NOVARTIS

Clinical Development

ACZ885 (Canakinumab)

CACZ885T2301 / NCT03447769

A phase III, multicenter, randomized, double blind, placebo-controlled study evaluating the efficacy and safety of canakinumab versus placebo as adjuvant therapy in adult subjects with stages AJCC/UICC v. 8 II-IIIA and IIIB (T>5cm N2) completely resected (R0) non-small cell lung cancer (NSCLC)

Statistical Analysis Plan (SAP) for close-out CSR

Author:	Trial Statistician [redacted]
Document type:	SAP Documentation
Document status:	Final
Release date:	03-Mar-2023
Number of pages:	10

Template Document History – Changes compared to previous version

Date	Version number	Summary of changes
03-Mar-2023	1.0	N/A

Table of contents

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the close-out clinical study report (CSR) of study CACZ885T2301, a phase III, multicenter, randomized, double blind, placebo-controlled study evaluating the efficacy and safety of canakinumab versus placebo as adjuvant therapy in adult subjects with stages American Joint Committee on Cancer (AJCC)/Union for International Cancer Control (UICC) v8 II-IIIA and IIIB (T > 5cm N2) completely resected (R0) non-small cell lung cancer (NSCLC).

The statistical methods for the study are documented in full in the SAP for the primary analysis CSR [CACZ885T2301 SAP Amendment 3 dated 21-Mar-2022]. The results of the primary and secondary endpoint data collected at the primary analysis cut-off date are described in the [CACZ885T2301 Primary Analysis CSR].

The close-out CSR will present the results of cumulative study data for secondary endpoints for which additional data were collected between the primary analysis cut-off date and the LPLV date. Thus the outputs planned for the close-out CSR will be a subset of the outputs planned for the primary analysis CSR that need to be updated after the primary analysis data cut-off date. The list of outputs planned for the close-out CSR is specified in Section 2.

The primary endpoint was not met for this study as described in the [CACZ885T2301 Primary Analysis CSR]; therefore, no formal statistical analysis will be performed for this close-out CSR. The outputs for the close-out CSR consist mainly of summary tables prepared using descriptive statistics as well as data listings.

2 Planned tables, figures and listings

The table below provides the list of outputs that need to be generated, with additional guidelines added under 'Notes'. Please follow the in-text layout given below, post-text output has been added for reference.

Add footnote: CACZ885T2301- *Cut-off date*: 07Feb2023

Output Number in close-out CSR	Output title in close-out CSR	Output number in primary CSR	Output title in primary CSR	Notes
Table 2-1	Subject disposition (Full analysis set)	Table 10-1	Subject disposition (Full analysis set)	
Table 2-2	Protocol deviation (≥ 0.4% for any treatment arm) (Full analysis set)	Table 10-2	Protocol deviation (≥ 0.4% for any treatment arm) (Full analysis set)	
Table 2-3	Duration of exposure to study drug (Safety set)	Table 10-7	Duration of exposure to study drug (Safety set)	
Table 2-4	Dose of study drug received (Safety set)	Table 10-8	Dose of study drug received (Safety set)	
Table 2-5	Dose interruptions and discontinuation of study drug (Safety set)	Table 10-9	Dose interruptions and discontinuation of study drug (Safety set)	
Table 2-6	Summary of overall survival by study arm (Full analysis set)	Table 11-3	Summary of overall survival by study arm (Full analysis set)	

Table 2-6a	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1%) (Full analysis set)	Table 11-3a	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1%) (Full analysis set)	
Table 2-6b	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1% and < 49% vs. >= 50%) (Full analysis set)	Table 11-3b	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1% and < 49% vs. >= 50%) (Full analysis set)	
Table 2-6c	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 50 % vs. >= 50%) (Full analysis set)	Table 11-3c	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 50 % vs. >= 50%) (Full analysis set)	
Table 2-6d	Summary of overall survival by study arm and median of baseline CD8 subgroup (Full analysis set)	Table 11-3d	Summary of overall survival by study arm and median of baseline CD8 subgroup (Full analysis set)	
Table 2-6e	Summary of overall survival by study arm and quartile of baseline CD8 subgroup (Full analysis set)	Table 11-3e	Summary of overall survival by study arm and quartile of baseline CD8 subgroup (Full analysis set)	
Table 2-7	Overview of adverse events (Safety set)	Table 12-1	Overview of adverse events (Safety set)	
Table 2-8	Adverse events by system organ class (Safety set)	Table 12-2	Adverse events by system organ class (Safety set)	
Table 2-9	On-treatment deaths by system organ class and preferred term (Safety set)	Table 12-3	On-treatment deaths by system organ class and preferred term (Safety set)	
Table 2-10	Most frequently reported serious adverse events by preferred term (≥2.0% all grades, both treatment arms) (Safety set)	Table 12-4	Most frequently reported serious adverse events by preferred term (≥2.0% all grades, both treatment arms) (Safety set)	
Table 2-11	Most frequently reported adverse events leading to dose interruption, by preferred term (≥2.0% all grades, both treatment arms) (Safety set)	Table 12-5	Most frequently reported adverse events leading to dose interruption, by preferred term (≥2.0% all grades, both treatment arms) (Safety set)	
Table 2-12	Overview of adverse events of special interest (Safety set)	Table 12-6	Overview of adverse events of special interest (Safety set)	
Table 2-13	Incidence of adverse events of special interest – Infections (Safety set)	Table 12-7	Incidence of adverse events of special interest – Infections (Safety set)	
Table 2-14	Incidence of adverse events of special interest – Opportunistic infections (Safety set)	Table 12-8	Incidence of adverse events of special interest – Opportunistic infections (Safety set)	
Table 2-15	Incidence of adverse events of special interest – Abnormal liver parameters (Safety set)	Table 12-9	Incidence of adverse events of special interest – Abnormal liver parameters (Safety set)	
Table 2-16	Incidence of adverse events of special interest – Neutropenia (Safety set)	Table 12-10	Incidence of adverse events of special interest – Neutropenia (Safety set)	

Table 2-17	Incidence of adverse events of special interest – Thrombocytopenia (Safety set)	Table 12-11	Incidence of adverse events of special interest – Thrombocytopenia (Safety set)	
Table 2-18	Incidence of adverse events of special interest – Pulmonary complications, pulmonary hypertension and interstitial lung disease (Safety set)	Table 12-12	Incidence of adverse events of special interest – Pulmonary complications, pulmonary hypertension and interstitial lung disease (Safety set)	source: Table 14.3.1-5.2
Table 2-19	Worst post-baseline hematology abnormalities based on CTC grades (Safety set)	Table 12-13	Worst post-baseline hematology abnormalities based on CTC grades (Safety set)	
Table 2-20	Worst post-baseline clinical chemistry abnormalities based on CTC grades (Safety set)	Table 12-14	Worst post-baseline clinical chemistry abnormalities based on CTC grades (Safety set)	
Table 2-21	Categorical analysis of hepatic laboratory values (Safety set)	Table 12-15	Categorical analysis of hepatic laboratory values (Safety set)	
Table 2-22	Notable vital sign values (Safety set)	Table 12-16	Notable vital sign values (Safety set)	
Table 2-23	Notable ECG values (Safety set)	Table 12-17	Notable ECG values (Safety set)	
Table 2-24	Canakinumab ADA and NAB incidence (Safety set)	Table 12-18	Canakinumab ADA and NAB incidence (Safety set)	
Figure 2-1	Arithmetic mean (SD) and geometric mean concentration-time profiles for canakinumab (linear view, Pharmacokinetic analysis set)	Figure 11-5	Arithmetic mean (SD) and geometric mean concentration-time profiles for canakinumab (linear view, Pharmacokinetic analysis set)	
Table 14.1-1.1	Subject disposition (Full analysis set)	Table 14.1-1.1	Subject disposition (Full analysis set)	
Table 14.1-1.5	Protocol deviations (Full analysis set)	Table 14.1-1.5	Protocol deviations (Full analysis set)	
Table 14.2-2.1	Summary of overall survival by study arm (Full analysis set)	Table 14.2-2.1	Summary of overall survival by study arm (Full analysis set)	
Table 14.2-2.1a	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1%) (Full analysis set)	Table 14.2-2.1a	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1%) (Full analysis set)	
Table 14.2-2.1b	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1% and < 49% vs. >= 50%) (Full analysis set)	Table 14.2-2.1b	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 1% vs. >= 1% and < 49% vs. >= 50%) (Full analysis set)	
Table 14.2-2.1c	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 50 % vs. >= 50%) (Full analysis set)	Table 14.2-2.1c	Summary of overall survival by study arm and baseline PD-L1 subgroup (Baseline PD-L1: < 50 % vs. >= 50%) (Full analysis set)	
Table 14.2-2.1d	Summary of overall survival by study arm and median of baseline CD8 subgroup (Full analysis set)	Table 14.2-2.1d	Summary of overall survival by study arm and median of baseline CD8 subgroup (Full analysis set)	

Table 14.2-2.1e	Summary of overall survival by study arm and quartile of baseline CD8 subgroup (Full analysis set)	Table 14.2-2.1e	Summary of overall survival by study arm and quartile of baseline CD8 subgroup (Full analysis set)	
Table 14.2-3.1	Summary of lung cancer specific survival by study arm (Full analysis set)	Table 14.2-3.1	Summary of lung cancer specific survival by study arm (Full analysis set)	
Table 14.2-4.1	Canakinumab serum concentrations (Pharmacokinetic analysis set)	Table 14.2-4.1	Canakinumab serum concentrations (Pharmacokinetic analysis set)	
Table 14.2-4.1a	Canakinumab serum concentrations by baseline ADA (Pharmacokinetic analysis set)	Table 14.2-4.1a	Canakinumab serum concentrations by baseline ADA (Pharmacokinetic analysis set)	
Table 14.2-4.1b	Canakinumab serum concentrations by on-treatment ADA (Pharmacokinetic analysis set)	Table 14.2-4.1b	Canakinumab serum concentrations by on-treatment ADA (Pharmacokinetic analysis set)	
Table 14.2-4.1c	Canakinumab serum concentrations by treatment-induced ADA (Pharmacokinetic analysis set)	Table 14.2-4.1c	Canakinumab serum concentrations by treatment-induced ADA (Pharmacokinetic analysis set)	
Table 14.2-4.1d	Canakinumab serum concentrations by treatment boosted ADA (Pharmacokinetic analysis set)	Table 14.2-4.1d	Canakinumab serum concentrations by treatment boosted ADA (Pharmacokinetic analysis set)	
Table 14.2-4.2b	Summary of overall survival by the quartile of average steady state evaluable Ctrough concentration (Pharmacokinetic analysis set)	Table 14.2-4.2b	Summary of overall survival by the quartile of average steady state evaluable Ctrough concentration (Pharmacokinetic analysis set)	
Table 14.2-5.1	Summary of EORTC QLQ-C30 Scores by time point and treatment arm (Full analysis set)	Table 14.2-5.1	Summary of EORTC QLQ-C30 Scores by time point and treatment arm (Full analysis set)	
Table 14.2-5.2	Summary of EORTC QLQ-LC13 Scores by time point and treatment arm (Full analysis set)	Table 14.2-5.2	Summary of EORTC QLQ-LC13 Scores by time point and treatment arm (Full analysis set)	
Table 14.2-5.3	Summary of EQ-5D-5L Scores by time point and treatment arm (Full analysis set)	Table 14.2-5.3	Summary of EQ-5D-5L Scores by time point and treatment arm (Full analysis set)	
Table 14.2-6.2	Canakinumab ADA and NAB incidence (Safety set)	Table 14.2-6.2	Canakinumab ADA and NAB incidence (Safety set)	
Table 14.2-7.1	Shift table of subjects with positive or negative anti-canakinumab antibodies overall and by visit (Full analysis set)	Table 14.2-7.1	Shift table of subjects with positive or negative anti-canakinumab antibodies overall and by visit (Full analysis set)	
Table 14.3-1.1	Duration of exposure to study drug (Safety set)	Table 14.3-1.1	Duration of exposure to study drug (Safety set)	
Table 14.3-1.2	Dose of study drug received (Safety set)	Table 14.3-1.2	Dose of study drug received (Safety set)	

Table 14.3-1.3	Interruptions and discontinuation of study drug (Safety set)	Table 14.3-1.3	Interruptions and discontinuation of study drug (Safety set)	
Table 14.3-3.5	Hematology laboratory results by time point (Safety set)	Table 14.3-3.5	Hematology laboratory results by time point (Safety set)	
Table 14.3-3.6	Biochemistry laboratory results by time point (Safety set)	Table 14.3-3.6	Biochemistry laboratory results by time point (Safety set)	
Table 14.3-3.7	Categorical analysis of hepatic laboratory values (Safety set)	Table 14.3-3.7	Categorical analysis of hepatic laboratory values (Safety set)	
Table 14.3-4.1	Notable vital sign values (Safety set)	Table 14.3-4.1	Notable vital sign values (Safety set)	
Table 14.3-5.1	Notable ECG values (Safety set)	Table 14.3-5.1	Notable ECG values (Safety set)	
Table 14.3.1-1.1	Adverse events by system organ class and preferred term (Safety set)	Table 14.3.1-1.1	Adverse events by system organ class and preferred term (Safety set)	
Table 14.3.1-1.2	Treatment-related adverse events by system organ class and preferred term (Safety set)	Table 14.3.1-1.2	Treatment-related adverse events by system organ class and preferred term (Safety set)	
Table 14.3.1-1.3	Adverse events leading to treatment discontinuation, by system organ class and preferred term (Safety set)	Table 14.3.1-1.3	Adverse events leading to treatment discontinuation, by system organ class and preferred term (Safety set)	
Table 14.3.1-1.4	Adverse events leading to dose interruption, by system organ class and preferred term (Safety set)	Table 14.3.1-1.4	Adverse events leading to dose interruption, by system organ class and preferred term (Safety set)	
Table 14.3.1-1.5	Adverse events requiring additional therapies, by system organ class and preferred term (Safety set)	Table 14.3.1-1.5	Adverse events requiring additional therapies, by system organ class and preferred term (Safety set)	
Table 14.3.1-1.6	Adverse events leading to fatal outcome, by system organ class and preferred term (Safety set)	Table 14.3.1-1.6	Adverse events leading to fatal outcome, by system organ class and preferred term (Safety set)	
Table 14.3.1-2.1	Non-serious adverse events (threshold=5%) by system organ class and preferred term (Safety set)	Table 14.3.1-2.1	Non-serious adverse events (threshold=5%) by system organ class and preferred term (Safety set)	
Table 14.3.1-3.1	On-treatment deaths and serious adverse events by system organ class and preferred term (Safety set)	Table 14.3.1-3.1	On-treatment deaths and serious adverse events by system organ class and preferred term (Safety set)	
Table 14.3.1-3.2	On-treatment deaths, by system organ class and preferred term (Safety set)	Table 14.3.1-3.2	On-treatment deaths, by system organ class and preferred term (Safety set)	
Table 14.3.1-3.3	All deaths, by system organ class and preferred term (Safety set)	Table 14.3.1-3.3	All deaths, by system organ class and preferred term (Safety set)	
Table 14.3.1-4.1	Serious adverse events by system organ class and preferred term (Safety set)	Table 14.3.1-4.1	Serious adverse events by system organ class and preferred term (Safety set)	
Table 14.3.1-4.1a	Serious adverse events by system organ class, preferred	Table 14.3.1-4.1a	Serious adverse events by system organ class, preferred	

	term and maximum grade (Safety set)		term and maximum grade (Safety set)	
Table 14.3.1-4.2	Treatment-related serious adverse events by system organ class and preferred term (Safety set)	Table 14.3.1-4.2	Treatment-related serious adverse events by system organ class and preferred term (Safety set)	
Table 14.3.1-4.2a	Treatment-related serious adverse events by system organ class, preferred term and maximum grade (Safety set)	Table 14.3.1-4.2a	Treatment-related serious adverse events by system organ class, preferred term and maximum grade (Safety set)	
Table 14.3.1-5.1	Overview of adverse events of special interest (Safety set)	Table 14.3.1-5.1	Overview of adverse events of special interest (Safety set)	
Table 14.3.1-5.2	Incidence of adverse event of special interest (Safety set)	Table 14.3.1-5.2	Incidence of adverse event of special interest (Safety set)	
Table 14.3.1-5.3	Incidence of adverse event of special interest including preferred term level (Safety set)	Table 14.3.1-5.3	Incidence of adverse event of special interest including preferred term level (Safety set)	
Figure 2-1	Kaplan-Meier plot of overall survival by treatment arm (Full Analysis Set)	Figure 11-2	Kaplan-Meier plot of overall survival by treatment arm (Full Analysis Set)	
Figure 2-2	Forest plot of hazard ratio with 95% confidence interval for overall survival from subgroup analysis (Full analysis set)	Figure 11-3	Forest plot of hazard ratio with 95% confidence interval for overall survival from subgroup analysis (Full analysis set)	
Figure 14.2-2.1	Kaplan-Meier plot of overall survival by treatment arm (Full analysis set)	Figure 14.2-2.1	Kaplan-Meier plot of overall survival by treatment arm (Full analysis set)	
Figure 14.2-2.2	Forest plot of hazard ratio with 95% confidence interval for overall survival from subgroup analysis (Full analysis set)	Figure 14.2-2.2	Forest plot of hazard ratio with 95% confidence interval for overall survival from subgroup analysis (Full analysis set)	
Figure 14.2-4.2	Arithmetic mean (SD) and geometric mean concentration-time profiles for canakinumab (Pharmacokinetic analysis set)	Figure 14.2-4.2	Arithmetic mean (SD) and geometric mean concentration-time profiles for canakinumab (Pharmacokinetic analysis set)	
Listing 16.2.1-1.2	Subject disposition (Full analysis set)	Listing 16.2.1-1.2	Subject disposition (Full analysis set)	
Listing 16.2.2-1.1	Protocol deviations (Full analysis set)	Listing 16.2.2-1.1	Protocol deviations (Full analysis set)	
Listing 16.2.7-1.1	Adverse events (Safety set)	Listing 16.2.7-1.1	Adverse events (Safety set)	
Listing 16.2.8-1.1	Subject laboratory profile: Hematology (Safety set)	Listing 16.2.8-1.1	Subject laboratory profile: Hematology (Safety set)	
Listing 16.2.8-1.2	Subject laboratory profile: Biochemistry - Liver (Safety set)	Listing 16.2.8-1.2	Subject laboratory profile: Biochemistry - Liver (Safety set)	
Listing 16.2.8-1.3	Subject laboratory profile: Biochemistry - Renal (Safety set)	Listing 16.2.8-1.3	Subject laboratory profile: Biochemistry - Renal (Safety set)	
Listing 16.2.8-1.4	Subject laboratory profile: Biochemistry - Other (Safety set)	Listing 16.2.8-1.4	Subject laboratory profile: Biochemistry - Other (Safety set)	

Listing 16.2.8-2.1	Canakinumab ADA sample status – subjects treated with Canakinumab (Safety set)	Listing 16.2.8-2.1	Canakinumab ADA sample status – subjects treated with Canakinumab (Safety set)	
Listing 16.2.8-2.2	Canakinumab ADA samples status for subjects with neutralizing antibodies (NAB) testing results – subjects treated with Canakinumab (Safety set)	Listing 16.2.8-2.2	Canakinumab ADA samples status for subjects with neutralizing antibodies (NAB) testing results – subjects treated with Canakinumab (Safety set)	
Listing 16.2.9-1.1	Vital signs (Safety set)	Listing 16.2.9-1.1	Vital signs (Safety set)	
Listing 16.2.9-1.2	ECG values (Safety set)	Listing 16.2.9-1.2	ECG values (Safety set)	
Listing 14.2-1.2	Overall survival (Full analysis set)	Listing 14.2-1.2	Overall survival (Full analysis set)	
Listing 14.2-1.3	Lung cancer specific survival (Full analysis set)	Listing 14.2-1.3	Lung cancer specific survival (Full analysis set)	
Listing 14.2-1.4	Serum PK concentrations for canakinumab (Full analysis set)	Listing 14.2-1.4	Serum PK concentrations for canakinumab (Full analysis set)	
Listing 14.3.2-1.1	Deaths (Safety set)	Listing 14.3.2-1.1	Deaths (Safety set)	
Listing 14.3.2-1.3	Serious adverse events (Safety set)	Listing 14.3.2-1.3	Serious adverse events (Safety set)	
Listing 14.3.2-1.4	Adverse events leading to discontinuation (Safety set)	Listing 14.3.2-1.4	Adverse events leading to discontinuation (Safety set)	
Listing 14.3.2-1.5	Adverse events leading to dose interruptions (Safety set)	Listing 14.3.2-1.5	Adverse events leading to dose interruptions (Safety set)	
Listing 14.3.2-1.6	Adverse events of special interest (Safety set)	Listing 14.3.2-1.6	Adverse events of special interest (Safety set)	
Listing 14.3.4-1.1	CTC grade 3 or 4 laboratory values (Safety set)	Listing 14.3.4-1.1	CTC grade 3 or 4 laboratory values (Safety set)	
Listing 14.3.4-1.2	Patients with a post-baseline TBILI > 2xULN, or ALT> 3xULN or AST > 3xULN (Safety set)	Listing 14.3.4-1.2	Patients with a post-baseline TBILI > 2xULN, or ALT> 3xULN or AST > 3xULN (Safety set)	

NOVARTIS

Clinical Development

ACZ885/Canakinumab

CACZ885T2301 / NCT03447769

A phase III, multicenter, randomized, double blind, placebo-controlled study evaluating the efficacy and safety of canakinumab versus placebo as adjuvant therapy in adult subjects with stages AJCC/UICC v. 8 II-IIIA and IIIB (T>5cm N2) completely resected (R0) non-small cell lung cancer (NSCLC)

Statistical Analysis Plan (SAP)

Author:	Trial statistician, [redacted]

Document type:	SAP Documentation
Document status:	Amendment 03 Final
Release date:	21-Mar-2022
Number of pages:	57

Document History – Changes compared to previous final version of SAP

Date	Section	Changes
14-Mar-2018		N/A- First version
10-Mar-2021	Section 1	Updated according to the latest protocol amendment version 02 prior to DFS futility analysis
	Section 1.1	Added 'The interim analysis for efficacy will only be carried out after all subjects have been randomized and have at least one post-baseline efficacy assessment if not withdrawn early'. Updated standard care: ' at a minimum of two cycles of cisplatin are required' and 'radiation therapy is allowed if indicated as per local guidance or practice'
	Section 1.2	Removed approximate timings of DFS and OS analysis as they were no longer accurate given slower than expected enrollment rate. This update applied throughout the document
	Section 2.1.1	Removed 'End of treatment date' from table 2-2
	Section 2.2	Removed Per protocol set (PPS) in line with ICH E9 (R1) Updated the PAS definition: 'PAS includes all subjects who receive at least one dose of canakinumab'
	Section 2.3	Added summary table of demographic and baseline characteristics by pandemic phase Added table 2-4: COVID-19 start and end dates by region/country Updated wording 'Medical history and current medical conditions, including cancer-related conditions and symptoms entered on (e) CRF will be summarized and listed by treatment arm'
	Section 2.3.1	Updated wording of 'discontinued post treatment follow-up' to 'discontinued study'
	Section 2.3.2	In light of COVID-19 pandemic, added COVID-19 related protocol deviations
	Section 2.4.1	To clarify the dose interruption: a dose interruption reported in the eCRF after the last dose of study drug, won't be considered as a dose interruption Added 'First new NSCLC anti-neoplastic medication after discontinuation of study treatment will also be summarized'
	Section 2.5.1	Added 'New primary lung malignancies are treated as recurrence, clinical deterioration is not considered as recurrence' Added 'For subjects having the event confirmed by biopsy results, radiology date will be used as DFS event date' In light of ICH E9 addendum, added Primary estimand (DFS)
	Section 2.5.3	Rewrote section 2.5.3 by adding a censoring table 2-5 and new censoring rules: i) subjects having new antineoplastic therapy (ANP) for NSCLC will be censored at

Date	Section	Changes
		the last assessment date prior to initiation of ANP; ii) subjects having 'evidence of unresected NSCLC or distant metastatic NSCLC disease on screening evaluation' will be censored at randomization date Included 'New primary lung malignancies' as recurrence Added 'Clinical review will be performed to determine if anti-neoplastic therapy is related to NSCLC in case of multiple malignancies'
	Section 2.5.4	Restructured supportive analyses into two subsections: sensitivity analyses and supplementary analyses
	Section 2.5.4.1	Removed PPS sensitivity analysis in line with recommendations from ICH E9 (R1) Moved ANP related sensitivity analyses to supplementary analyses Added DFS sensitivity analysis with stratification factors from eCRF Added sensitivity analysis of censoring subjects with two or more times TAs missing
	Section 2.5.4.2	Clarified supplementary analyses related NSCLC ANP: i) DFS event after initiating new NSCLC anti-neoplastic therapy will be counted as event; ii) new NSCLC anti-neoplastic therapy prior to disease recurrence will be considered as DFS event Added supplementary analysis for DFS to assess the treatment effect had the pandemic not occurred. Hypothetical strategy was used to handle intercurrent events due to COVID-19.
	Section 2.5.4.2.1	Added subsection 2.5.4.2.1 'Medications treating suspected or confirmed COVID-19' Analysis of protocol deviations related to COVID-19 pandemic were added. Supplementary analyses were added for DFS and OS to assess the treatment effect had the pandemic not occurred. Hypothetical strategy was used to handle intercurrent events due to COVID-19.
	Section 2.6.1	In light of ICH E9 addendum, added key secondary estimand (OS) Added justification for key-secondary estimand
	Section 2.6.4	Restructured supportive analyses into two subsections: sensitivity analyses and supplementary analyses Added justification for the key second estimand
	Section 2.6.4.1	Removed PPS sensitivity analysis in line with recommendations from ICH E9 (R1)
	Section 2.6.4.2	Added supplementary analysis for OS to assess the treatment effect had the pandemic not occurred. Hypothetical strategy was used to handle intercurrent events due to COVID-19.
	Section 2.8.1.1	AESI groups in table 2-6 updated according to current eCRS.

Date	Section	Changes
		Removed 'Infections/Opportunistic infection' and separated them into 2 different AESIs.
	Section 2.8.3	CTCAE 5.0 will be used for all safety data summary and listing
		Updated liver function parameters for summary as per Novartis internal guidance
	Section 2.8.4.1	Rewrote ECG section to summary ECG abnormality for baseline and any post baseline assessments
		Added respiratory rate to table 2-7 clinically notable changes in vital signs
	Section 2.9	Removed 'along with other studies to generate post-hoc estimates of pharmacokinetic parameters using NONMEM' from population pharmacokinetic analysis
	Section 2.10.1	Removed landmark analysis
		Added summaries on 'treatment-induced and treatment-boosted' ADA and summary of ADA-positive Nab samples at baseline and on-treatment'
	[redacted]	[redacted]
	Section 5.1.1	Clarified the imputation for infusion date: 'no imputation of the start date of infusion or end date for infusion will be applied'
	Section 5.4.1	Rewrote section 5.4.1 primary analysis
	Section 5.5	Removed the SAS code for PRO mixed model
	[redacted]	[redacted]
	[redacted]	[redacted]
	Section 5.6	Specified the methods to calculate the time windows for one and two (D1 and D2) missing assessments
	Section 6	Added three new reference related to FDA, EMA and Novartis COVID-19 guidance
		Removed reference for calculating rates of multiple occurrences given the removal of analysis
07-Jan-2022	Rationale	The study planned to randomize approximately 1500 subjects to observe ~392 DFS events required for final DFS analysis. Since all planned subjects were not randomized and have not had at least one post-baseline efficacy assessment prior to efficacy interim analysis (as per protocol requirement), the planned efficacy interim analysis was not performed. The final DFS analysis is planned to be performed when approximately 392 DFS events are documented irrespective of the accrued sample size (even if less than 1500 subjects) to ensure the trial is not over-powered. At the time of the data cut-off for final DFS analysis, all subjects randomized prior to the data cut-off date for

Date	Section	Changes
		final DFS analysis who have had at least one post baseline recurrence assessment (i.e. all subjects have a minimal follow-up of 3 months) or discontinued study earlier will be included.
	Section 1	Specified that the efficacy interim analysis of DFS was not performed and the primary CSR will be based on the final DFS analysis
	Section 1.1	Specified that the efficacy interim analysis of DFS and first OS IA were not performed
	Section 2.1.1	According to the protocol version 02 visit schedules, Table 2-1 was modified to include schedules for subjects who discontinue study treatment and safety follow-up for all subjects
	Section 2.2.1	Due to EGFR mutant status (mutant vs. non-mutant) being assessed from *local* laboratories only, replaced "central testing" with "local testing"
	Section 2.5.2	Specified that the efficacy interim analysis of DFS was not performed
	Section 2.5.4.2	Specified all levels of baseline demographic and disease characteristics, including age, gender, race and ECOG status and smoking history, for the first supplementary analysis estimand of DFS
	Section 2.6.2	Specified that the first OS IA was not performed
	Section 2.6.4.2	Specified all levels of baseline demographic and disease characteristics, including age, gender, race and ECOG status and smoking history, for the first supplementary analysis estimand of OS
	Section 2.8.3	Added Table 2-7 for time windows for laboratory assessments
	Section 2.8.4.2	Removed respiratory rate from Table 2-8 since it was not collected
	[redacted]	[redacted]
	Section 2.13.1	Specified that the efficacy boundary and observed *p*-value at the final DFS analysis will take into account only the interim analysis actually performed (i.e. futility analysis)
	Section 2.13.2	Removed "At the time of interim analyses for DFS, an interim analysis for the key secondary endpoint of OS will be performed by an independent statistician" since the efficacy interim analysis of DFS was not performed
	Section 2.13.2	Specified that the OS efficacy boundaries at the interim and final looks will be calculated according to actual events observed at performed analysis timepoints
	Section 2.13.3	Specified the final DFS analysis, all OS IAs and final OS analyses will be performed by Sponsor's team
	Section 3.1	Specified that the efficacy DFS IA was not performed and final DFS analysis is planned to be performed when approximately 392 DFS events are documented irrespective of the accrued sample size (even if less than 1500 subjects) to ensure the trial is not over-powered.

Date	Section	Changes
	Section 3.2	Specified that the OS efficacy boundaries will be calculated based on actual events and the number of interim analyses that are actually performed.
	Section 5.3	The formula for corrected calcium was updated: Corrected Calcium (mmol/L) = Calcium (mmol/L) + 0.020 [40 - albumin (g/L)]
	Section 6	All uncited references have been removed
21-Mar-2022	Rationale	In line with the protocol amendment 03 (03-Feb-2022), the purpose of this amendment is to remove the second interim analysis for DFS; include the comparison between the canakinumab and placebo arms of DFS and OS in subgroups defined by PD-L1 and CD8 expression as secondary endpoints; add PRO related endpoints: time to first 10 point deterioration for pain, cough and dyspnea per QLQ-LC13; time to first 10 point deterioration for global health status/QoL, shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30; time to 10 point definitive deterioration for role functioning, physical functioning and fatigue per QLQ-C30
	Section 1	Specified the SAP is based on the protocol amendment 03 (03-Feb-2022) and the primary CSR will be based on the final DFS analysis
	Section 1.1	Modified the study design by removing the second DFS interim analysis and the first OS interim analysis originally planned
	Section 1.2	Updated the study objectives and endpoints by adding two biomarker endpoints: DFS and OS in PD-L1 and CD8 subgroup; time to first 10 point deterioration for pain, cough and dyspnea per QLQ-LC13; time to first 10 point deterioration for global health status/QoL, shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30; time to 10 point definitive deterioration for role functioning, physical functioning and fatigue per QLQ-C30; [redacted]
	Section 2.2.1	Added PK and immunogenicity for Japanese subgroup analysis; non-Japanese subgroup analysis and removed Chinese subgroup analysis since a dedicated SAP will address it Added subgroups of PD-L1 categories (<1%, >= 1 %, < 50%, >= 50%, >= 1 % and < 49%) and CD8 categories by median or quartile (including its complements) for both efficacy and safety
	Section 2.3	Added analyses of baseline characteristics by PD-L1 categories (<1%, >= 1 %, < 50%, >= 50%, >= 1 % and < 49%), CD8 categories by median and quartile (including its complements)
	Section 2.5.2	Updated the interim analysis plan for DFS
	Section 2.6.2	Updated the interim analysis plan for OS
	Section 2.7	Added two biomarker secondary efficacy objectives (PD-L1 and CD8)
	Section 2.11	Added PRO related endpoints: time to first 10 point deterioration symptom scores of pain, cough and dyspnea per QLQ-LC13; time to first 10 point deterioration of global health status/QoL, shortness of breath, pain, role functioning, physical

Date	Section	Changes
		functioning and fatigue per QLQ-C30 questionnaire; time to 10 point definitive deterioration of role functioning, physical functioning and fatigue per QLQ-C30 Added the plots of change from baseline by visit for all scores per QLQ-LC13, QLQ-C30 and EQ-5D-5L Added tables of compliance and completion rate over time for QLQ-LC13, QLQ-C30 and EQ-5D-5L by treatment arm Added first deterioration analysis at 5 point cut off in sensitivity analysis Added a reference (Osoba 1998)
	Section 2.13.1	Updated the interim analysis plan for DFS and modified all boundary values; updated table 2-9 with the new interim analysis plan
	Section 2.13.2	Updated the interim analysis plan for OS; updated table 2-10 with the new interim analysis plan
	Section 2.13.3	Updated DMC analysis plan with only futility interim analysis of DFS; removed efficacy interim analysis DFS responsibility from DMC
	Section 3.1	Updated interim analysis plan for DFS

Table of contents

List of abbreviations

AE	Adverse event
AESI	Adverse events of special interest
ATC	Anatomical Therapeutic Classification
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DFS	Disease Free Survival
DMC	Data Monitoring Committee
FAS	Full Analysis Set
eCRF	Electronic Case Report Form
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
OS	Overall Survival
PAS	Pharmacokinetic analysis set
PK	Pharmacokinetics
PRO	Patient-reported Outcomes
PT	Preferred Term
QoL	Quality of Life
RAP	Report and Analysis Process
SAP	Statistical Analysis Plan
SOC	System Organ Class
TFLs	Tables, Figures, Listings
WHO	World Health Organization

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the primary Clinical Study Report (CSR) of study CACZ885T2301, a phase III, multicenter, randomized, double blind, placebo-controlled study evaluating the efficacy and safety of canakinumab versus placebo as adjuvant therapy in adult subjects with stages AJCC/UICC v8 II-IIIA and IIIB (T>5cm N2) completely resected (R0) non-small cell lung cancer (NSCLC). As specified in Section 4.3 of the study protocol version 03 (dated 03-Feb-2022), the final analysis of DFS will occur when approximately 392 DFS events are reached and the primary CSR will be based on the final DFS analysis.

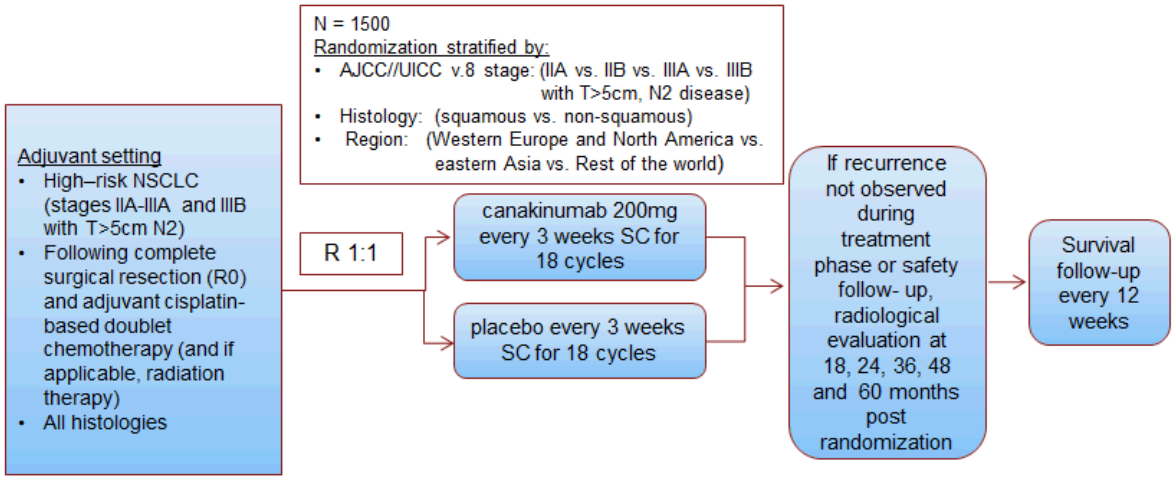
1.1 Study design

This phase III study CACZ885T2301 will enroll adult subjects with completely resected (R0) NSCLC AJCC/UICC v. 8 stages II-IIIA and IIIB (T>5cm N2) disease. Subjects may be screened after undergoing complete surgical resection of their NSCLC and having R0 status confirmed (negative margins on pathologic review), after completing adjuvant cisplatin-based doublet chemotherapy if applicable, (and, if applicable, radiation therapy for stage IIIA N2 or IIIB N2 disease) and after all entry criteria are met. Subjects must not have had preoperative neo-adjuvant chemotherapy or radiotherapy to achieve the R0 status. Approximately 1500 subjects will be randomized in 1:1 ratio to canakinumab 200 mg subcutaneously (s.c.) every 3 weeks or matching placebo s.c. every 3 weeks (Figure 1-1). Randomization will be stratified by AJCC/UICC v. 8 stage (IIA versus IIB versus IIIA versus IIIB with T>5cm, N2 disease), histology (squamous versus non-squamous), and region (Western Europe and North America vs. eastern Asia vs. Rest of the world (RoW)). Subjects will continue their assigned treatment until they complete 18 cycles or experience any one of the following: disease recurrence as determined by the investigator, unacceptable toxicity that precludes further treatment, treatment discontinuation at the discretion of the investigator or subject, death, or lost to follow-up, whichever occurs first. It is postulated that the one year adjuvant treatment will provide an acceptable benefit for subjects who have intermediate or high risk of developing disease recurrence. If disease recurrence is not observed during the treatment phase, subjects will be followed up for five years until disease recurrence, withdrawal of consent by the subject, lost to follow up, death, or termination of the study by the sponsor. All subjects who discontinue from the study treatment will be followed up every 12 weeks for survival until death, lost to follow-up, or withdrawal of consent.

The purpose of this multicenter, randomized, double-blind, placebo-controlled phase III study is to evaluate the efficacy and safety of canakinumab as adjuvant therapy, following standard of care for completely resected (R0) AJCC/UICC v. 8 stages II-IIIA and stage IIIB (T>5cm N2) NSCLC subjects.

Standard of care includes complete resection of the NSCLC with margins free of cancer. At a minimum, two cycles of cisplatin-based doublet chemotherapy are required for all stage IIB-IIIA and IIIB (T>5cm N2) disease subjects; chemotherapy is recommended but not mandatory for stage IIA with no nodal involvement. Radiation therapy is allowed if indicated as per local guidelines or practice. All subjects must have had complete surgical resection of their NSCLC to be eligible for study entry; and margins must be pathologically reviewed and documented as negative. Comparisons will be made between the arms for efficacy: DFS, OS, LCSS and Quality of Life measures (EQ-5D-5L and EORTC QLQ-C30/LC13), and for safety.

Figure 1-1 Study Design



Disease free survival (DFS) by local investigator assessment is the primary endpoint in this study. Overall survival (OS) is the key secondary endpoint.

One futility interim analysis (IA) is planned for DFS when *approximately* 196 (50%) of the 392 DFS events have been observed. The intent is to determine whether there is a need to stop the study early for lack of efficacy (futility), there is no plan to stop the study for efficacy at the futility IA.

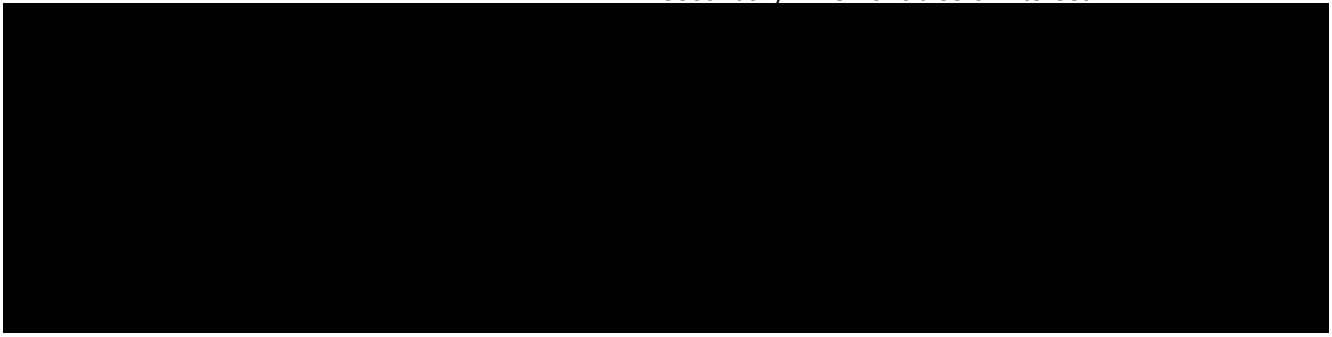
A hierarchical testing procedure will be adopted in this study. OS analysis will be performed *only* if the primary efficacy endpoint DFS is statistically significant. There are a maximum of three OS analyses planned, including two OS IAs and final OS analysis:

- At the time of final DFS analysis (provided DFS is statistically significant at final DFS analysis), at which point a total of *approximately* 318 (63%) deaths are expected
- An additional IA for OS when a total of *approximately* 418 (83%) deaths are expected
- A final analysis for OS when *approximately* 504 deaths are expected

This study has a data monitoring committee (DMC), which will function independently of all other individuals associated with the conduct of this clinical trial, including the site investigators participating in this study. The DMC will be responsible to review safety data approximately 6 months after the first randomized subject has started study treatment, provided at least 50 subjects have been treated, and approximately every one year thereafter until all subjects have completed 18 cycles (approximately 54 weeks) of treatment or final DFS analysis, whichever occurs earlier. The DMC will also be responsible to review efficacy data at DFS futility IA. This includes but does not limit the role of the DMC to evaluate these data and to provide recommendations to the sponsor to continue, modify, or stop the study early. The IA will be performed by an independent statistician external to Novartis and the results will be provided to the DMC by the independent statistician. A separate DMC SAP specifies the analyses to be performed for the DMC safety/efficacy reviews.

1.2 Study objectives and endpoints

Objective	Endpoint
Primary	
The primary objective is to compare the Disease-free survival (DFS) in the canakinumab versus placebo arms as determined by local investigator assessment.	DFS determined by local investigator assessment. See Section 2.5.1 for the primary estimand
Key secondary	
To compare overall survival (OS) in the canakinumab arm versus placebo arm	OS. See Section 2.6.1 for the key secondary estimand
Other secondary	
1. To compare DFS by local investigator assessment and OS in the canakinumab versus placebo arms in subgroups defined respectively by PD-L1 and CD8 expression levels	1. DFS by local investigator assessment and OS in PD-L1 and CD8 subgroups
2. To compare lung cancer specific survival in the canakinumab arm versus placebo arm	2. Lung cancer specific survival (LCSS)
3. To characterize the safety profile of canakinumab	3. Frequency of AEs, ECGs and laboratory abnormalities
4. To characterize the pharmacokinetics of canakinumab therapy	4. Serum concentration-time profiles of canakinumab and appropriate individual PK parameters based on population PK model
5. To characterize the prevalence and incidence of immunogenicity (anti-drug antibodies, ADA) of canakinumab	5. Serum concentrations of anti-canakinumab antibodies
6. To assess the effect of canakinumab versus placebo on PROs (EORTC QLQ-C30 with QLQ-LC13 incorporated and EQ-5D) including functioning and health-related quality of life	6. Time to 10 point definitive deterioration of symptom scores of pain, cough and dyspnea per QLQ-LC13 questionnaire are primary PRO variables of interest. Time to 10 point definitive deterioration of global health status/QoL,shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30 questionnaire, time to first 10 point deterioration for symptom scores of pain, cough, dyspnea per QLQ-LC13 questionnaire, time to first 10 point deterioration for global health status/QoL, shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30 questionnaire together with the utilities derived from EQ-5D-5L are secondary PRO variables of interest



2 Statistical methods

2.1 Data analysis general information

The futility IA of DFS will be performed by an independent statistician external to Novartis. The final DFS analysis will be the basis of the primary CSR. Novartis will perform the analyses specified in this SAP. SAS version 9.4 or later will be used for analysis.

Data included in the analysis

There is one IA and a final analysis planned for the primary efficacy endpoint of DFS. Up to two IAs and a final analysis may be performed for the key secondary endpoint of OS if DFS is statistically significant. A unique cut-off date will be established after the targeted number of events for the planned interim and final analyses has been documented. For each of the analyses, all statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date, e.g. vital sign assessment date or start date of an adverse event, prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date or after subjects withdraw their consent forms will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

For any CSR after the primary CSR, a separate set of TFL shells will detail the planned outputs. If futility is claimed at the first DFS IA, neither DFS and OS is statistically significant, or the study is terminated by the sponsor, it can result in an either close-out CSR or abbreviated CSR.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of subjects enrolled at centers, no center effect will be assessed.

Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables by treatment arm; a missing category will be included as applicable. Percentages will be calculated using the number of subjects in the relevant population or subgroup as the denominator.

Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (e.g. mean, standard deviation, median, percentiles, minimum, and maximum) by treatment arm.

2.1.1 General definitions

Study treatment

Study treatment will refer to canakinumab or placebo.

Date of first administration of study treatment

The date of first administration of study treatment is defined as the first date when a non-zero dose of canakinumab or placebo is administered and recorded on the Dosage Administration Record (DAR) (e) CRF. The date of first administration of study treatment will also be referred as start of study treatment.

Date of last administration of study treatment

The date of last administration of study treatment is defined as is the last date when a nonzero dose of canakinumab or placebo is administered and recorded on DAR eCRF. The date of last administration of study treatment will also be referred as end of study treatment.

Study day

The study day, describes the day of the event or assessment date, relative to the reference start date.

The study day is defined as:

- the date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- the date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference start date for safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, pk etc.) is the start of study treatment.

The reference start date for all other, non-safety assessments (i.e., disease recurrence, survival, LCSS, ECOG performance status, and patient reported outcomes (PRO)) is the date of randomization.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Time unit

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

Baseline

For efficacy evaluations, the last non-missing assessment, including unscheduled assessments on or before the date of randomization is taken as “baseline” value or “baseline” assessment. In the context of baseline definition, the efficacy evaluations also include PRO and performance status.

For safety evaluations, the last available assessment on or before the date of start of study treatment is taken or ‘‘baseline” assessment.

If subjects have no value as defined above, the baseline result will be missing.

On-treatment assessment/event and observation periods

For adverse event reporting the overall observation period will be divided into three mutually exclusive segments:

- ***pre-treatment period***: from day of subject’s informed consent to the day before first administration of study treatment
- ***on-treatment period***: from date of first administration of study treatment to 130 days (approximately five terminal half-lives of canakinumab) after date of last actual administration of any study treatment (including start and stop date)
- ***post-treatment period***: starting at day 131 after last administration of study treatment.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for death including on treatment and post treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (***treatment-emergent*** AEs).

However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

Windows for multiple assessments

Time windows will be defined for descriptive summary of PRO and ECOG performance status data by visit and longitudinal data analysis (Table 2-1). If more than one assessment is available in the same time window, the assessment closest to the planned date will be considered. If two assessments are obtained with the same time difference compared to the scheduled visit day, the assessment obtained prior to visit will be considered. Data obtained at the end of treatment will be classified as other assessment in the corresponding time window.

Table 2-1 Time windows for PRO and ECOG PS assessments

Time Window	Planned Visit Timing	Time Window Definition
Treatment phase		
Baseline	On or before Study Day 1[a]	≤ Study Day 1
Week 3	Study Day 22	Study Days 2 to 32
Week 6	Study Day 43	Study Days 33 to 53
.....	.....	.....
Week k	Study Day 21*(k/3)+1	Study Day
(with k = 9, 12…)		21*(k/3)-9 to 21*(k/3)+11
		If last dose date is in the window, upper bound of this time window will be EOT visit date +7

		"Note: EOT will be included if EOT is performed within 7 days of permanent discontinuation of study treatment"
Safety follow-up phase		
Safety follow-up 1	26 days after treatment discontinuation	[EOT visit date+8; EOT date +39]
Safety follow-up 2	52 days after treatment discontinuation	[EOT visit date+40; EOT date +65]
Safety follow-up 3	78 days after treatment discontinuation	[EOT visit date+66; EOT date +91]
Safety follow-up 4	104 days after treatment discontinuation	[EOT visit date+92; EOT date +117]
Safety follow-up 5	130 days after treatment discontinuation	[EOT visit date+118; EOT date +143]
Post treatment efficacy follow-up phase		
Post treatment follow-up (Efficacy follow up)	Every 26 weeks in Year 2 and 3; Every 52 weeks in Year 4 and 5	First time window: [upper bound of the last previous time windows with assessment + 1; TA date + TA interval at next visit/2] Otherwise : [TA date - TA interval at this visit/2; TA date+TA interval at next visit/2]
7 days post disease progression (PRO only)	Within 7 days post disease progression	[disease progression date+1; disease progression date +21]
28 days post disease progression (PRO only)	28 days post disease progression	[disease progression date+22; disease progression date +42]

[a] Study Day 1 = randomization date and PRO assessment on Cycle 1 Day1 will be baseline for PRO if randomization date is not Cycle 1 Day1.

Last contact date

The last contact date will be derived for subjects not known to have died at the analysis cut-off using the last complete date among the following (Table 2-2):

Table 2-1 Last contact date data sources

Source data	Conditions
Date of Randomization	No condition

Source data	Conditions
Last contact date/last date subject was known to be alive from Survival Follow-up page	Subject status is reported to be alive, lost to follow-up or unknown.
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term.
Start/End dates from drug administration record	Non-missing dose. Doses of 0 are allowed.
Date of disease recurrence	No condition
Laboratory/PK collection dates	Sample collection marked as 'done'.
Vital signs date	At least one non-missing parameter value
Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date. The cut-off date will not be used for last contact date, unless the subject was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring is coming from 'Survival information' eCRF.

The last contact date will be used for censoring of subjects in the analysis of overall survival.

2.2 Analysis sets

Full Analysis Set

The Full Analysis Set (FAS) comprises of all subjects to whom study treatment has been assigned by randomization.

According to the intent to treat principle, subjects will be analyzed according to the treatment and strata (AJCC/UICC v8 stage: IIA versus IIB versus IIIA versus IIIB with T>5cm N2 disease; Histology: squamous versus non-squamous; and Region: Western Europe and North America vs. eastern Asia vs. Rest of the world (RoW)) to which they have been assigned during the randomization procedure.

Safety

The Safety Set includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment they received, either canakinumab or placebo. The treatment received is defined as the randomized treatment if the subject took at least one dose of that treatment or the first treatment received if the randomized treatment was never received.

Pharmacokinetic analysis set (PAS)

The Pharmacokinetic analysis set (PAS) includes all subjects who receive at least one dose of canakinumab and provide at least one evaluable PK sample. A PK sample for canakinumab is considered evaluable when subjects have:

- at least one dose of canakinumab prior to sampling, except C1D1 pre-dose sample
- Pre-dose samples collected before the next dose administration
- Pre-dose samples (except for Cycle 1 Day 1) collected between 16 days (384 hours) and 26 days (624 hours) after the last 200 mg canakinumab dose administration
- Received 200 mg of canakinumab prior to post-dose PK sampling

Subject Classification:

Subjects may be excluded from the analysis populations defined above based on the protocol deviations entered in the database and/or on specific subject classification rules defined in Table 2-3.

Table 2-2 Subject classification based on protocol deviations and non-PD criteria

Analysis set	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS	No written inform consent	Not applicable
Safety set	No written inform consent	No dose of study medication
PAS	No written inform consent	See definition of PAS

Withdrawal of Informed Consent

Any data collected in the clinical database after a subject withdraws informed consent from all further participation in the trial, will not be included in the analyses. The date on which a subject withdraws full consent is recorded in the eCRF.

Death events may be used in the analysis if captured from public records, local law, and subject informed consent permitting.

Additional data for which there is a separate informed consent, e.g. PK, [redacted] etc., collected in the clinical database without having obtained that consent, will not be included in the analyses. These data will be excluded by the presence of the appropriate protocol deviation criterion.

2.2.1 Subgroup of interest

Efficacy

The primary endpoint (DFS) and key secondary endpoint (OS) will be summarized by the following subgroups to examine homogeneity of treatment effect, provided that the corresponding primary efficacy analysis and key secondary efficacy analysis based on the FAS are statistically significant respectively:

- AJCC/UICC v. 8 stage (IIA vs. IIB vs. IIIA vs. IIIB with T>5cm, N2 disease) based on randomization data from IRT
- AJCC/UICC v. 8 stage (IIA vs. IIB vs. IIIA vs. IIIB with T>5cm, N2 disease) based on eCRF data
- Histology (squamous vs. non-squamous) based on randomization data from IRT

- Histology (squamous vs. non-squamous) based on eCRF data
- Region (Western Europe and North America vs. eastern Asia vs. Rest of the world (RoW)) based on randomization data from IRT
- Sex (Male vs. Female)
- Race (White vs. Black vs. Asian vs. Others (includes "Native Hawaiian or Other Pacific Islander" and "American Indian or Alaska Native"))
- Age category (<50 years vs. (≥50 years and <65 years) vs. (≥65 years and <75 years) vs. ≥75 years)
- ECOG PS status (0 vs. ≥ 1)
- Smoking history (never vs. former smoker vs. current smoker)
- [redacted]
- [redacted]
- EGFR mutant status (mutant vs. non-mutant) based on *local* testing during screening
- Prior chemotherapy (yes vs. no) in subjects with stage IIA and N0 disease based on eCRF data
- Prior radiotherapy (yes vs. no) in subjects with N2 disease based on eCRF data
- PD-L1 categories (<1%, >= 1 %, < 50%, >= 50%, >= 1 % and < 49%)
- CD8 categories by median or quartile (including its complements)

No formal statistical test of hypotheses will be performed for the subgroups. Kaplan-Meier method will be used for summarizing medians and their corresponding 95% C.I. of DFS or OS. Hazard ratio (HR) for the treatment effect and its 95% confidence intervals will be summarized based on Cox regression model stratified by randomization stratification factors. The objective of the efficacy subgroup analysis is to demonstrate homogeneity of treatment effect in the above subgroups.

Safety

Key safety analyses, including AEs and AESIs, will be repeated on safety set in the following subgroups:

- Sex (Male vs. Female)
- Race (White vs. Black vs. Asian vs. Others (includes "Native Hawaiian or Other Pacific Islander" and "American Indian or Alaska Native"))
- Age category (<50 years vs. (≥50 years and <65 years) vs. (≥65 years and <75 years) vs. ≥75 years)
- ECOG PS status (0 vs. ≥ 1)
- Smoking history (former or never smoker vs. current smoker)
- Prior chemotherapy and no prior radiotherapy vs. no prior chemotherapy and no prior radiotherapy

- Prior radiotherapy (yes vs. no) in subjects with N2 disease based on eCRF data
- BMI ≥ 35kg/m^2 at baseline (yes vs. no)
- PD-L1 categories (<1%, >= 1 %, < 50%, >= 50%, >= 1 % and < 49%)
- CD8 categories by median or quartile (including its complements)

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to a subgroup of subjects, or safety issues that are more commonly observed in a subgroup of subjects.

The following safety summaries will be performed by subgroup:

- AEs, regardless of study treatment, by primary system organ class and preferred term
- AEs with suspected relationship to study treatment by primary system organ class and preferred term
- Adverse Event of Special Interest, irrespective of causality, by grouping, preferred term, maximum grade, and treatment

Country-specific subgroup analyses

Key efficacy and safety outputs, including baseline characteristics, PK and immunogenicity, will be repeated for Japanese subjects and Non-Japanese subgroup respectively.

2.3 Subject disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline and demographic summaries and listings unless otherwise specified. Summaries will be reported by treatment arm and for all subjects. In addition, listings will be reported by treatment arm to assess baseline comparability. No inferential statistics will be provided.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment. Categorical data (e.g. gender, age groups: <50 years vs. (≥50 years and <65 years) vs. (≥ 65 years and < 75 years) vs. ≥ 75 years, race, WHO performance status, and smoking history), PD-L1 categories (<1%, >= 1 %, < 50%, >= 50%, >= 1 % and < 49%) and CD8 categories by median and quartile (including its complements) will be summarized by frequency counts and percentages; the number and percentage of subjects with missing data will be provided. Continuous data (e.g. age, weight, height, and body mass index (BMI)) will be summarized by descriptive statistics (e.g. N, mean, median, standard deviation, percentiles, minimum, and maximum). BMI (kg/m^2) will be calculated using weight and height at baseline.

In light of COVID-19 pandemic impacting the study during the enrollment phase, summary of demographic and baseline characteristics will be generated by groups of subjects who are enrolled in different pandemic phases to have an understanding of whether there is a shift in subject population characteristics. The pandemic phase will be divided into two phases: pre-pandemic period and pandemic period. Table 2-4 specifies COVID-19 start and end dates by region/country to determine the pandemic phase.

Table 2-4 COVID-19 start and end dates by region/country

Region/Country	Start Date	End Date
China	01Jan2020	Not defined yet
South Korea	20Feb2020	Not defined yet
Japan	21Feb2020	Not defined yet
Italy	23Feb2020	Not defined yet
Rest of the world	01Mar2020	Not defined yet

Baseline stratification factors

The number (%) of subjects in each randomization stratum based on data obtained from the IRT system will be summarized overall and by treatment for the FAS. Discordances between the stratum recorded in IRT at the time of randomization and the actual stratum recorded in the clinical database through the data collected on eCRF will be cross-tabulated and listed.

Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following: primary site of cancer, details of tumor histology/cytology, and TNM pathological staging after resection.

Medical history

Medical history and current medical conditions at screening, including cancer-related conditions and symptoms entered on (e) CRF, will be summarized and listed by primary system organ class (SOC), preferred term (PT), and treatment. They will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

Other

All data collected at baseline, including source of subject referral and other informed consents than main study informed consent, will be listed.

2.3.1 Subject disposition

All screened and randomized subjects will be summarized by country and center. The number (%) of randomized subjects will be presented overall and by treatment. The number (%) of screened subjects and not-randomized subjects, along with the reasons of screening failure, will be displayed. The number (%) of randomized subjects who are still on-treatment, who discontinued the treatment or study, along with the reasons for discontinuation, will be presented overall and by treatment.

The following summaries will be provided (with % based on the total number of FAS subjects):

- Number (%) of subjects who were randomized (based on data from IRT system)
- Number (%) of subjects who were randomized but not treated (based on ‘DAR’ eCRF page not completed for study treatment)
- Primary reason for not being treated (based on “Treatment disposition” eCRF page)

- Number (%) of subjects who were treated (based on 'DAR' eCRF pages of study treatment completed with non-zero dose administered)
- Number (%) of subjects who are still on-treatment (based on the 'Treatment disposition' page not completed);
- Number (%) of subjects who discontinued the study treatment (based on the 'Treatment disposition' page)
- Primary reason for treatment discontinuation (based on the 'Treatment disposition' page)
- Number (%) of subjects who discontinued study (based on the 'Study disposition' page);
- Reasons for study discontinuation (based on 'Study disposition'page)

2.3.2 Protocol deviations

Protocol deviations, including COVID-19 pandemic related and not related, will be summarized in total by treatment. In addition to the pre-defined standard protocol deviation terms, the following protocol deviations related to COIVD-19 in line with "FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency" (March 2020) and "Guidance on the management of clinical trials during the COVID-19 (coronavirus) pandemic" (April 2020) from EMA will be summarized:

- Missing visits
- Changes in procedures and assessments
- Planned visits not done at sites
- Changes in drug supply method
- Treatment not given
- Subject discontinuation due to COVID-19 situation
- Dose interruption greater than 84 days

The relationship of protocol deviation to COVID-19 will also be summarized using the following descriptions:

- COVID-19 Health Status: (i.e. subject's COVID-19 infection led to this PD)
- COVID-19 *Site issue: (e.g. site closed, personnel not available)
- COVID-19 Lockdown: (e.g. site is active but subject not allowed to come)
- COVID-19 Subject/Subject concern: (e.g. site is active, subject/subject could come but refused to come / do assessment)
- COVID-19 Drug supply issue (e.g. drug was delivered to home)
- COVID-19 Other: (e.g. situation not already covered by the information above)

* For deviation with multiple relationships, 'Site Issue' was reported as the primary relationship to COVID-19 if another relationship was also applicable.

In addition, COVID-19 related outcomes (e.g., COVID-19 AEs, treatment discontinuation and study discontinuation due to COVID-19, death due to COVID-19) will be descriptively summarized by country, site, and treatment.

All protocol deviations will be listed.

2.3.3 Analysis sets

The number (%) of subjects in each analysis set will be summarized by treatment and randomization stratum.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

2.4.1 Study treatment / compliance

Duration of exposure, actual cumulative dose, dose intensity (DI), and relative dose intensity (RDI) will be summarized by treatment. Duration of exposure will be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of subjects in each interval. The number (%) of subjects who have dose interruptions, along with interruption reasons, will be summarized by treatment. No dose modification is allowed as per study protocol.

Subject level listings of all doses administered on-treatment, along with dose interruption reasons, will be produced.

The safety set will be used for all summaries and listings of study treatment.

Duration of exposure to study treatment

Last date of exposure and duration of exposure are defined as below:

- Last date of exposure to study treatment = last dosing date of study treatment + 20 days (If the subject died or was lost to follow-up before the derived last date of exposure, the last date of exposure to study treatment is the date of death or the date of last contact before lost to follow-up, respectively. If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.)
- Duration of exposure to study treatment (days):
  last date of exposure to study treatment − first dosing date of study treatment + 1 (periods of interruption are not excluded)

Summary of duration of exposure of study treatment in months will include categorical summaries and continuous summaries (i.e. mean, standard deviation, etc.).

Cumulative dose

Cumulative dose of a study treatment is defined as the total dose given during the study treatment exposure.

The **planned cumulative dose** for a study treatment refers to the total planned dose per the protocol up to the last date of study treatment.

- Cumulative planned dose for a study treatment:
  planned dose per cycle as per protocol * number of complete cycles the subject should have been over the duration of exposure
- Number of complete cycles: (last dosing date of study treatment – first dosing date of study treatment + 21)/21
- Planned dose per cycle as per protocol: 200 mg/cycle

The **actual cumulative dose** refers to the total actual dose administered, over the duration for which the subject is on the study treatment as documented in the Dose Administration eCRF.

For subjects who did not take any drug, the cumulative dose is equal to zero.

The actual cumulative dose should be defined based on the days when the subject is assumed to have taken a non-zero dose during dosing periods.

- Actual cumulative dose (mg):
  total dose of study treatment taken by a subject in the study calculated as the sum of amount of dose administered

Dose intensity and relative dose intensity

Dose intensity (DI) for subjects with non-zero duration of exposure is defined as follows:

- DI (mg/cycle) = Actual Cumulative dose (mg) / Duration of exposure to study treatment (days)*21 (days/cycle)

For subjects who did not take any drug, the DI is equal to zero.

Planned dose intensity (PDI) is defined as follows:

- PDI (mg/cycle) = Planned Cumulative dose (mg) / Duration of exposure to study treatment (days)*21 (days/cycle)

Relative dose intensity (RDI) is defined as follows:

- RDI = DI (mg/cycle) / PDI (mg/cycle)

DI and RDI will be summarized by treatment arm.

Dose interruptions or permanent discontinuations

The number of subjects who have dose permanent discontinuations or interruptions, and the reasons, will be summarized by treatment. Dose reduction is not allowed per study protocol.

'Dose interrupted' will be determined if the total "Dose Administered" equal to 0 from the Dosage Administration CRF pages (DAR). The corresponding reasons will be summarized as well. If multiple interruptions are entered on consecutive cycles with different reasons, they will be counted as separate interruptions, whereas they will be counted as one single interruption. Any dose interruption reported after the last dose of study drug followed by permanent discontinuation won't be considered as a dose interruption.

2.4.2 Prior, concomitant and post therapies

Prior anti-cancer therapy

The number and percentage of subjects, who received prior anti-neoplastic medications, radiotherapies, or surgeries, will be summarized by treatment. Prior anti-neoplastic medications will also be summarized by therapy type, setting, lowest ATC class, preferred term, and treatment. Summaries will include total number of regimens. For radiotherapies, time since last radiotherapy, locations, and setting of last therapy will be summarized by treatment. For prior surgeries, time since last surgery, procedure, and residual disease of last therapy will be summarized by treatment. In addition, the summaries for prior antineoplastic medications and radiotherapies will be repeated by AJCC/UICC v.

8 stage (including Stage IIA, T2b, N0 T (>4 - 5cm) vs. IIB (T1abc-2ab) vs. IIB T3 >5cm & ≤7cm, N0 vs. IIIA T3 or T4 Tis>5 cm vs. IIIA T≤5cm vs. IIIB T>5cm).

Separate listings will be produced for prior anti-neoplastic medications, radiotherapies, and surgeries.

Anti-neoplastic medications will be coded using the WHO Drug Dictionary (WHO-DD); anti-neoplastic surgery will be coded using MedDRA. The above analyses will be performed using the FAS.

Post treatment anti-cancer therapy

Anti-neoplastic therapies since discontinuation of study treatment will be summarized and listed by ATC class, preferred term, and treatment by means of frequency counts and percentages using FAS. First new Anti-neoplastic therapies related to NSCLC since discontinuation of study treatment will be summarized as well.

Concomitant medications

Concomitant therapies are defined as interventions (therapeutic treatment or procedure), other than the study treatment, administered to a subject in the study treatment period. They includes concomitant medications and therapies.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using frequency counts and percentages. Summaries will include:

- Medications starting on or after the start of study treatment but no later than 130 days after last dose of study treatment and
- Medications starting prior to start of study treatment and continuing after the start of study treatment

All concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 130 days after the last date of study treatment will be flagged. Surgical and medical procedures will be coded using MedDRA and summarized by SOC, preferred term, and treatment.

The safety set will be used for all tables and listings of concomitant medications and therapies.

2.5 Analysis of the primary objective

The primary objective is to compare the DFS in the canakinumab versus placebo arms as determined by local investigator assessment.

2.5.1 Primary estimand

The primary efficacy variable of the study is DFS, defined as the time from the date of randomization to the date of the first documented disease recurrence as assessed by local investigator radiologically or death due to any cause. Disease recurrence includes diagnoses of new primary lung malignancies. Clinical deterioration is not considered as a recurrence of disease. In case of non-conclusive radiological evidence, a biopsy should be performed to confirm recurrence. If a biopsy is not feasible, the subject will be followed up until recurrence can be confirmed as per protocol (radiologically conclusive and/or biopsy)"; radiological assessment date will be used as DFS event date. The primary

analysis will be based on FAS and will include all data observed up to the cut-off date. Censoring conventions are provided in Section 2.5.3.

The scientific objective guiding the primary estimand is to estimate the treatment effect based on the primary endpoint of DFS for canakinumab compared to placebo, for the randomized subjects enrolled with inclusion/exclusion criteria, which were defined in the protocol, irrespective of any treatment discontinuation reasons, any new malignancy not related to lung cancer and had the post-treatment antineoplastic therapies for NSCLC not been administered. It is expected that new therapies such as PD-L1s for adjuvant NSCLC may be approved during the conduct of this trial. In case placebo subjects in this study were to switch to these new approved therapies prior to disease recurrence, it can have a potential impact on the treatment effect in the placebo arm. Hence, the primary estimand uses a hypothetical strategy to assess the treatment effect based on DFS had the post-treatment antineoplastic therapies for NSCLC not been administered. Additional intercurrent events such as discontinuation of study treatment, appearance of new primary malignancy not related to lung, post-treatment antineoplastic therapies not related to lung, and intercurrent events due to COVID-19 pandemic are expected in both treatment arms. Hence, the primary estimand of interest will be to assess the treatment effect based on DFS, irrespective of these intercurrent events.

The primary estimand is described by the following five attributes:

1. The **target population** includes all randomized subjects (FAS).
2. The **primary variable** is DFS defined as the time from the date of randomization to the date of the first documented disease recurrence, assessed by local investigator radiologically, or death due to any cause. Any new primary lung malignancy will be handled as disease recurrence.
3. The **treatment of interest** is the randomized treatment (canakinumab arm or placebo arm).
4. The remaining intercurrent events describe how events that may occur after randomization are considered when assessing the treatment effect.
    - Discontinuation of study treatment: DFS will take into account all DFS assessments, irrespective of the study treatment discontinuation reasons (treatment policy)
    - Initiation of post-treatment anti-neoplastic therapies for NSCLC: DFS will be censored at the last adequate tumor assessment prior to the initiation of post-treatment antineoplastic therapy (hypothetical strategy)
    - New primary malignancy not related to Lung cancer: DFS will take into account all DFS assessments, irrespective of the occurrence of any new malignancy not related to lung cancer (treatment policy)
    - Initiation of post-treatment anti-neoplastic therapies for malignancies not related to lung cancer: DFS will take into account all DFS assessments, irrespective of initiation of any post-treatment antineoplastic therapies for new malignancy not related to lung cancer (treatment policy)
    - Any unforeseen intercurrent events due to COVID-19 pandemic will be handled by treatment policy strategy
5. The **summary measure** is the hazard ratio (HR) for DFS between two treatment arms. It will be estimated using Cox proportional hazard model stratified by randomization stratification factors.

2.5.2 Statistical hypothesis, model, and method of analysis

Assumed proportional hazards model for DFS, the following statistical hypotheses will be tested to address the primary efficacy objective:

H_{01} (null hypotheses): $\Theta_1 \geq 0$ vs. H_{a1} (alternative hypotheses): $\Theta_1 < 0$

Where Θ_1 is the log hazard ratio of DFS in the canakinumab (investigational) vs. placebo (control).

The primary efficacy analysis to test this hypothesis and compare the two arms will involve a stratified log-rank test at an overall one-sided 2.5% level of significance. The stratification will be based on the following randomization stratification factors: AJCC/UICC v8 stage IIA versus IIB versus IIIA versus IIIB with T>5cm N2 disease; Histology: squamous versus non-squamous; and Region: Western Europe and North America vs. eastern Asia vs. Rest of the world (RoW). A group sequential design is employed for the primary efficacy variable DFS. One interim analysis, along with a final analysis, is planned, using a Lan-DeMets (O'Brien-Fleming) α-spending function for efficacy and a non-binding user-defined β-spending function (gamma function with γ=-3) for futility. There is only futility assessed at the IA.

The DFS distribution will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves will be plotted for each arm. The DFS Kaplan-Meier estimate along with 95% confidence intervals will be presented by year for each arm. The hazard ratio for DFS will be calculated, along with its 95% confidence interval, from a stratified Cox model using the same stratification factors as those for the log-rank test.

2.5.3 Handling of missing values/censoring/discontinuations

For subjects who received any subsequent anti-neoplastic therapy for NSCLC, the following rules will apply:

- If new NSCLC anti-neoplastic therapy is started without evidence of documented disease recurrence, then DFS will be censored at the date of last tumor assessment before the date of initiation of anti-cancer therapy
- If a subject does not have any tumor assessment before the date of initiation of new anti-neoplastic therapy for NSCLC, DFS will be censored at the date of randomization.

Clinical review will be performed to determine if anti-neoplastic therapy is NSCLC related in case of multiple malignancies.

Any new primary malignancies *not* related to lung or clinical deterioration will not be considered as a disease recurrence.

Event/censoring dates for the primary estimand will be determined by Table 2-5:

Table 2-5 Outcome and event/censoring dates for primary DFS analysis

Situation	Date	Outcome
Evidence of unresected NSCLC or distant metastatic NSCLC disease on screening evaluation	Date of randomization	Censored
Death before any tumor assessment	Date of death	Event

Situation	Date	Outcome
Recurrence or death documented between scheduled visits	Date of tumor assessment (or death)	Event
No recurrence (or death)	Date of last adequate tumor assessment	Censored
New antineoplastic therapy (including surgery) for NSCLC given prior to recurrence or death	Date of last adequate assessment on or prior to starting new anti-cancer therapy	Censored
New antineoplastic therapy (including surgery) for NSCLC prior to any scheduled visit	Date of randomization	Censored
New primary *lung* malignancies	Date of tumor assessment	Event

The DFS censoring reasons are defined in the following way. If the date is the earliest of the following dates:

- Analysis cut-off date
- Date of consent withdrawal
- Date of new NSCLC antineoplastic therapy
- Date of lost to follow-up

Then the DFS censoring reason will be:

- Ongoing without event
- Withdrew consent
- New NSCLC antineoplastic therapy
- Lost to follow-up

2.5.4 Supportive analyses for primary estimand

2.5.4.1 Sensitivity analyses for primary estimand

First sensitivity analysis estimand: The target population, the primary variable, the treatment of interest, intercurrent events, and the summary measure of this endpoint are the same as those for the primary estimand. To assess the impact of stratification, the two treatment arms will be compared using the unstratified log-rank test. In the summary tables, this approach is referred to as '*unstratified DFS sensitivity analysis*'.

Second sensitivity analysis estimand: the target population, the primary variable, the treatment of interest, intercurrent events, and the summary measure of this endpoint are the same as those for the primary estimand. The two treatment arms will be compared using the stratified log-rank test with stratification factors derived from the clinical database. In the summary tables, this approach is referred to as '*DFS sensitivity analysis with stratification factors from eCRF*'.

Third sensitivity analysis estimand: the target population, the primary variable, the treatment of interest, intercurrent events, and the summary measure of this endpoint are the same as those for the primary estimand. Subjects with two or more times missing tumor assessments will be censored at the last adequate tumor assessment. Algorithm to identify missing tumor assessments are defined in Appendix 5.6. In the summary tables, this approach is referred to as '*DFS sensitivity analysis with censoring for two or more missing assessments*'.

2.5.4.2 Supplementary analyses for primary estimand

First supplementary analysis estimand: the target population, the primary variable, the treatment of interest, and intercurrent events are the same as those for the primary estimand. A Cox regression model, stratified by randomization stratification factors, will be fitted to evaluate the effect of the baseline demographic and disease characteristics, which include age (< 50 years vs. (≥ 50 years and < 65 years) vs. (≥ 65 years and < 75 years) vs. ≥75 years), gender (Male vs. Female), race (White vs. Black vs. Asian vs. Others (includes "Native Hawaiian or Other Pacific Islander" and "American Indian or Alaska Native"), ECOG status (0 vs. ≥1), and smoking history (never vs. former smoker vs. current smoker). In the summary tables, this approach is referred to as '*DFS supplementary analysis adjusted for baseline covariates*'.

Second supplementary analysis estimand: the target population, intercurrent events, and the summary measure of this endpoint are the same as those for the primary estimand. The treatment of interest is the randomized treatment (canakinumab arm or placebo arm without any NSCLC antineoplastic therapy post randomization). The primary variable is defined as the time from the date of randomization to the date of the first documented disease recurrence, date of death due to any cause, or date of new NSCLC antineoplastic therapy, whichever occurs first. Any new NSCLC antineoplastic therapy post randomization prior to recurrence/death will be considered as a DFS event (composite strategy). In the summary tables, this approach is referred to as '*ANP leading to DFS event supplementary analysis*'.

Third supplementary analysis estimand: the target population, treatment of interest, and the summary measure of this endpoint are the same as those for the primary estimand. The intercurrent event of initiation of post-treatment anti-neoplastic therapies for NSCLC will be handled by treatment policy, i.e., DFS assessments after initiation of post-treatment anti-neoplastic therapies for NSCLC will be taken into account in assessment of treatment effect. In the summary tables, this approach is referred to as '*DFS supplementary analysis regardless of ANP*'.

Fourth supplementary analysis estimand: In light of COVID-19 pandemic and its potential impact on treatment effect, an important scientific question of interest is to estimate the treatment effect based on DFS, had COVID-19 pandemic not occurred. The target population, treatment of interest, and the summary measure of this endpoint are the same as those for the primary estimand. The primary variable is defined as the time from the date of randomization to the date of the first documented recurrence or date of death due to non-COVID-19 pandemic reasons, whichever occurs first. The remaining intercurrent events will be handled as follows:

- **Discontinuation of study treatment due to any non-COVID-19 pandemic reasons**: DFS will take into account all DFS events, irrespective of the study treatment discontinuation reasons (treatment policy)
- **Discontinuation of study treatment due to COVID-19 pandemic reasons**: DFS will be censored at the date of the last adequate tumor assessment prior to discontinuation of treatment

due to COVID-19 pandemic (hypothetical strategy). The discontinuation reason due to COVID-19 will be identified from the defined COVID-19 protocol deviations.

- **Medications used for treating suspected or confirmed COVID-19 cases:** DFS will be censored at the date of the last adequate tumor assessment prior to the administration of COVID-19 medication (hypothetical strategy). Details to identify medications in Section 2.5.4.2.1
- **Death due to COVID-19**: DFS will be censored at the date of the last adequate tumor assessment prior to death due to COVID-19 (hypothetical strategy)

In the summary tables, this approach is referred to as '*COVID-19 DFS supplementary analysis*'.

2.5.4.2.1 Medications treating suspected or confirmed COVID-19

The identification of the drug used to treat COVID-19 depends on whether COVID-19 is confirmed or suspected.

For **confirmed COVID-19**, the drug treating COVID-19 is selected based on:

- Drug indication that contains key word "COVID" or "SARS-COV"
- Drug class (CM ATC level 4) is aminoquinolines, protease inhibitors, glucocorticoids, macrolides, antiviral or antiretroviral; or drug name (CM standard medicine name) contains key word "hydroxychloroquine" or "chloroquine"

For **suspected COVID-19**, the drug treating COVID-19 is selected based on:

- AEs with the preferred term of "SARS-CoV-2 test negative" or "suspected COVID-19"
- Drug class (CM ATC level 4) is aminoquinolines, protease inhibitors, glucocorticoids, macrolides, antiviral or antiretroviral; drug name (CM standard medicine name) contains key word "hydroxychloroquine" or "chloroquine"; drug class (CM ATC level 4) is glucocorticoids and drug indication contains key word "PNEUM" or equals to "DYSPNEA"
- Selecting the first drug with start date within a time window of +/- 7 days of AE start date

2.5.5 Subgroup analyses for the primary estimand

If the primary efficacy analysis is *statistically significant*, the primary endpoint of DFS will be summarized for the subgroups specified in Section 2.2.1:

- Median DFS with its 95% CI from Kaplan-Meier estimation of the DFS distribution
- Hazard ratio with its 95% CI using Cox proportional hazards model stratified by randomization stratification factors

Subgroup analyses are intended to explore the consistency (homogeneity) of treatment effect. A forest plot, including sample size, number of events, and HR with its 95% CI, will be produced for subgroups. *p*-values will not be generated.

2.6 Analysis of the key secondary objective

The key secondary objective is to determine whether canakinumab prolongs OS, compared with placebo.

2.6.1 Key secondary estimand

The key secondary endpoint is OS, defined as the time from the date of randomization to the date of death due to any cause. If a subject is not known to have died, then OS will be censored at the last date the subject was known to be alive (on or before the cut-off date). The OS analysis will be based on the FAS and will include all data up to the cut-off date.

The scientific objective guiding the key secondary estimand is to estimate the treatment effect based on the key secondary endpoint of OS for canakinumab compared to placebo for the randomized subjects with completely resected (R0) NSCLC AJCC/UICC v.8 stages II-IIIA and IIIB (T> 5 cm N2) disease, irrespective of any treatment discontinuation reasons and any post-treatment antineoplastic therapies administered. The justification for the primary estimand is driven by the rationale to assess treatment effect based on overall survival, irrespective of intercurrent events such as treatment discontinuation, new malignancy, and post treatment anti-neoplastic therapies using treatment policy strategy.

The key secondary endpoint analysis will be described by the following five attributes:

1. The **target population** includes all randomized subjects (FAS).
2. The **primary variable** is OS, defined as the time from the date of randomization to the date of death due to any cause.
3. The treatment of interest is the randomized treatment (canakinumab arm or the matching placebo arm) with or without any new anti-neoplastic therapy administered as needed.
4. The remaining **intercurrent event** describes how events that may occur after randomization are considered when assessing the treatment effect.
    - Discontinuation of study treatment: OS will take into account all deaths and assessments, irrespective of the study treatment discontinuation reasons (treatment policy).
    - New primary malignancy: OS will take into account all deaths and assessments, irrespective of the occurrence of any new malignancy (treatment policy)
    - Any unforeseen intercurrent events due to COVID-19 pandemic will be handled by treatment policy strategy.
5. The **summary measure** is the HR for OS between two treatment arms. It will be estimated using Cox proportional hazard model stratified by randomization stratification factors.

2.6.2 Statistical hypothesis, model, and method of analysis

Assumed proportional hazards model for OS, the following statistical hypotheses will be tested *only* if DFS is statistically significant:

H_{02} (null hypotheses): $\Theta_2 \geq 0$ vs. H_{a2} (alternative hypotheses): $\Theta_2 < 0$

Where Θ_2 is the log hazard ratio of OS in the canakinumab (investigational) vs. placebo (control). The analysis to test these hypotheses will involve a stratified log-rank test at an overall one-sided 2.5% level of significance. The stratification will be based on the following randomization stratification factors: AJCC/UICC v. 8 stage IIA versus IIB versus IIIA versus IIIB T>5cm N2 disease; Histology: squamous versus non-squamous; and Region: Western Europe and North America vs. eastern Asia vs. Rest of the world (RoW).

A group sequential design using a Lan-DeMets (O'Brien-Fleming) α-spending function is employed for OS analyses. OS will be tested hierarchically as follows:

- If DFS is statistically significant at the final DFS analysis, OS will be tested. If OS is not statistically significant at this time, an additional IA and final OS analysis will be planned.
- If DFS is not statistically significant at the final DFS analysis, OS will not be tested.

The OS distribution will be estimated using the Kaplan-Meier method. Kaplan-Meier curves, medians OS and its 95% CI will be presented for each arm. The OS Kaplan-Meier estimate along with its 95% CI will be presented by year as well. The HR and its 95% CI for OS will be calculated using a stratified Cox model. The number of subjects censored for OS and their reasons for censoring will be summarized and listed by treatment.

2.6.3 Handling of missing values/censoring/discontinuations

Subjects not known to have died will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than 14 weeks, i.e., the protocol specified interval between the survival follow-up assessments plus 2 weeks.

2.6.4 Supportive Analyses for key secondary estimand

2.6.4.1 Sensitivity analyses for key secondary estimand

First sensitivity analysis estimand: The target population, the primary variable, the treatment of interest, intercurrent events, and the summary measure of this endpoint are the same as those for the key secondary estimand. To assess the impact of stratification, the two treatment arms will be compared using the unstratified log-rank test. In the summary tables, this approach is referred to as '*unstratified OS sensitivity analysis*'.

Second sensitivity analysis estimand: the target population, the primary variable, the treatment of interest, intercurrent events, and the summary measure of this endpoint are the same as those for the key secondary estimand. The two treatment arms will be compared using the stratified log-rank test with stratification factors derived from the clinical database. In the summary tables, this approach is referred to as '*OS sensitivity analysis with stratification factors from eCRF*'.

2.6.4.2 Supplementary analyses for key secondary estimand

First supplementary analysis estimand: the target population, the primary variable, the treatment of interest, and intercurrent events are the same as those for the key secondary estimand. A Cox regression model, stratified by randomization stratification factors, will be fitted to evaluate the effect of the baseline demographic and disease characteristics, which include age (< 50 years vs. (≥ 50 years and < 65 years) vs. (≥ 65 years and < 75 years) vs. ≥75 years), gender (Male vs. Female), race (White vs. Black vs. Asian vs. Others (includes "Native Hawaiian or Other Pacific Islander" and "American Indian or Alaska Native"), ECOG status (0 vs. ≥1), and smoking history (never vs. former smoker vs. current smoker). In the summary tables, this approach is referred to as '*OS supplementary analysis adjusted for baseline covariates*'.

Second supplementary analysis estimand: In addition, analysis to assess the treatment effect based on OS, had COVID-19 pandemic not occurred, will be conducted. The target population, treatment of interest, and the summary measure of this endpoint are the same as those for the key secondary estimand. The variable is defined as the time from the date of randomization to the date of death due to non-COVID-19 pandemic reasons. The remaining intercurrent events will be handled as follows:

- **Discontinuation of study treatment due to any non-COVID-19 pandemic reasons**: OS will take into account all deaths and assessments, irrespective of the study treatment discontinuation reasons (treatment policy)
- **Discontinuation of study treatment due to COVID-19 pandemic reasons**: OS will be censored on the date of discontinuation of treatment due to COVID-19 pandemic (hypothetical strategy). The discontinuation reason due to COVID-19 pandemic will be identified from the defined COVID-19 protocol deviations.
- **Medications used for treating suspected or confirmed COVID-19 cases:** OS will be censored on the date of administration of COVID-19 medication (hypothetical strategy). Details to identify medications are shown in Section 2.5.4.2.1
- **Death due to COVID-19**: OS will be censored on the date of death due to COVID-19 (hypothetical strategy)

In the summary tables, this approach is referred to as '*COVID-19 OS supplementary analysis*'.

2.6.4.3 Subgroup analyses for key secondary estimand

If the key secondary efficacy analysis is *statistically significant*, OS will be summarized for the subgroups specified in Section 2.2.1:

- Median OS with its 95% CI from Kaplan-Meier estimation of the OS distribution
- Hazard ratio with its 95% CI using Cox proportional hazards model stratified by randomization stratification factors

A forest plot, including sample size, number of events, and HR with its 95% CI, will be produced for subgroups. *p*-values will not be generated.

2.7 Analysis of secondary efficacy objectives

Along with the OS as the key secondary endpoints, three secondary efficacy objectives are also evaluated:

- To compare DFS by local investigator assessment and OS in the canakinumab versus placebo arms in subgroups defined by PD-L1 expression levels
- To compare DFS by local investigator assessment and OS in the canakinumab versus placebo arms in subgroups defined by CD8 expression levels
- To compare lung cancer-specific survival (LCSS) in the canakinumab arm versus placebo arm is an important secondary efficacy objective.

2.7.1 Secondary endpoints

In order to evaluate the effect of canakinumab within PD-L1 subgroups, DFS and OS analyses will be performed by PD-L1 subgroups with 1% cut-off on tumor cells (TC) to define positive or negative status as determined by IHC (<1% and ≥1%). Additionally, PD-L1 subgroups will also be assessed using ≥ 50% cut-off (<50% and ≥50%). A comparison between <1%, ≥1-49%, and ≥ 50% PD-L1 expression level subgroups will also be done.

In order to evaluate the effect of canakinumab within CD8 subgroups, DFS and OS analyses will be performed by CD8 subgroups with the median and quartiles (including quartile-complements, e.g.

below lower quartile and above lower quartile) of baseline CD8 as cut-offs. In addition, the relation between CD8 expression in resected tumor sample and clinical outcome will be evaluated using CD8 as continuous, potentially log-transformed, co-variate as well as categorical covariates in a Cox model. These analyses will support the establishment of CD8 as a prognostic/predictive factor if there is an expression level correlating to clinical outcome regardless of treatment (prognostic) and/or associated with canakinumab treatment (predictive).

LCSS is a secondary efficacy endpoint, defined as the time from the date of randomization to the date of death due to lung cancer. For subjects who have died due to reasons other than lung cancer, LCSS will be censored at the death date. For subjects who is not known to have died, LCSS will be censored at the last date the subject was known to be alive (on or before the cut-off date); they will be censored for 'Lost to follow-up' if the time between their last contact date and the analysis cut-off date is longer than 14 weeks, i.e., the protocol specified interval between the survival follow-up assessments plus 2 weeks.

The LCSS distribution will be estimated using the Kaplan-Meier method. Kaplan-Meier curves, medians and its 95% CI will be presented for each arm. The HR and its 95% CI for LCSS will be calculated using a stratified Cox model. The number of censored subjects will be summarized by treatment. Reasons for censoring ('Alive', 'Lost to follow-up', 'Death due to reasons other than lung cancer'), along with death cause, will be summarized and listed by treatment.

2.7.2 Duration of follow-up

Summary information regarding the follow-up of subjects in the FAS will be displayed by arm in order to describe the maturity of data and quality of follow-up.

Randomization period, duration between randomization and cut-off date, and follow-up times for DFS/OS are defined as follows:

- Randomization (recruitment) period = (Date of last subject randomized - Date of first subject randomized + 1)/30.4375 (months).
- Duration between randomization date and data cut-off date = (Data cut-off date – Date of randomization + 1)/30.4375 (months).
- Follow-up time elapsed from study start reference date= (Date of event or censoring – randomization date + 1)/30.4375 (months) regardless of censoring. Date of event or censoring is the same as the one defined for the primary analyses.

The follow-up time for DFS/OS derived from KM method will also be summarized.

2.7.3 New primary cancer other than lung cancer

Summary statistics will be tabulated for subjects with new primary cancer other than lung cancer based on "Diagnosis and extent of new primary Cancer Other than Lung" eCRF page. This analysis will include the following: number of subjects with new primary cancer, primary site of cancer, the time from randomization to initial diagnosis, predominant histology/cytology, staging system and metastatic sites.

2.8 Safety analyses

All safety analyses will be based on the safety set.

2.8.1 Adverse events (AEs)

AE summaries will include all AEs occurring during on-treatment period. All AEs collected in the AE (e)CRF page will be listed along with the information collected on those AEs e.g. AE relationship to study treatment, AE outcome, etc. AEs with start date outside of on-treatment period will be flagged in the listings.

AEs will be summarized by number and percentage of subjects by primary system organ class (SOC), preferred term (PT), and treatment. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. A subject with multiple CTCAE grades for the same preferred term will be summarized under the maximum grade. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries, the primary system organ class will be presented alphabetically and the PTs will be sorted within primary SOC in descending frequency. The sort order for the PTs will be based on their frequency in the investigational treatment.

The following AE summaries will be produced by treatment; overview of AEs and deaths, AEs by SOC and PT, summarized by relationship, seriousness, leading to treatment discontinuation, leading to dose interruption, requiring additional therapy, and leading to fatal outcome. In addition, a summary of serious adverse events with number of occurrences will be produced (an occurrence is defined as > 1 day between start and prior end date of record of same PT).

2.8.1.1 Adverse events of special interest / grouping of AEs

Data analysis of AESIs

An AE of special interest (AESI) is a grouping of AEs that are of scientific and medical concern specific to canakinumab. These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs. Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. The groups are defined according to the MedDRA terms defined in the program Case Retrieval Strategy (CRS) document and will be summarized. The latest version of the CRS document available at the time of the analyses will be used. For each specified AESI, number and percentage of subjects with at least one event of the AESI occurring during on treatment period will be summarized.

Summaries of these AESIs will be provided by treatment (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose interruption, hospitalization, death, etc.).

A listing of all grouping levels down to the MedDRA PTs used to define each AESI will be generated.

Table 2-6 provides current AESI groupings. CRS at the time of database lock will govern how these groups will be summarized:

Table 2-6 AESI groupings

AESI grouping
Infections
Opportunistic infections
Opportunistic infections (based on SMQ Broad)

Opportunistic infections (based on SMQ Narrow)
Neutropenia
Abnormal liver parameters
Thrombocytopenia
Immunogenicity/allergenicity
Autoimmunity reactions
Malignancy
Interactions with vaccines
Pulmonary complications: pulmonary hypertension and interstitial lung disease
Injection site reactions

Time to onset of AESI

Time to onset of CTC grade ≥ 2 AESI is defined as the time from the start of treatment to the start date of the first incidence of an event of CTC grade ≥ 2, i.e., time in days is calculated as (start date of first occurrence of the event) – (date of first dose of study treatment) +1.

In the absence of an event during the on-treatment period, the censoring date applied will be the earliest of the following dates:

- End date of on-treatment period (end of study treatment + 130 days).
- death date
- Start date of new antineoplastic therapy before experiencing any CTC grade ≥ 2 AESI.
- Data cut-off date.
- withdrawal of informed consent date

Time to onset of CTC grade ≥ 2 AESI will be summarized using the Kaplan-Meier method by AESI grouping and treatment. Median time to onset and its 95% CI will be provided and ascending Kaplan-Meier plots will be generated by treatment. The analysis of time to onset of AESI will be performed *only* for AESI groupings of infections, opportunistic infections, neutropenia, DILI (Hepatic transaminase and bilirubin elevations) and thrombocytopenia (if there are sufficient number of events).

The same analysis will be repeated for time to onset of CTC grade ≥ 3 AESI with similar censoring rules applied.

2.8.2 Deaths

Summaries for on-treatment and all deaths will be produced separately by SOC, PT, and treatment. Deaths will be listed with post treatment deaths flagged. A listing of deaths prior to starting treatment will be provided for all screened subjects.

2.8.3 Laboratory data

On analyzing laboratory, data from all sources, including central and local laboratories, will be combined. The summaries will include all assessments available for the lab parameter collected no later than 130 days after the last study treatment administration date. Grading of laboratory values will be assigned programmatically per NCI CTCAE 5.0. The time windows for lab assessments are shown in Table 2-7:

Table 2-7 Time windows for lab assessments

Time Window	Planned Visit Timing	Time Window Definition
On treatment		
Baseline	On or before Study Day 1[a]	≤ Study Day 1
Week 3	Study Day 22	Study Days 2 to 32
Week 6	Study Day 43	Study Days 33 to 53
.....	.....	.....
Week k (with k = 9,12 …)	Study Day 21*(k/3)+1	Study Day 21*(k/3)-9 to 21*(k/3)+11
Safety follow-up 1	26 days after treatment discontinuation	[EOT visit date+8; EOT date +39]
Safety follow-up 2	52 days after treatment discontinuation	[EOT visit date+40; EOT date +65]
Safety follow-up 3	78 days after treatment discontinuation	[EOT visit date+66; EOT date +91]
Safety follow-up 4	104 days after treatment discontinuation	[EOT visit date+92; EOT date +117]
Safety follow-up 5	130 days after treatment discontinuation	[EOT visit date+118; EOT date +143]

[a] Study Day 1 = first treatment date

The following summaries will be produced for hematology and biochemistry laboratory data by laboratory parameter and treatment:

- Worst post-baseline CTC grade, regardless of the baseline status. Each subject will be counted only for the worst grade observed post-baseline
- Shift tables using CTC grades to compare baseline with the worst on-treatment value. Hypo and hyper worst grade will be summarized separately if applicable
- For laboratory tests where CTC grades are not defined, shift tables using the low/normal/high classification to compare baseline with the worst on-treatment value
- Trends of lab parameter values over time, including baseline and selected on-treatment time points, will be displayed via boxplots

The following listings will be produced for the laboratory data:

- Listings of all laboratory data, with CTC grades and classification relative to the laboratory normal range. Lab data collected during the post-treatment period will be flagged
- Listing of all CTC grade 3 or 4 laboratory toxicities

Liver function parameters

Liver function parameters of interest are total bilirubin (TBL), ALT, AST, and alkaline phosphatase (ALP). The number (%) of subjects with worst post-baseline values per Novartis Liver Toxicity guidelines will be summarized by treatment:

The following summaries will be produced:

- ALT > 3xULN
- ALT > 5xULN
- ALT > 10xULN
- ALT > 20xULN
- AST > 3xULN
- AST > 5xULN
- AST > 10xULN
- AST > 20xULN
- ALT or AST > 3xULN
- ALT or AST > 5xULN
- ALT or AST > 8xULN
- ALT or AST > 10xULN
- ALT or AST > 20xULN
- TBL > 2xULN
- TBL > 3xULN

Combined elevations post-baseline:

For AST and ALT ≤ ULN at baseline

- ALT or AST > 3xULN & TBL > 2xULN
- ALT or AST > 3xULN & TBL > 2xULN & ALP < 2xULN
- ALT or AST > 3xULN & TBL > 2xULN & ALP ≥ 2xULN

For ALT or AST > ULN at baseline (Bsl)

- Elevated ALT or AST (*) & TBL (> 2x Bsl and 2xULN)
- Elevated ALT or AST (*) & TBL (> 2x Bsl and 2xULN) & ALP < 2xULN
- Elevated ALT or AST (*) & TBL (> 2x Bsl and 2xULN) & ALP ≥ 2xULN

* Elevated AST or ALT defined as: >3x ULN if =< ULN at baseline, or (>3x Bsl or 8x ULN) if > ULN at baseline

In addition, a listing of all TBILI, ALT, AST, and ALP values for subjects with a post-baseline TBILI > 2xULN, ALT> 3xULN or AST > 3xULN will be provided.

Time to onset of grade 2 or worse and time to onset of grade 3 or worse liver function test abnormalities will be summarized for the following liver function test parameters:

- AST or ALT (whichever occurs first)
- Total bilirubin

Time to onset will be summarized using Kaplan-Meier method. Median and its 95% CI will be summarized.

Time to onset of LFT abnormalities is defined as the time from the start of treatment to the start date of the first incidence of grade 2 or worse LFTs post-baseline (or grade 3 or worse), i.e., time in days is calculated as (start date of first occurrence of LFT abnormalities) – (date of first dose of study treatment) +1. A subject will be censored for time to onset if:

- the subject dies without experiencing the LFT abnormality
- the subject receives a new anticancer therapy without experiencing the LFT abnormality or before LFT abnormality has occurred
- the subject discontinues from the study treatment without experiencing the LFT abnormality (up to 130 days after study treatment discontinuation)
- the subject is still ongoing at the analysis cut-off without experiencing the LFT abnormality

The censoring date will be the earliest date from the following dates: last date of study treatment in the treatment phase + 130 days, analysis cut-off, the day before new anticancer therapy start date, death date, and last on-treatment laboratory sampling date (for the particular parameter) during on-treatment period. For the time to onset of grade 2 or worse LFTs, subjects with grade 2 or worse at baseline will be excluded from the analysis and the same for the time to onset of grade 3 or worse LFTs.

Time to onset of neutropenia and thrombocytopenia:

Time to onset of grade 2 or worse and time to onset of grade 3 or worse neutropenia and thrombocytopenia will be summarized respectively. The censoring rule will be the same as those used at time to onset of LFT abnormalities analysis.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

ECG abnormality for baseline and any post baseline assessments will be summarized by treatment. A listing of all ECG assessments will be produced by treatment with notable values flagged.

2.8.4.2 Vital signs

Vital sign assessments are performed in order to characterize the basic body function. The following parameters will be collected: height (cm), weight (kg), body temperature (°C), pulse rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Data handling

Vital signs collected on treatment will be summarized. Values measured outside of on-treatment period will be flagged in the listings.

Data analysis

For analysis of vital signs, the clinically notable vital sign criteria are provided in Table 2-8:

Table 2-8 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Weight (kg)	increase > 10% from Baseline	decrease > 10% from Baseline
Systolic blood pressure (mmHg)	>=180 with increase from baseline of >=20	<=90 with decrease from baseline of >=20
Diastolic blood pressure (mmHg)	>=105 with increase from baseline of >=15	<=50 with decrease from baseline of >=15
Pulse rate (bpm)	>=100 with increase from baseline of >25%	<=50 with decrease from baseline of > 25%
Body temperature (°C)	>= 39.1	-

The number and percentage of subjects with notable vital sign values (high/low) will be summarized by treatment.

A listing of all vital sign assessments will be produced by treatment with notable values and assessments, collected outside of on-treatment, flagged.

2.9 Pharmacokinetic endpoints

PAS will be used in the pharmacokinetic data analysis.

PK concentrations

Descriptive statistics (n, m (number of non-zero concentrations), arithmetic mean, CV% mean, SD, median, geometric mean, CV% geo-mean, minimum and maximum) for canakinumab concentration will be presented at each scheduled time point for the PAS.

Individual concentration-time profiles for canakinumab evaluable concentrations with median will be displayed graphically for PAS on the semi-log view. In addition, the mean (+/- SD) and geometric mean concentration-time profiles for canakinumab over time will be displayed graphically for PAS on the linear and semi-log view. Only time points with $n \geq 4$ observations will be shown on the figure.

All individual plasma canakinumab concentration data will be listed by treatment for the FAS.

Handling of PK data below LLOQ or missing

All concentration values below the lower limit of quantitation (LLOQ) are set to zero by the Bioanalyst, and will be displayed in the listings as zero and flagged. LLOQ values will be treated as zero in any calculations of summary statistics, and treated as missing for the calculation of the geometric means and their CV%. The number of non-zero concentrations will also be reported in the summary statistics.

Missing values for any PK data will not be imputed and will be treated as missing.

Population pharmacokinetic analysis

If there is adequate amount of data, a mixed-effects model may be applied to the serum canakinumab concentration-time data from this study to characterize canakinumab exposure and to determine the effects of intrinsic (i.e. demographic factors) and extrinsic covariates (e.g. concomitant medications) on canakinumab exposure. If there is sufficient data for analysis, the details of the population pharmacokinetic analyses may be provided in a separate reporting and analysis plan, and the results may be reported in a separate population pharmacokinetic report.

2.10 PD and PK/PD analyses

2.10.1 Analysis of relationship between canakinumab exposure and efficacy/safety endpoints

2.10.1.1 Evaluable Ctrough concentration

For canakinumab, steady state pre-dose evaluable concentrations (as defined in PAS) collected before dose administration on and after Day 1 of Cycle 6 and beyond are steady state evaluable Ctrough.

2.10.1.2 Average evaluable Ctrough concentration

Average steady state evaluable Ctrough concentration will be used in the exposure/response and exposure/safety analyses. Each subject may have one or more steady state evaluable Ctrough concentration values. The average steady state evaluable Ctrough is defined as:

Cmin_avg_ss = geometric mean of all steady state *evaluable Ctrough* for each subject, considering the distribution of PK concentration is generally lognormal

Quartile of average trough concentration is defined as:

- <Q1: <25th percentile
- Q1 to <Q2: ≥25th percentile and <50 percentile
- Q2 to <Q3: ≥50th percentile and <75th percentile
- ≥Q3: ≥75th percentile

Subject demographic characteristics will be summarized for all subjects and by quartile of Cmin_avg_ss.

2.10.1.3 Correlations between PK exposure and efficacy measures

Correlations between efficacy endpoints and trough concentrations will be explored using the PAS. Analyses will be focused on the efficacy endpoints: DFS and OS.

The relationship between DFS/OS and PK exposure will be analyzed using Kaplan-Meier methodology.

Steady state evaluable Ctrough concentrations as defined in Section 2.10.1.2 will be included in calculating average steady state trough concentration with the following additional rules:

- for subjects who had a DFS event, only steady state trough concentrations observed on or before the date of DFS event will be included;

- for subjects who did not have a DFS event, all steady state trough concentrations observed on or before the date of last recurrence assessment prior to the cut-off date will be included.

Kaplan-Meier curves for DFS/OS will be fitted for subjects in different quartiles of average steady state evaluable Ctrough concentration (with above rules applied). DFS/OS will also be summarized by quartiles of average steady state evaluable Ctrough concentration (with above rules applied).

2.10.1.4 Correlations between PK exposure and safety measures

Correlation between safety endpoints and PK concentrations will be explored using the PAS.

AESI will also be summarized by the quartile of Cmin_avg_ss, preferred term and maximum CTC grade, for each AESI grouping, as defined in Section 2.10.1.2.

Kaplan-Meier curves for time to onset of grade 2 or worse AESI and time to onset of grade 3 or worse AESI will be fitted for subjects in different quartiles of Cmin_avg_ss concentration. Time to onset of AESI will also be summarized by quartiles of Cmin_avg_ss.

[redacted]

2.10.1.6 Immunogenicity

Immunogenicity data will be summarized using the safety set. Immunogenicity will be characterized descriptively by tabulating anti-drug antibodies (ADA) incidence at baseline, ADA incidence on-treatment, treatment induced and treatment boosted ADA.Treatment induced ADA-positive subject is defined as a subject with ADA-negative sample at baseline and at least one treatment-induced ADA-positive sample post baseline. Treatment-boosted ADA-positive subject is defined as a subject with ADA-positive sample at baseline and at least one ADA-positive sample post-baseline with a titer greater than ADA-positive baseline titer by the defined titer fold change.

The ADA-positive NAb samples at baseline and on-treatment will also be summarized. A shift table of subjects with positive or negative anti-canakinumab antibodies overall and by visit will be produced ("positive" corresponds to "worst"). A listing will be provided by subject with supporting information (i.e. ADA sample status at each time point (including titer for positive samples) and subject ADA status). In addition, a listing will also be provided for subjects with neutralizing antibodies (NAB) testing results.

2.11 Patient-reported outcomes

Three patient-reported outcomes (PRO) questionnaires will be assessed: EORTC QLQ-C30, with its QLQ-LC13 lung cancer module, and the EQ-5D-5L. QLQ-C30 and QLQ-LC13 will be considered as the primary scale. Scoring of PRO data and methods for handling of missing items or missing assessments will be handled according to the scoring manual and user guide for each respective patient questionnaire (Fayers 2001, Van Reenen 2015). No imputation procedures will be applied for missing items or missing assessments.

The FAS will be used for analyzing PRO data. For all PRO analysis, nominal *p*-values will be presented without any statistical inference since there is no adjustment for multiplicity. The baseline is defined as the last PRO assessment on or prior to randomization. In the absence of a better definition

to define clinically relevant changes in this population (lung cancer, adjuvant setting) a 10 point deterioration will be assumed to be clinically meaningful for the QLQ-LC13 and QLQ-C30 (Osoba 1998).

The primary PRO variables of interest are: time to definitive 10 point deterioration symptom scores for each of chest pain, cough and dyspnea per QLQ-LC13 questionnaire

The secondary PRO variables of interest are:

- Utilities derived from the EQ-5D-5L
- Time to first 10 point deterioration symptom scores of pain, cough and dyspnea per QLQ-LC13 questionnaire
- Time to first 10 point deterioration of global health status/QoL, shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30 questionnaire
- Time to 10 point definitive deterioration in global health status/QoL, shortness of breath, pain, role functioning, physical functioning and fatigue per QLQ-C30

The time to 10 point definitive deterioration is defined as the time from the date of randomization to the date of event, which is defined as at least 10 points absolute worsening from baseline of the corresponding scale score, with no later change below this threshold, i.e. <10 points was observed or if this worsening was observed at the last assessment for the subject, or death due to any cause (whichever occurs earlier).

The time to first 10 point deterioration is defined as the time from the date of randomization to the first onset of at least 10 points absolute increase from baseline (worsening) in symptoms scores or at least 10 points absolute decrease from baseline (worsening) in global health status/QoL and functioning scales (role functioning, physical functioning), or death due to any cause, whichever occurs earlier.

If a subject has not had an event, time to definitive deterioration or first deterioration will be censored at the date of the last adequate assessment. If deterioration occurs at the last adequate assessment, this will also be considered definitive. Subjects receiving any further anti-neoplastic therapy related to NSCLC before definitive worsening will be censored at the date of their last assessment before starting this therapy. If a definitive deterioration is observed after two or more missing assessments, subject is censored at the date of their last available questionnaire prior to the deterioration. Subjects with no baseline data will be censored at day 1.

Death is considered as a definitive deterioration event when it occurs within a period of time defined by 2 times the period between two assessments as planned in the study protocol. This avoids overestimating the time to definitive worsening in subjects dying after an irregular assessment scheme. Subjects who die after more than twice the planned period between two assessments since the last assessment are censored at the date of their last available questionnaire.

Censoring reasons for time to definitive deterioration or first deterioration event will be summarized.

All assessments will be included in the time to definitive deterioration or first deterioration analysis. The distribution will be presented descriptively using Kaplan-Meier curves. Summary statistics from Kaplan-Meier distributions will be determined, including the median time to definitive 10 point deterioration along with two-sided 95% confidence interval. Log-rank test stratified by randomization stratification factors will be performed. A Cox regression stratified by randomization stratification factors from IRT will be used to estimate the hazard ratio (HR), along with two-sided 95% confidence interval.

Sensitivity analysis of time to definitive and first deterioration with 5 points will also be performed using the same methodology.

Descriptive statistics will be used to summarize the original scores, as well as change from baseline, of the QLQ-C30, QLQ-LC13, and EQ-5D-5L at each scheduled assessment time point for each treatment arm using time windows as described in Table 2.1. Changes from the baseline by treatment arm at each visit will be ploted for all scales of QLQ-LC13, QLQ-C30 and EQ-5D-5L. Additionally, change from baseline in the scale and subscale values at the time of each assessment will be summarized. Subjects with an evaluable baseline score and at least one evaluable post baseline score during the treatment period will be included in the change from baseline analyses. Tables of compliance and completion rates over time for QLQ-LC13, QLQ-C30 and EQ-5D-5L by treatment arm will be generated.

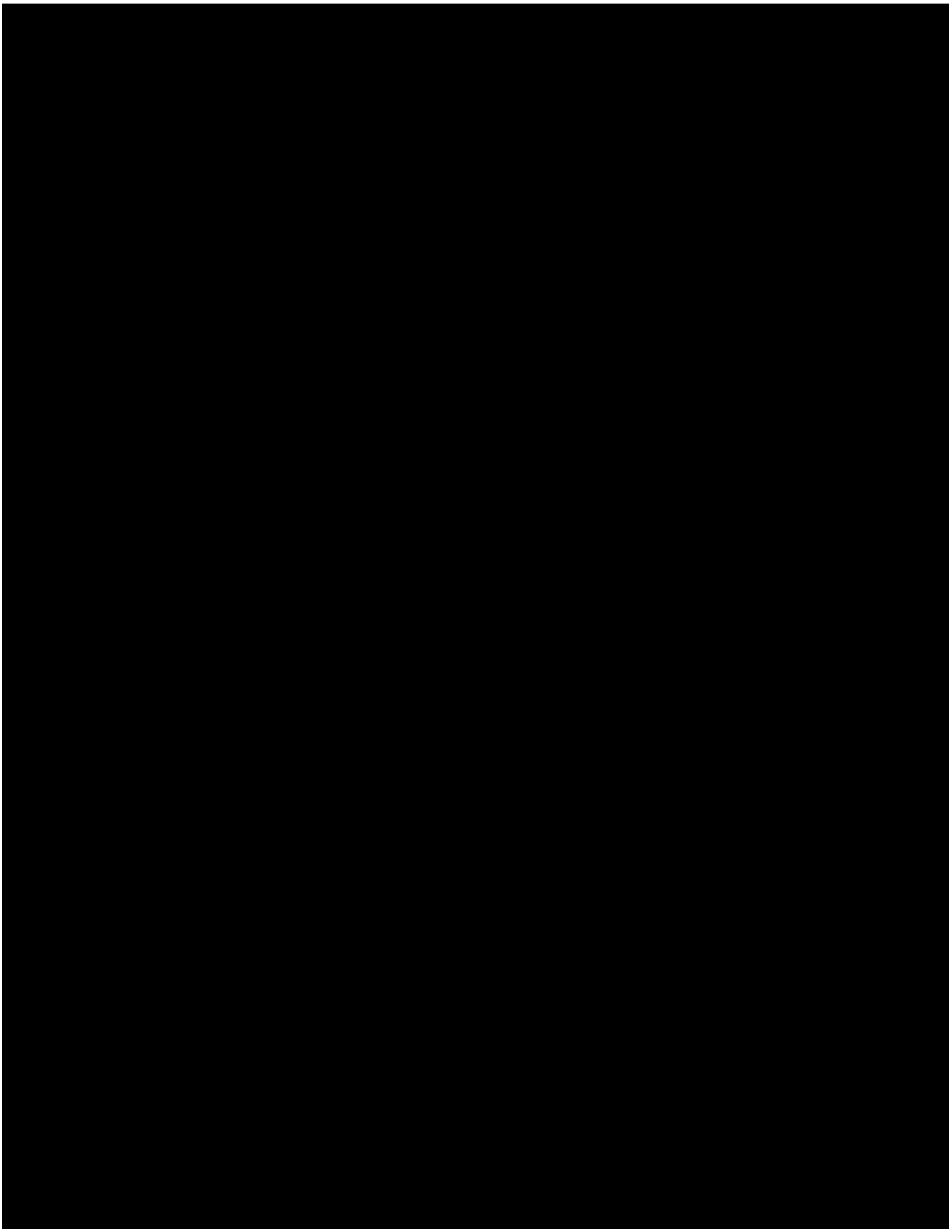
In addition, a repeated measures model for longitudinal data will be used to estimate differences in EORTC QLQ-C30/QLQ-LC13 domains as well as the VAS and utility scores of the EQ-5D-5L between treatment arms. The models will be first based on all assessments. In additional, the models will be repeated to include only assessments collected within 130 days of last study treatment. Any assessments collected after the start of further anti-tumor therapy will not be included in the models.

The modeling will mainly be done on the actual score. Note that the modeling of the change in score or the actual score is equivalent since adjustment for baseline score is considered. This repeated measures model will include terms for treatment, the stratification factors, time, and baseline value as main effects, and an interaction term for treatment by time. This analysis will be restricted to subjects with an evaluable baseline score and at least one evaluable post-baseline score. The differences in least square means between the treatment arms and corresponding 95% confidence interval at selected time points will be presented.

Time will be considered as a continuous variable expressed in weeks, i.e. considering that PRO follow a linear trend. As a first approach, an unstructured correlation matrix will be used to model the correlation within subjects. Other structures of the correlation matrix, including AR (1), will be investigated and simplified using likelihood ratio tested if appropriate.

If PRO is found not to follow a linear trend, the time variable might be considered as a categorical variable instead of continuous in the model.

[redacted]



2.13 Interim analysis

2.13.1 Disease free survival (DFS)

One interim analysis is planned after approximately 196 of the approximately 392 targeted DFS events, i.e., at approximately 50% information fraction, respectively, has been documented. The primary intent of the interim analysis is to stop early for lack of efficacy (futility). There is no intent to carry out an analysis to declare superior efficacy at the time of the first interim analysis.

An α-spending function according to a 2-look Lan-DeMets group sequential design with O'Brien-Fleming type stopping boundary will be used to construct the efficacy stopping boundaries (Lan and DeMets 1983). A user-defined gamma function with $\gamma = -3$ stopping boundary will be used as a β-spending function to determine the non-binding futility boundary. The choice of non-binding nature of the futility stopping boundary ensures that the efficacy stopping boundaries are not affected. Based on the choice of α-spending and β-spending function described above and if the interim analysis is performed *exactly* at 196 DFS events, the futility boundary expressed on *p*-value (or Z-statistic) at the first interim is calculated as 0.411 (or 0.226), i.e., the observed *p*-value has to be greater than 0.411 (or calculated Z-statistic is < 0.226) to conclude futility.

Since the observed number of events at the interim analysis may not be exactly equal to the planned 196 DFS events, the futility boundaries will need to be recalculated based on the *exact* information fraction. The observed *p*-value (or Z-test statistic) at the interim analyses will then be compared against the re-calculated futility boundary.

If the study continues to the final DFS analysis, the final DFS analysis will be performed when *approximately* 392 DFS events have been documented. If exactly 196 events are observed at the interim analysis, exactly 392 events are obtained at the final analysis, the observed *p*-value will have to be less than 0.024 (or the observed Z statistics has to be > 1.969) to declare statistical significance. With consideration of the *p*-value has been spent at interim analyses, in practice, the boundary for the final analysis will be derived accordingly from the pre-specified α-spending function such that the overall significance level across all analyses is maintained at 0.025.

Since the planned efficacy interim analysis was not performed, the efficacy boundary and observed *p*-value at the final DFS analysis will take into account only the interim analysis actually performed (i.e. futility analysis). The statistical properties of the group sequential design are summarized for DFS in Table 2-9 below.

Table 2-9 Simulated probabilities to stop for futility or efficacy at the interim or final DFS analysis

Scenario	Look	# DFS events	Simulated cumulative probabilities (%)		Simulated incremental probabilities (%)	
			Stop for efficacy	Stop for futility	Stop for efficacy	Stop for futility
Under H_{01} (HR=1)	1st IA (futility)	196	-	59.34	-	59.34
	Final	392	2.36	-	2.18	-
Under H_{a1} (HR=0.716)	1st IA (futility)	196	-	1.83	-	1.83
	Final	392	89.56	-	63.7	-
Under HR=0.8	1st IA (futility)	196	-	9.28	-	9.28
	Final	392	58.65	-	50.49	-

Note: Simulation is performed in East 6.4 with number of simulations = 10,000 and randomization seed = 1234

The interim analysis will be performed by an independent statistician (not involved with the conduct of the study). Further details are described in the DMC charter. The results of the interim analyses will be provided to the DMC by the independent statistician.

2.13.2 Key secondary endpoint: Overall survival (OS)

A hierarchical testing procedure will be adopted and the statistical tests for OS will be performed only if the primary efficacy endpoint DFS is statistically significant.

A maximum of three analyses is planned for OS; at the time of final analysis for DFS provided DFS is significant, at which point a total of approximately 318 (63%) deaths are expected, an additional IA for OS when a total of approximately 418 (83%) deaths are expected, and a final analysis for OS when approximately 504 deaths are expected.

A α-spending function according to Lan-DeMets (Lan and DeMets 1983) (O'Brien-Fleming) α-spending function along with the testing strategy outlined below will be used to maintain the overall type I error probability. This guarantees the protection of the 2.5% overall level of significance across the two hypotheses and the repeated testing of the OS hypotheses in the interims and the final analysis (Glimm 2010).

The trial allows for the stopping of the study for a superior OS result, provided the primary endpoint DFS has already been shown to be statistically significant favoring the test treatment arm. Further, the

exact nominal *p*-values that will need to be observed to declare statistical significance at the time of these analyses for OS will depend on the number of OS events that have been observed at the time of these analyses and the α for OS already spent at the time of earlier analyses.

Unblinded results from the interim analysis for DFS will not be communicated to the Sponsor's clinical team or to any party involved in the study conduct (apart from the independent statistician and DMC members) until the DMC has determined that either i) DFS analysis has crossed the pre-specified boundary for efficacy, or ii) DMC determines if the study needs to be terminated due to any cause including futility or safety reasons. Further details will be described in the DMC Charter. At the time of final DFS analysis, both DFS and interim OS analysis will be performed by the Sponsor's clinical team. Investigators and subjects will remain blinded to study treatment and all subjects will continue to be followed for OS until the final OS analysis (or earlier if OS reaches statistical significance at any of the interim analyses).

The statistical properties of the group sequential design are summarized for OS in Table 10-2 below.

Table 2-3 Simulated cumulative power to stop for efficacy on overall survival at final DFS analysis or final OS analysis

Look	Months after randomization of the first subject (approximation)	# DFS Events	Cumulative power against a hazard ratio of 0.716	# OS Events	Cumulative power[b] against a hazard ratio of 0.776
1st DFS IA (futility)	27	196 (50.0%)	-	-	-
Final DFS	42	392 (100%)	89.56%	318 (63.1%)[a]	36.26%
OS IA	52	-	-	418 (82.9%)[a]	64.09%
Final OS	63	-	-	504 (100%)[a]	79.25%

a: To be performed only if the final DFS analysis is significant

b: Power conditional on DFS being significant

Simulations performed in East 6.4 with number of simulations = 10,000 and randomization seed =1234.

2.13.3 Confidentiality of interim results

At the time of the interim analysis for DFS, interim analysis will be performed by an independent statistician. Analysis results from the interim analysis for DFS will not be communicated to the sponsor’s clinical team or to any party involved in the study conduct until the DMC has determined that the study needs to be terminated due to any cause including futility or safety reasons. Further details will be described in the DMC Charter.

At the time of final DFS analysis, both DFS and interim OS analysis will be performed by the sponsor's clinical team. All subjects will continue to be followed for OS until the final OS analysis or earlier if OS reaches statistical significance at any of the interim analyses. Any subsequent OS interim analysis and final OS analysis will also be performed by the Sponsor’s clinical team.

3 Sample size calculation

3.1 Primary analysis

The sample size calculation is based on the primary variable DFS. The hypotheses to be tested and details of the testing strategy are described in Sections 2.5.2.

Based on available data, the median DFS in the placebo arm is around 48 months (Kelly 2015). It is expected that treatment with canakinumab will result in a 28.4% reduction in the hazard rate for DFS (corresponding to an increase in median DFS from 48 months to 67 months under the exponential model assumption).

Then, in order to ensure 90% power to test the null hypothesis: DFS hazard ratio = 1, versus the specific alternative hypothesis: DFS hazard ratio = 0.716, it is calculated that a total of 392 DFS events need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, subjects randomized to the two treatment arms in a 1:1 ratio, and a 2-look group sequential design with a Lan-DeMets (O'Brien-Fleming) α-spending function and a user-defined gamma function with $\gamma = -3$ to define a non-binding futility rule at the interim analysis, using an information fraction of 50% for the interim analysis (futility only). Assuming that enrolment will continue for approximately 30 months at a rate of 50 subjects per month and a 5% annual dropout rate by the time of the final DFS analysis, a total of approximate 1500 subjects will need to be randomized to observe the targeted 392 DFS events at about 12 months after the randomization date of the last subject, i.e., 42 months after the randomization date of the first subject. These calculations were made using the software package East 6.4.

The final DFS analysis is planned to be performed when *approximately* 392 DFS events are documented to ensure the trial is not over-powered. At the time of the data cut-off for final DFS analysis, all subjects randomized prior to the data cut-off date for final DFS analysis who have had at least one post baseline recurrence assessment (i.e. all subjects have a minimal follow-up of 3 months) or discontinued study earlier will be included.

3.2 Power for analysis of key secondary variables

OS, as the key secondary variable, will be formally statistically tested, provided that the primary variable DFS is statistically significant. The hypotheses to be tested and details of the testing strategy are provided in Section 2.6.2. Based on available data, the median OS in the placebo arm is expected to be approximately 5 years (Pignon 2008). In addition, the OS result presented in (Kelly 2015) was not mature as the median was not reached at 5 years, but the overall trend indicates a median OS greater than 5 years. After taking these into consideration, the median OS for the placebo arm is assumed to be approximately 66 months. It is hypothesized that treatment with canakinumab will result in a 22.4% reduction in the hazard rate for OS, i.e., an expected hazard ratio of 0.776. It corresponds to an increase in median 66 to 85 months under the exponential model assumption. Then, in order to ensure 80% power to test the null hypothesis: OS hazard ratio = 1, versus the specific alternative hypothesis: OS hazard ratio = 0.776, it is calculated that a total of 504 deaths need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, subjects randomized to the two treatment arms in a 1:1 ratio, and a group sequential design with a Lan-DeMets α-spending function using information fractions according to actual events observed at performed analysis time points. All the efficacy boundaries will be calculated based on actual events and the number of interim analyses.

4 Change to protocol specified analyses

Not applicable

5 Appendix

This will be used later for drafting CSR Appendix 16.1.9.

5.1 Imputation rules

5.1.1 Study treatment

No imputation of the start date of infusion or end date for infusion will be applied. Complete dates are required as per eCRF.

5.1.2 AE, ConMeds and safety assessment date imputation

Table 5-1 Imputation of start dates (AE, CM) and assessments (LB, EG, VS)

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates
day, month	• If available year = year of study treatment start date then • If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY • Else set start date = study treatment start date. • If available year > year of study treatment start date then 01JanYYYY • If available year < year of study treatment start date then 01JulYYYY
day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY. o Else set start date = study treatment start date. • If available month and year > month and year of study treatment start date then 01MONYYYY • If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates (AE, CM)

Missing Element	Rule (*=last treatment date plus 130 days not > (death date, cut-off date, withdrawal of consent date))
day, month, and year	• Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	• If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	• If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

Any AEs and ConMeds with partial/missing dates will be displayed as such in the data listings.

Any AEs and ConMeds which are continuing as per data cut-off will be shown as 'ongoing' rather than the end date provided.

5.1.2.1 Other imputations

Not applicable

5.2 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death. This grading system inherently places a value on the importance of an event, although there is not necessarily proportionality among grades (a grade 2 is not necessarily twice as bad as a grade 1).

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters. The latest available version of the document based on the underlying CTCAE version 5.0 at the time of analysis will be used. For laboratory tests where grades are not defined by CTCAE version 5.0, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

5.3.1 Imputation Rules

CTC grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of WBC.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X.

The following rules will be applied to derive the WBC differential counts when only percentages are available for an xxx differential

xxx count = (WBC count) * (xxx %value / 100)

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

Corrected Calcium (mmol/L) = Calcium (mmol/L) + 0.020 [40 - albumin (g/L)]]

In order to apply the above formula, albumin values in g/L will be converted to g/dL by multiplying by 0.1), calcium values in mmol/L will be converted to mg/dL by dividing by 0.2495. For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

CTC grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) and corrected calcium will be assigned as described above for grading

5.4 Statistical models

5.4.1 Primary analysis

The LIFETEST procedure in SAS with the TIME statement including a variable with survival times and a (right) censoring variable, and with STRATA statement including variables of stratification factors and with GROUP option under STRATA statement. As an output of the procedure, the rank statistic S and variance var (S) will be obtained. Under the null hypothesis, the test statistic $Z=S/\sqrt{}$ [var (S)] is approximately normally distributed. The one-sided *p*-value will be obtained from normally distributed Z statistic.

This Section gives additional details regarding the analyses described in Section 2.5.2.

Kaplan-Meier estimates

An estimate of the survival function in each treatment arm will be constructed using Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

Median survival for each treatment arm will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of (Brookmeyer and Crowley 1982). Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula (Collett 1994).

Hazard ratio

Hazard ratio will be estimated by fitting the Cox proportional hazards model using SAS procedure PHREG (with TIES=EXACT option in the MODEL statement).

A stratified unadjusted Cox model will be, i.e. the MODEL statement will include the treatment arm variable as the only covariate and the STRATA statement will include stratification variable(s).

Hazard ratio with two-sided 95% confidence interval will be based on Wald test.

Treatment of ties

The STRATA statement in LIFETEST procedure will be used to analyze time to event data with ties. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

Checking proportionality of hazard assumption

Plots (SURVIVAL LOGSURV LOGLOGS) generated by LIFETEST procedure in SAS will be used to provide visual checks of the proportional hazard assumption.

- LOGLOGS plots log (cumulative hazard). The LOGLOG plot will show parallel curves if hazards

are proportional.

5.4.2 Key secondary analysis

Same statistical methodologies will be used as in the primary efficacy analysis

5.5 Patient reported outcomes: EORTC QLQ-C30/LC13 and EQ-5D-5L/VAS

The text below gives more detailed instructions and rules needed for programming of the analyses described in Section 2.11.

EORTC QLQ-C30 scale scores will be generated by first obtaining the raw scores adding up the item responses on the questions which make up each domain and then applying the linear transformation to the raw scores in accordance with the respective scoring manual provided by the developers. Scores in each scale will be generated if at least half of the items comprising the scale have been answered. For single item scales with missing responses and scales where less than half of the items have not been answered, scale scores will be set to missing.

The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales/single items of the QLQ-C30. The dyspnea scale of the QLQ-LC13 is the only multi item scale (all others are single item scales) and should only be used if all items comprising the scale have been answered.

For the calculation of EQ-5D-5L index value, the EQ-5D crosswalk value set for the UK using the time trade-off method will be used. If any one of the five dimensions of health state is missing, the index value will be set to missing.

A repeated measures model for longitudinal data will be used to compare the two treatment arms in terms of the domain scores over time. This longitudinal model will include terms for treatment, the randomization stratification factors, time of visit (duration in days from the time of baseline measurement to the time of a particular post baseline measurement), baseline value as fixed effects, and an interaction term for treatment by time. This analysis will be restricted to subjects with an evaluable baseline score and at least one evaluable post-baseline score. Time will be considered as a continuous variable in this analysis. As a first approach, an unstructured correlation matrix will be used to model the correlation within subjects. Other structures of the correlation matrix, including AR (1), will be investigated and simplified using likelihood ratio tested if appropriate. In particular situations, the non-convergence of the model may be caused by few subjects with assessments at later time points. The possibility of removing few assessments later than a certain time point will be investigated if appropriate.

5.6 Missing tumor assessments

The term ‘missing adequate tumor assessment’ is defined as a tumor assessment not done. For the sake of simplicity, a ‘tumor assessment’ will be referred to as an ‘assessment’ and a ‘missing adequate tumor assessment’ will be referred to as a ‘missing assessment’.

The presence and the number of missing assessments affects DFS censoring and event date.

An exact rule to determine whether there is no, one or two missing assessments is therefore needed. This rule is based:

- on the duration between LATA date and the event date
- on the duration between LATA date and the cut-off date

In this study, the protocol defined schedule of assessments is

- every 12 weeks during the first year (i.e. at weeks 12, 24, 36 and 48),
- every 26 weeks through year 2 and year 3 (i.e. at weeks 74, 100,126 and 152),
- every 52 weeks through year 4 and year 5 (i.e. at weeks 204 and 256).

The scheduled date of assessments (in weeks from randomization), protocol specified windows for assessment, and the thresholds for LARA to belong to a visit can be found in Table 5-3.

Table 5-3 Schedule for assessments, time windows and D1/D2 (weeks)

assessment schedule		Scheduled date – one week*	Schedule date (weeks from randomization)	Scheduled date + one week*	Threshold (weeks)**	Boundary of assessment*** (start, end]	D1/D2****
Every 12 weeks for the first year	1st (C5D1)	11	12	13	18	(0,18]	14/26
	2nd (C9D1)	23	24	25	30	(18,30]	14/26
	3rd (C13D1)	35	36	37	42	(30,42]	14/40
	4th (C17D1)	47	48	49	61	(42,61]	28/54
Every 26 weeks for the second and third year	5th	73	74	75	87	(61,87]	28/54
	6th	99	100	101	113	(87,113]	28/54
	7th	125	126	127	139	(113,139]	28/80
	8th	151	152	153	178	(139,178]	54/106
Every 52 weeks for the four and fifth year	9th	203	204	205	230	(178,230]	54/106
	10th	255	256	257	282	(230, 282]	54/106

* These start and end study weeks are used to determine if a visit is on schedule. For example, a visit between study week 23 and study week 25 is considered the 2nd 'on schedule' assessment.

** The threshold correspond to the mid-point between current and next visit and to the upper limit for LARA to be matched to a certain scheduled assessment e.g. if LARA is at week 45, this is after the threshold for 3rd recurrence assessment (C13D1) and before that for the 4th recurrence assessment, so the matching schedule assessment is the 4th recurrence assessment (C17D1).

*** These boundaries are defined according to assessment thresholds to calculate D1 and D2. For example: (30, 42] means any assessment falls into this time interval will be considered as the 3rd assessment visit.

**** D1 and D2 are defined per the protocol-specified recurrence assessment schedule

- The threshold D1 is defined as the protocol-specified time interval between the assessments plus 2x the protocol-allowed time window around the assessments (i.e. = 2*1 week=2 weeks).
- The threshold D2 is defined as twice the protocol-specified time interval between the assessments plus 2x the protocol-allowed time window around the assessments

Since there are changes in schedule for assessments during the 5 year tumor assessment period, D1 and D2 are defined differently depending on when LATA occurs. Here are the rules based on when LATA occurs:

- **Rule 1:** if LATA happens within 30 weeks after randomization (≤ study day 210 day), the applicable scheduled tumor assessments are the 1st and 2nd assessment (on C5D1 and C9D1), D1= 12 + 2 = 14 weeks and D2=12*2 + 2 = 26 weeks.
- **Rule 2:** if LATA happens after 30 weeks but within 42 weeks, the matched scheduled tumor assessment is 3rd assessment (on C13D1), the study period: ≥ study day 211 and ≤ study day 294, D1= 12 + 2 = 14 weeks and D2=12 + 26 + 2 = 40 weeks.
- **Rule 3:** if LATA happens after 42 weeks but within 113 weeks, the applicable scheduled tumor assessments are 4th, 5th and 6th assessments, the study period: ≥ study day 295 and ≤ study day 791, D1= 26 + 2 = 28 weeks and D2=26*2 + 2 = 54 weeks.
- **Rule 4:** if LATA happens after 113 weeks but within 139 weeks, the matched scheduled tumor assessment is 7th assessment, the study period: ≥ 792 days and ≤ 973 days, D1= 26 + 2 = 28 weeks and D2=26 + 52 + 2 = 80 weeks.
- **Rule 5:** if LATA happens after 139 weeks, the applicable scheduled tumor assessments are 8th, 9th and 10th assessments, the study period: ≥ 974 days, D1= 52 + 2 = 54 weeks and D2=52*2 + 2 = 106 weeks.

The number of missing events is defined as:

- An event after LATA+D1 weeks will be considered as having ≥1 missing assessment
- An event after LATA+D2 weeks will be considered as having ≥2 missing assessments

6 Reference

1. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
2. Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall.
3. Fayers P.M. et al. on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd edition). Published by: European Organization for Research and Treatment of Cancer, Brussels 2001.
4. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. Journal of clinical oncology. 1998;16(1):139-144.
5. Fleming TR, Harrington DP and O'Brien PC (1984). Designs for group sequential tests. Controlled Clinical Trials, 5, 348-361.
6. Kelly K et al. Adjuvant Erlotinib versus Placebo in Subjects with Stage IB-IIIA Non–Small-Cell Lung Cancer (RADIANT): A Randomized, Double-Blind, Phase III Trial, Journal of Clinical Oncology 2015; 33, no. 34 4007-4014.
7. Lan KKG, DeMets DL (1983). Discrete sequential boundaries for clinical trials. Biometrika, 70, 649-53.
8. Pignon JP et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. J Clin Oncol 2008; Jul 20; 26(21):3552-9.
9. Van Reenen M. et al. EQ-5D-5L User Guide – Basic information on how to use the EQ-5D-5L instrument. http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_ Flyers/EQ-5D-5L_UserGuide_2015.pdf. Accessed Oct 2017.
10. European Medicines Agency (2020). Guidance on the management of clinical trials during the COVID-19 (coronavirus) pandemic
11. Food and Drug Administration (2020). FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic: Guidance for Industry, Investigators, and Institutional Review Boards (2020)