listed and summarized with standard descriptive statistics (such as the number of observations, arithmetic and geometric means, median, standard deviation, standard error, coefficient of variation, minimum, and maximum) by visit.

For the descriptive statistics, Ctrough following any dose modification (delay or reduction as defined in Section 3.6.1 will be excluded. Ceoi and Ceoi+1h will be excluded following dose reduction.

A graphical representation of mean Ctrough (±SD) profile over time, will be provided throughout the course of treatment.

3.7.1.1.1 Population PK analysis

The population PK analyses will be described in a specific document and the results will be presented separately from the clinical study report.

3.7.1.1.2 IgG

In the all-treated population, the level of IgG in blood at pre-infusion of Cycle 1 Day 1 will be summarized with standard descriptive statistics (such as the number of observations, arithmetic and geometric means, median, standard deviation, coefficient of variation, minimum and maximum). A listing will also be provided.

3.7.1.2 Immunogenicity analyses

Participant's ATA status, response variable, and kinetics of ATA responses (see definitions below) will be summarized on the ATA population.

Kinetics of ATA responses will be described for participants with treatment-induced ATA and for participants with treatment-boosted ATA, separately. Time to ATA onset and duration of ATA will be described with minimum, Q1, median, Q3, and maximum statistics.

Peak titer will be described with minimum, Q1, median, Q3, and maximum statistics for participants with treatment-induced ATA and for participants with treatment-boosted ATA, separately.

Sample status (negative, positive, inconclusive) and titers will also be described overtime using descriptive statistics.

The impact of positive immune response on efficacy, PK, and safety variables may be further explored, depending on ATA incidence.

Participant's ATA status

- Participants with **pre-existing ATA**s correspond to participants with ATAs present in samples drawn before first administration of intervention. Participants with missing ATA sample at baseline will be considered as without pre-existing ATA.
- Participants with **treatment-emergent ATA** correspond to participants with at least 1 treatment-induced/boosted ATA.
  - Participants with **treatment-induced ATA**s correspond to participants with ATAs that developed at any time after IMP administration and without pre-existing ATA (including participants without pre-treatment samples).
  - Participants with **treatment-boosted ATA**s correspond to participants with pre-existing ATAs that are boosted at any time after IMP administration to a significant higher titer than the baseline. A 2-fold serial dilution schema is used during titration, so at least a 4-fold increase will be considered as significant.
- Participants with **unclassified ATA** correspond to participants with pre-existing ATAs that cannot be classified as treatment-boosted ATA because of missing titer(s) (ie, a positive ATA sample at any time after IMP administration in a participant with pre-existing ATA but with missing titer at this sample or at baseline).
- Participants **without treatment-emergent ATA** correspond to participants without treatment-induced/boosted ATA and without any inconclusive sample nor unclassified ATA at any time after IMP administration .
- Participants **with inconclusive ATA** are defined as participants which cannot irrefutably be classified as with or without treatment-emergent ATA.

Kinetics of ATA response

Kinetics of ATA response will be derived for participants with treatment-induced/boosted ATA.

- **Time to onset of ATA response** is defined as the time period between the first IMP administration and the first treatment-induced/boosted ATA.
- **Duration of ATA response** is defined as the time between the first treatment-induced/ boosted ATA and the last treatment-induced/boosted ATA, irrespective of negative samples or positive samples not reaching the boosted threshold in-between. ATA duration will be summarized only for participants with persistent ATA response.

- **Persistent ATA response** is defined by treatment-induced/boosted ATA with a duration of ATA response of at least 16 weeks.
- **Transient ATA response** is defined by treatment-induced/boosted ATA with a duration of ATA response of less than 16 weeks and the last sample is not treatment-induced/boosted.
- **Indeterminate ATA response** is defined by treatment-induced/boosted ATA that are neither persistent nor transient.

ATA response variable:

- **ATA incidence** is defined as the proportion of participants found to have seroconverted (treatment-induced ATAs) or boosted their pre-existing ATA response (treatment-boosted ATAs) at any time point after first IMP administration.

3.7.2 Biomarker analyses

Several exploratory objectives related to pharmacodynamic/genomic endpoints will be considered in this study.

Unless otherwise specified, biomarkers analyses will be performed on participants with available data from the all-treated population.

3.7.2.1 Circulating CEA

The circulating CEA will be considered as a quantitative variable and circulating CEA values below the LLOQ will be replaced by half of the LLOQ.

3.7.2.1.1 *Circulating CEA levels before IMP and CEACAM5 IHC*

Circulating CEA levels before IMP will be correlated with IHC CEACAM5 expression status (for prescreened participants). To this end, the closest circulating CEA assessment to biopsy (or the oldest value before IMP if the date of the biopsy is missing) will be considered. The circulating CEA levels (quantitative) will be presented using descriptive statistics, according to IHC CEACAM5 expression status (negative vs. moderate expressers; or according to the % of tumoral cells expressing CEACAM5 at intensity ≥2+). The association between the circulating CEA levels (quantitative) and the CEACAM5 expression will be also assessed and visualized graphically. The time from tumor biopsy collection and circulating CEA assessment (the closest to biopsy) will be also described and investigated in the correlation between circulating CEA levels and IHC CEACAM5 expression.

In order to evaluate the variability of circulating CEA values, an intraclass correlation coefficient will be calculated between pre-screening and screening circulating CEA values (when both available). A graphical visualization may be provided to observe the degree of concordance. Similarly, the concordance between the circulating CEA and the circulating CEACAM5 will be evaluated.

3.7.2.1.2 Modulations of circulating CEA as a potential PD biomarker of response to tusamitamab ravtansine treatment

The circulating CEA at different timepoints and its relative change from baseline will be presented using descriptive statistics by response status and overall. Depending on the data available, some timepoints may not be presented when the sample size is too small.

A graphical visualization (spaghetti plot) will be presented to visualize the relative change from baseline of circulating CEA values for each participant, annotated with the best overall response.

3.7.2.2 Circulating free DNA (cfDNA)

3.7.2.2.1 Relationship between the tumor mutation profiles detected in the circulating free DNA (cfDNA) at baseline with efficacy outcomes

Genomic alterations collected locally from the eCRF at pre-screening will be described. For each gene, participants will be grouped into two groups: “altered” (ie, participants with at least one alteration for the gene) and “wild-type” (ie, participants without alteration for the gene) and described using descriptive statistics by IHC CEACAM5 expression status and overall. The confirmed objective response rate and two-sided 95% confidence intervals using the Clopper-Pearson method will be presented for each gene in the all-treated population.

Circulating-free DNA (cfDNA) centrally assessed will be collected and analyzed in the same way as local data from eCRF.

3.7.2.3 To explore potential sets of biomarkers from tumor DNA and RNA analyses as potential biomarkers of response to tusamitamab ravtansine

The same tumor biopsy as the one sent for CEACAM5 assessment at pre-screening may be used for additional tumor DNA and/or RNA analyses, if enough relevant data is collected.

Additional exploratory analyses on biomarkers from tumor DNA and RNA analyses, beside target expression, as potential biomarkers of response to tusamitamab ravtansine might be conducted..

3.7.3 Subgroup analyses

Subgroup analyses of the primary efficacy endpoint (defined in Section 3.2.1) will be performed to assess the homogeneity of the treatment effect across the following subgroups (categories with fewer than 5 participants may be combined with other categories):

- Race (White, Other)
- Age group (<65, ≥65 years)
- Gender (Male, Female).
- ECOG PS at baseline (0,1)
- Smoking status (Never, Former and Current)

- Programmed death ligand-1 (PD-L1) expression (<1%, ≥1%)
- Circulating CEA at baseline (≥100 µg/L),
- Circulating CEA at baseline and CEACAM5 expression (CEACAM5 1-49% and CEA <100 µg/L, CEACAM5 1-49% and CEA ≥100 µg/L, CEACAM5 <1% and CEA <100 µg/L, CEACAM5 <1% and CEA ≥100 µg/L)
- Prior treatment with EGFR, ALK, or ROS1 inhibitors (Yes, No)
- Time from initial diagnosis (<12 months, ≥12 months)
- Brain metastases (Yes, No)
- Number of organs involved (including primary tumor location) (<3, ≥3)

If a sufficient number of responses is observed, tables will be provided, otherwise subgroup information will be displayed in listings only. For tables, Objective response rate estimate and the corresponding 95% CI, will be provided for each subgroup, using the same method as applied to the primary analysis. Forest plots will be provided.

3.8 INTERIM ANALYSES

An interim analysis for futility based on the number of confirmed objective responses observed in the activity population will be performed.

The study cut-off date for the interim analysis corresponds to the date on which the first 20 evaluable treated participants have had at least 2 postbaseline tumor assessments, experienced confirmed objective response, or have discontinued the study for any reason. For participants with 2 postbaseline tumor assessments and occurrence of response at the second postbaseline tumor assessment, the confirmatory assessment will also be included.

If 0 or 1 confirmed objective response is observed among the first 20 treated participants evaluable for anti-tumor activity, the enrollment will be stopped. Otherwise, the enrollment will continue with the 18 additional evaluable participants.

The study analysis will then be conducted in 2 steps.

The first step analysis will be conducted when all evaluable treated participants have had at least 2 postbaseline tumor assessments, experienced confirmed objective response, or have discontinued the study treatment for any reason. This study cut off can be up to approximately 20 weeks (16 weeks for 2 tumor assessments and 4 weeks for confirmation of response, if needed) after the last participant's first IMP administration. Of note, DCR will also be assessed at this cut-off.

The second step analysis for analysis of the secondary efficacy endpoints will be conducted approximately 4 months after the cut-off date for the primary analysis. At that time, the primary analysis for ORR and DCR will also be updated.

The final analysis will be conducted at the end of the study. Only safety analyses will be updated at this time.

For each analysis step, the analysis cut-off date will be defined as the date of the database extraction that will be performed for the analysis after the cut-off (study cut-off for the interim analysis, as well as the first and second analysis steps).

Analyses methods and conventions described in the other sections of this SAP will be applied for all analyses as applicable. The following additional rules will apply for analyses performed at interim analysis and at first analysis step:

- Participants without end of treatment visit performed at the time of the cut-off date will be considered as ongoing and exposed up to the cut-off date. Therefore:
  - Participants who did not complete treatment period nor prematurely discontinued the study intervention at cut-off date will be analyzed as “ongoing” in the disposition summary.
  - Their TE period, treatment period and concomitant medication period will end at the cut-off date.
  - Their treatment duration will be derived by considering date of cut-off as last IMP date.

3.9 CHANGES TO PROTOCOL-PLANNED ANALYSES

Not applicable.

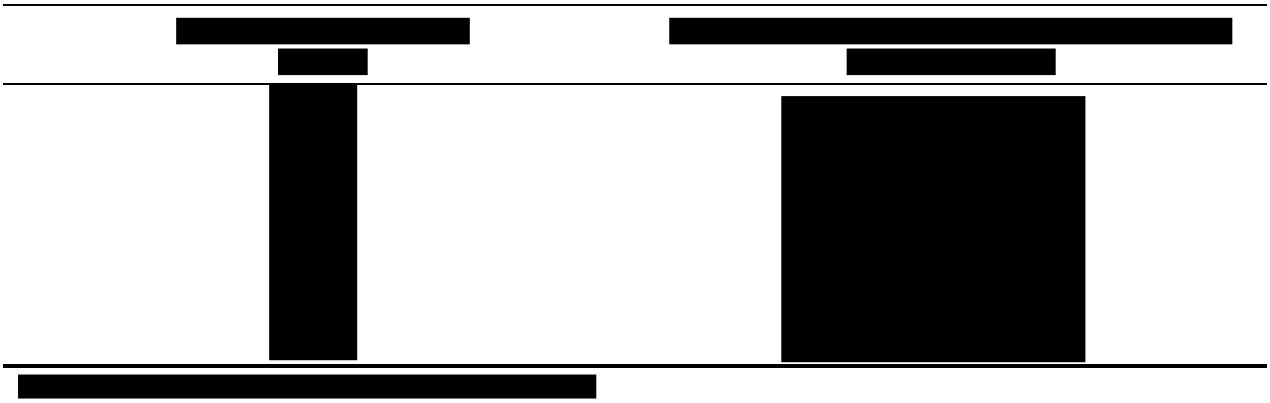
4 SAMPLE SIZE DETERMINATION

Assuming a pre-screening failure rate of 84% and a study screening failure rate of 15%, approximately 285 participants will be pre-screened to achieve approximately 38 treated participants in the study.

The initial plan is to treat a total of 38 participants evaluable for anti-tumor activity (at least 1 postbaseline tumor assessment, early clinical progression, or death due to disease progression).

Estimated ORR and 95% exact CIs by number of responders from a sample size of 38 evaluable participants for anti-tumor activity are listed in Table 7.

Table 7 - [redacted]



5 SUPPORTING DOCUMENTATION

5.1 APPENDIX 1 LIST OF ABBREVIATIONS

ADI:	actual dose intensity
AE:	adverse event
AESIs:	adverse events of special interest
ALP:	alkaline phosphatase
ALT:	alanine aminotransferase
ATA:	anti therapeutic antibody
ATC:	anatomical therapeutical chemical
BCVA:	best corrected visual acuity
BOR:	best overall response
BSA:	body surface area
BUN:	blood urea nitrogen
CEA:	carcinoembryonic agent
CEACAM5:	carcinoembryonic antigen-related cell adhesion molecule 5
Ceoi:	concentration at end of infusion
CI:	confidence interval
CMQ:	customized MedDRA queries
COVID-19:	coronovirus disease 2019
CR:	complete response
Ctrough:	through concentration
DOR:	duration of response
ECG:	electrocardiogram
ECOG:	Eastern Cooperative Oncology Group
eCRF:	electronic Case Report Form
FACT-GP5:	functional assessments of cancer therapy item GP-5
HLGT:	high-level group term
HLT:	high-level term
ICI:	immune checkpoint inhibitor
IMP:	investigational medicinal product
INR:	international normalized ratio
LLN:	lower limit of normal
LLOQ:	lower limit of quantification
LLT:	lower-level term
MedDRA:	Medical Dictionary for Regulatory Activities
NCI:	National Cancer Institute
NCI-CTCAE:	National cancer institute common terminology for adverse events
NSCLC-SAQ:	non-small-cell lung cancer symptom assessment questionnaire
NSQ NSCLC:	non-squamous non small cell lung cancer
ORR:	objective response rate
PCSA:	potentially clinically significant abnormality
PD:	progressive disease
PDI:	planned dose intensity

PFS:	progression-free survival
PGIC-LCS:	patient global impression of change- lung cancer symptom
PGIS-LCS:	patient global impression of severity- lung cancer symptom
PK:	pharmacokinetic
PR:	partial response
PRO:	patient-reported outcomes
PS:	performance status
PT:	preferred term
QTcF:	Fredericia's correction formuka
RBC:	red blood cells
RECIST:	response evaluation criteria in solid tumors
SAE:	serious adverse event
SAP:	Statistical Analysis Plan
SD:	stable disease
SMQ:	standardized MedDRA queries
SOC:	system organ class
TEAE:	treatment-emergent adverse event
ULN:	upper limit of normal
ULOQ:	upper limit of quantification
WHO-DD:	World Health Organization-drug dictionary

5.2 APPENDIX 2 PARTICIPANT DISPOSITIONS

The number (%) of participants included in each of the analysis populations listed in will be summarized. Reasons for exclusion from the population without trial impact (disruption) due to COVID-19 will be summarized.

Pre-screen failures are defined as participants who consent to participate in the pre-screening phase of the study but are not subsequently screened. The number (%) of pre-screen failures and reasons for pre-screen failures will be provided in the pre-screened population.

Screen failures are defined as participants who consent to participate in the study but are not subsequently enrolled. The number (%) of screen failures and reasons for screen failures will be provided in the screened population.

The number (%) of participants in the following categories will be provided:

- Enrolled participants
- Enrolled but not exposed participants
- Enrolled and exposed participants
- Participants who discontinued the study intervention and main reason for permanent intervention discontinuation
- Participants still on study intervention

- Participants who discontinued the study period and main reason for study discontinuation

Reasons for permanent study intervention and study discontinuation "adverse event" and "other reasons" will be split as related versus not related to COVID-19, if applicable.

The number (%) of exposed and not enrolled participants will also be summarized.

In addition, the number (%) of participants pre-screened, pre-screen failed, screened, screen-failed, enrolled, with permanent intervention discontinuation and with early study discontinuation will be provided by country and site.

For all categories of participants (except for the pre-screened, screened and non-enrolled categories) percentages will be calculated using the number of participants in the all-treated population as the denominator.

Protocol deviations

Critical and major protocol deviations (automatic or manual) will be summarized in the all-treated population.

5.3 APPENDIX 3 DEMOGRAPHICS AND BASELINE CHARACTERISTICS, PRIOR OR CONCOMITANT MEDICATIONS

Demographics, baseline characteristics, medical surgical history

The following demographics and baseline characteristics, medical and surgical history and disease characteristics at baseline will be summarized using descriptive statistics in the all-treated population.

Demographic and baseline characteristics

- age in years as quantitative variable and in categories (<65, 65 to <75, ≥75)
- gender (Male, Female)
- race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Not reported, Unknown)
- ethnicity (Hispanic or Latino, not Hispanic or Latino, Not Reported, Unknown)
- ECOG PS (0,1)
- Weight in kg as a quantitative variable
- BSA in m^2 as quantitative variable

Baseline safety and efficacy parameters (apart from those listed above) will be presented along with the safety and efficacy summaries.

Medical (or surgical) history includes relevant history of previous pathologies and surgeries. Medical and surgical history will be coded to a LLT, PT, HLT, HLGT, and associated primary SOC using the MedDRA version currently in effect at Sanofi at the time of database lock.

Specific disease history includes diagnosis, histopathology types, stage at diagnosis, disease extent at study entry, time from initial diagnosis to first administration of IMP (in months).

Specific disease status at study entry includes extent of disease, number and type of organs involved (including primary tumor location), CEACAM5 expression (<1%, 1-49%), PD-L1 expression (<1%, 1 – 49% and ≥ 50%), circulating CEA, smoking status (current, former, never) and smoking habits (in pack-years). Type of tumor biopsy (archival /fresh) will also be described.

Prior or concomitant medications

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version currently in effect at Sanofi at the time of database lock.

- Prior medications are those the participant received prior to first IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment period.
- Concomitant medications are any medications received by the participant concomitantly to the IMP from the first administration of IMP to the last IMP intake + 30 days.
- Post-treatment medications are those the participant received in the period running from the end of the concomitant medications period up to the end of the study.
- A given medication can be classified as a prior medication and/or as a concomitant medication and/or as post-treatment medication. If it cannot be determined whether a given medication was taken prior or concomitantly or post, it will be considered as prior, concomitant, and post-treatment medication.

The prior and concomitant medications will be summarized for all-treated population, by anatomic and therapeutic level. Participants will be counted once in each ATC category (anatomic or therapeutic) linked to the medication.

Anticancer therapies

Prior anticancer therapies include chemotherapy, surgery and radiotherapy.

- Prior anti-cancer therapies:
  - Number (%) of participants with prior anti-cancer therapies including neoadjuvant/adjuvant/radiosensitizer (ie, concurrent chemoradiotherapy with or without maintenance treatment) and advanced regimen,
  - Number of participants with intent: neoadjuvant and adjuvant and chemoradiotherapy and advanced, neoadjuvant and/or adjuvant and advanced, chemoradiotherapy and advanced, neoadjuvant or adjuvant or chemoradiotherapy only, advanced only,
  - Number of prior anti-cancer therapies in the advanced setting,

  - A regimen in the advanced setting consists of a single agent, combination or a sequential therapeutic strategy with several drugs, given until a PD is documented. Participant who had adjuvant /neoadjuvant treatment and relapsed as metastatic disease during or within 6 months of treatment will be considered as first line treatment in the advanced setting,
- Type of prior treatment including neoadjuvant/adjuvant/chemoradiotherapy and advanced regimen
  - Biologics/Small molecules:
    - Tyrosine kinase inhibitor: EGFR inhibitor, Anti-angiogenic (Bevacizumab…), BRAF Kinase inhibitors, ALK inhibitors, RAS/RAF/MEK/ERK signaling pathway inhibitors
    - ICI: anti PD1/PDL1, CTL4A inhibitors
    - Others
  - Chemotherapy: Platinum, Pemetrexed, Taxanes, Vinca alcaloids…, Others
  - Antibody Drug Conjugate
  - Others
- Prior Immune checkpoint inhibitor (anti PD1/PDL1) as per eCRF: sequential vs in combination with chemotherapy,
- Summary of last prior anti-cancer therapy:
  - Time from completion of last regimen of treatment to randomization (months)
  - Main treatments of last regimen
    - ICI monotherapy
    - ICI in combination with chemotherapy
    - ICI and other biologics and small molecules
    - No ICI: Chemotherapy, Biologic, Chemotherapy and biologics
- Reason for discontinuation of the last line
- Best response to the last line

- Prior surgery: number (%) of participants with any prior surgery related to lung cancer, type of surgery and time from the last surgery to first administration of IMP (in months).
- Prior radiotherapy: number (%) of participants with any prior radiotherapy related to lung cancer, intent, analgesic intent if palliative and time from last dose of radiotherapy to first administration of IMP (in months).

Further anticancer therapies (including systemic anticancer therapies, surgeries and radiotherapies) after discontinuation of study intervention will be summarized based on WHO-DD coding.

5.4 APPENDIX 4 DATA HANDLING CONVENTIONS

Unscheduled visits

Unscheduled visit measurements of laboratory data, vital signs, ECG, physical examination, ophthalmological examination and ATA will be used for computation of baseline, the last on-treatment value, analysis according to PCSAs and/or NCI grades, and the shift summaries for safety.

Unscheduled tumor assessments will be used for computation of efficacy endpoints based on radiological assessments of tumor burden.

5.4.1 General conventions

The following formulas will be used for computation of parameters.

Time unit

A month length is 30.4375 days (365.25 / 12). If duration is to be reported in months, duration in days is divided by 30.4375. If duration is to be reported in years, duration in days will be divided by 365.25.

Demographic formula

Body surface area value will be derived using the variation of DuBois and DuBois formula:

$$BSA = 0.0007184 \times Weight(kg)^{0.425} \times Height(cm)^{0.725}$$

5.5 GENERIC RANGES FOR HEMATOLOGICAL AND BIOCHEMISTRY PARAMETERS

The current list of generic ranges for hematological parameters (for adults) is provided in the table below:

Table 8 - Generic ranges for hematological parameters

LBTESTCD	LBTEST	GENDER	LBSTRESU	LBGNNRLO- LBGNNRHI
HGB	Hemoglobin	F	g/L	120-160
HGB	Hemoglobin	M	g/L	135-175
LYM	Lymphocytes		10^9/L	1-2
NEUT	Neutrophils		10^9/L	1.8-3.15
PLAT	Platelets		10^9/L	150-350
WBC	Leukocytes		10^9/L	4.5-11
EOS	Eosinophils		10^9/L	0-0.4
BASO	Basophils		10^9/L	0-0.15
MONO	Monocytes		10^9/L	0.18-0.5
HCT	Hematocrit	F	%	0.36-0.46
HCT	Hematocrit	M	%	0.41-0.53
RBC	Erythrocytes	F	10^12/L	4-5.2
RBC	Erythrocytes	M	10^12/L	4.5-5.9
INR	INR		Ratio	0.8-1.2

Based on NEJM (N Engl J Med 2004;351:1548-63.): “Laboratory Reference Values”. Alexander Kratz. M.D.. Ph.D.. M.P.H.. Maryjane Ferraro. Ph.D.. M.P.H.. Patrick M. Sluss. Ph.D.. and Kent B. Lewandrowski. M.D.

The current list of generic ranges for biochemistry parameters (for adults) is provided in the table below:

Table 9 - Generic ranges for biochemistry parameters

LBTEST	LBSTRESU	LBGNNRLO-LBGNNRHI
Albumin	g/L	35-55
Blood Urea Nitrogen (BUN)	mmol/L	3.6-7.1
Corrected Calcium	mmol/L	2.2-2.6
Chloride	mmol/L	80-115
Glucose	mmol/L	3.9-7
Bicarbonate (HCO3)	mmol/L	22-29
Potassium	mmol/L	3.5-5
Magnesium	mmol/L	0.8-1.2
Sodium	mmol/L	136-145
Phosphate	mmol/L	1-1.4
Protein	g/L	55-80
Urea Nitrogen	mmol/L	3.6-7.1

6 REFERENCES

1. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000;92(3):205-16.

2. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

3. Liepa A, Bushnell D, Atkinson T, Debusk K, Mccarrier K, Martin M, et al. P1. 06-004 Evaluating the Non-Small Cell Lung Cancer Symptom Assessment Questionnaire (NSCLCSAQ): Preliminary Results from the Quantitative Pilot Study: Topic: Advanced General. J Thorac Oncol. 2017;12(1):S666-7.

4. FDA. Clinical Outcome Assessment (COA) Qualification Program DDT COA #000009: NonSmall Cell Lung Cancer Symptom Assessment Questionnaire (NSCLC-SAQ) Full Qualification Package [Online]. Updated 2018 Oct 12 [Retrieved 2021 Aug 05]. Available from: URL:https://www.fda.gov/media/119251/download.

5. Williams P, Burke T, Norquist JM, Daskalopoulou C, Speck RM, Samkari A, et al. NSCLC symptom assessment questionnaire: measurement properties and estimated clinically meaningful thresholds from a phase 3 study. JTO Clin Res Rep. 2022;3(4):100298.

6. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. J Clin Oncol. 1993;11:570-9.

7. Brucker PS, Yost K, Cashy J, Webster K, Cella D. General population and cancer patient norms for the Functional Assessment of Cancer Therapy-General (FACT-G). Eval Health Prof. 2005;28(2):192-211.

8. Akpek EK, Amescua G, Farid M, Garcia-Ferrer FJ, Lin A, Rhee MK, et al. Dry Eye Syndrome Preferred Practice Pattern®. Ophthalmology. 2019;126(1):P286-334.

Signature Page for VV-CLIN-0644910 v1.0
act17241-16-1-9-sap

Approve & eSign	[redacted] Clinical [redacted]

Approve & eSign	[redacted] Clinical [redacted]