Document Name: mm398010304 16.1.9 statistical analysis plan v4.0 2021nov11

Clinical	PPD 12-Nov-2021 14:38:54 GMT+0000

NonClinical	PPD 12-Nov-2021 15:33:22 GMT+0000

Clinical	PPD 12-Nov-2021 16:04:58 GMT+0000

Clinical	PPD 23-Nov-2021 15:17:03 GMT+0000

Clinical	PPD 24-Nov-2021 23:01:02 GMT+0000

Approved

Statistical Analysis Plan

MM-398-01-03-04

RESILIENT: A Randomized, Open Label Phase 3 Study of Irinotecan Liposome Injection (ONIVYDE ®) versus Topotecan in Patients with Small Cell Lung Cancer Who Have Progressed on or after Platinum-based First-Line Therapy.

Author: PPD

Version Number and Date: Final V4.0 11NOV2021

CCI

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

Statistical Analysis Plan V4.0 (Dated 11NOV2021) for Protocol MM-398-01-03-04 version 8.0.

	Name	Signature	Date
Author:	PPD		DDMMMYYYY
Position:	PPD		
Company:	IQVIA		

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

	Name	Signature	Date
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	IQVIA		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		

CCI

Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		

CCI

OUTPUT TEMPLATES SIGNATURE PAGE

Output Templates V3.0 (Dated 18OCT2021) for Protocol MM-398-01-03-04 version 8.0.

	Name	Signature	Date
Author:	PPD		DDMMMYYYY
Position:	PPD		
Company:	IQVIA		

Upon review of this document, the undersigned approves this version of the Output Templates, authorizing that the content is acceptable for the reporting of this study.

	Name	Signature	Date
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	IQVIA		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc.		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc.		
Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		

CCI

Approved By:	PPD		DDMMMYYYY
Position:	PPD		
Company:	Ipsen Bioscience, Inc		

CCI

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
1.0	23MAY2019	PPD	Not Applicable – First Version
2.0	01MAR2021	PPD	Post Part 1 lock, changes for Part 2: - Adding PP population Adding working for PRO analyses Incorporating text from PRO SAP for the secondary endpoints Adding summary's split by platinum sensitivity and PK analysis. Implementing updates from protocol amendments 4, 5, 6 and 7.
2.1	12MAR2021	PPD	Included sensitivity subgroup analysis for original platinum sensitivity
3.0	18OCT2021	PPD	Include details of supplemental interim analysis. Add new subgroup for prior immunotherapy.
4.0	11NOV2021	PPD	Updated details of the OS futility analysis to match protocol updates

CCI

TABLE OF CONTENTS

CCI

CCI [redacted]

CCI [redacted]

CCI [redacted]

1. List of Abbreviations

Abbreviation	Definition
AE	Adverse Event
ALT	Alanine aminotransferase
ANCOVA	Analysis of Covariance
AST	Aspartate aminotransferase
BICR	Blinded Independent Central Review
BLQ	Below Limit of Quantification
BOR	Best Overall Response
BUN	Blood urea nitrogen
CI	Confidence Interval
CNS	Central Nervous System
CR	Complete Response
CRF	Case Report Form
CTCAE	Common Terminology Criteria for Adverse Events
CV	Coefficient of Variation
DBP	Diastolic Blood Pressure
DCR	Disease Control Rate
DLT	Dose Limiting Toxicity
DMC	Data Monitoring Committee
DOR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer

CCI

EORTC-QLQ-C30	EORTC quality-of-life core 30 questionnaire
EORTC-QLQ-LC13	EORTC quality-of-life questionnaire lung cancer supplement
EQ-5D-5L	EuroQoL 5 dimension health status questionnaire (5 level)
HR	Hazard Ratio
ITT	Intent-to-Treat
IV	Intravenous
IWRS	Interactive Web Response System
LS	Least Square
MAR	Missing At Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
MNAR	Missing Not At Random
NCI	National Cancer Institute
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive disease or disease progression
PFS	Progression Free Survival
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetics
PR	Partial Response
PRO	Patient Reported Outcome
PT	Preferred Term
QoL	Quality of Life

CCI

[redacted]

QTcF	QT interval, Fridericia correction
RANO	Response Assessment in Neuro-Oncology
RBC	Red Blood Cell
RECIST	Response Evaluation Criteria in Solid Tumors
RS	Raw Score
SBP	Systolic Blood Pressure
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCLC	Small Cell Lung Cancer
SD	Stable Disease
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TTF	Time to Treatment Failure
UGT1A1	Uridine diphosphate Glucuronosyl Transferase 1A1
UGT1A1*28	UGT1A1 pharmacogenetic variant *28
ULN	Upper Limit of Normal
WBC	White Blood Cell

CCI

2. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for Protocol MM-398-01-03-04. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed. This analysis plan does not cover PK and Biomarker analyses which will be described in a separate plan. The Patient Reported Outcomes (PRO) will also be covered in a separate plan prepared by the Real-World Analytics and Solutions group.

This statistical analysis plan (SAP) is based on protocol version 8.0 dated 07OCT2021.

3. STUDY OBJECTIVES

This study will be conducted in two parts: an open-label, single-arm, safety run-in (Part 1) to confirm the Part 2 dose of irinotecan liposome injection, followed by a randomized, open-label comparison of irinotecan liposome injection and topotecan in patients with Small Cell Lung Cancer (SCLC).

3.1. PART 1 OBJECTIVES

3.1.1. PRIMARY OBJECTIVES

The primary objectives of Part 1 are:

- Describe the safety and tolerability of irinotecan liposome injection monotherapy administered every 2 weeks
- To determine the irinotecan liposome injection monotherapy dose (85 mg/m^2 or 70 mg/m^2 administered every 2 weeks) for Part 2 of this study.

3.1.2. SECONDARY OBJECTIVES

The secondary objectives of Part 1 are to assess the preliminary efficacy of irinotecan liposome injection (at either the 85 mg/m^2 dose level or the 70 mg/m^2 dose level) as determined by

- Objective response rate (ORR)
- Progression free survival (PFS)
- Overall survival (OS)

3.1.3. EXPLORATORY OBJECTIVES

- To describe QTcF following treatment with irinotecan liposome injection

CCI [redacted]

- To explore the biomarkers associated with toxicity and efficacy following treatment with irinotecan liposome injection in this patient population
- To describe the association between UGT1A1*28 and other UGT1A1 genotypes, SN-38 concentration and safety
- To evaluate the pharmacokinetics and the relationship between pharmacokinetic exposure and efficacy and safety following irinotecan liposome injection in this patient population
- To explore patient-reported outcomes (PROs) between arms using European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC-QLQ-C30) and European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Lung Cancer 13 (EORTC-QLQ-LC13), Patient Global Impression of Severity (PGI-S), Patient Global Impression of Change (PGI-C), and EuroQoL 5 dimensions health status questionnaire (5 level) (EQ-5D-5L)

3.2. Part 2 Objectives and Estimands

3.2.1. Primary Objectives

The primary objective of Part 2 is to compare overall survival following treatment with irinotecan liposome injection with overall survival following treatment with IV topotecan.

3.2.2. Secondary Objectives

The secondary objectives of Part 2 are to compare the following between the treatment arms:

- Progression free survival (PFS)
- Objective Response rate (ORR)
- Patient Reported Outcomes (PRO)
- Safety profile

3.2.3. Exploratory Objectives

- To explore the biomarkers associated with toxicity and efficacy following treatment with irinotecan liposome injection in this patient population
- To describe the association between UGT1A1*28 and other UGT1A1 genotypes, SN-38 concentration (irinotecan liposome injection treated patients only) and safety
- To evaluate the pharmacokinetics and the relationship between pharmacokinetic exposure and efficacy and safety following irinotecan liposome injection in this patient population
- To compare the rate of development/time to development of central nervous system (CNS) progression

CCI [redacted]

and development of new CNS metastases between treatment arms

- To compare time to treatment failure (TTF) between treatment arms
- To assess the proportion of patients with improvement in symptoms as measured by EORTC QLQ-C30/LC13 dyspnea scale
- To assess the proportion of patients with improvement in symptoms as measured by the EORTC QLQ-LC13 cough scale
- To compare the effect of irinotecan liposome injection verses topotecan on symptoms (other than dyspnea and cough), functioning and global health status as measured by EORTC-QLQ-C30, EORTC QLQ-LC13 and EQ-5D-5L

3.2.4. ESTIMANDS

The primary and secondary efficacy estimands for part 2 are described in the following table.

CCI

Table 1: Estimands

Estimand	Attributes					
	Treatment	Population	Variable (or endpoint)	Intercurrent events (ICEs)		Population level summary
				Description	Handling strategy	
Primary	Randomized irinotecan liposome injection or IV topotecan	Patients with SCLC who have progressed on or after platinum-based first-line therapy as specified in the inclusion and exclusion criteria (Protocol Section 4.1 and Section 4.2), respectively	Overall survival, defined as the time from randomization date to the date of death	1. Withdrawal from study and lost to follow up 2. Start of new anticancer treatment 3. Permanently discontinuation from study treatment due to any reasons	1. While alive strategy : Patients with ICE 1 will be censored at their last known alive date. 2. Treatment policy strategy: ICE 2 and 3 will not affect the assessment of OS.	Median OS calculated with Kaplan-Meier (KM) estimator for each treatment group. The difference in OS between treatment groups will be tested using a stratified log-rank test. The hazard ratio (HR) will be estimated in a stratified Cox proportional hazards model.
Secondary	As described for the primary estimand analysis	As described for the primary estimand analysis	Progression free survival, defined as time from randomization to first documented objective disease progression (PD) or death.	1. Start of new anticancer treatment 2. Permanently discontinuation from study treatment due to any reasons 3. Two and more consecutively missing scheduled tumor assessments	1. Hypothetical strategy : Patients with ICE 1, 2 and 3 will be censored at last tumor assessment documenting no PD or death prior to ICEs	Median PFS calculated with Kaplan-Meier (KM) estimator for each treatment group. The difference in PFS between treatment groups will be tested using a stratified log-rank test. The hazard ratio (HR) will be estimated in a stratified Cox proportional hazards model.
Secondary	As described for the primary estimand analysis	As described for the primary estimand analysis	Objective Response Rate, defined as proportion of	1. Permanently discontinuation from study treatment due to	1. Composite strategy: Patients with BOR as Not Evaluable (NE) and ICE 1 prior to	The difference in ORR between treatment groups will be compared using Cochran-Mantel-Haenszel

CCI [redacted]

			patients with a Best Overall Response (BOR) [1] characterized as either a Complete Response (CR) or Partial Response (PR)	any reasons	disease progression or death will be treated as non-responder.	method, incorporating analysis stratification factors (region and platinum sensitivity)
Secondary	As described for the primary estimand analysis	As described for the primary estimand analysis	Change from Baseline to Week 12 in EORTC-C30/LC13 dyspnea, regardless of premature study treatment discontinuation	1. Permanently discontinuation from study treatment prior to Week 12 visit due to any reasons 2. Start of new anticancer treatment 3. Death	1. Treatment policy strategy: Any values post-treatment discontinuation (until start new anticancer treatment or death) will be collected and included in the analysis. 2. Hypothetical strategy: Any values after ICE 2 and 3 will be considered missing.	The difference of Change from Baseline to Week 12 in EORTC-C30/LC13 dyspnea score between treatment groups will be tested using an ANCOVA model. Missing values will be imputed using multiple imputations under a MAR assumption.
Secondary	As described for the primary estimand analysis	As described for the primary estimand analysis	Change from Baseline to Week 12 in EORTC-LC13 cough, regardless of premature study treatment discontinuation	Same as dyspnea	Same as dyspnea	The difference of Change from Baseline to Week 12 in EORTC-LC13 cough score between treatment groups will be tested using an ANCOVA model. Missing values will be imputed using multiple imputations under a MAR assumption.

CCI [redacted]

4. STUDY DESIGN

4.1. GENERAL DESCRIPTION

This study will be conducted in two parts: an open-label, single-arm, safety run-in period (Part 1) followed by a randomized period (Part 2) assessing irinotecan liposome injection versus intravenous (IV) topotecan in patients with SCLC who have progressed on or after platinum-based first line therapy.

Part 1 is an open-label, single-arm, safety run-in evaluation of irinotecan liposome injection administered every 2 weeks, intended to confirm the anticipated Part 2 regimen (85 mg/m^2), based on safety and preliminary efficacy. A contingency has been included for evaluation of irinotecan liposome injection at a dose level of 70 mg/m^2, should the higher dose level of 85 mg/m^2 result in unacceptable toxicity. At either dose level of irinotecan liposome injection, up to 24 patients will be enrolled. The safety assessment and the corresponding expansion will be conducted according to a "6+6" design followed by enrolment of an additional 12 patients (as described below). Patients will initially be treated with irinotecan liposome injection 85 mg/m2 every 2 weeks. Dose limiting toxicities (DLTs) will be evaluated for the first 12 patients treated, during the first 28 days of treatment (or up to 14 days after the second dose of study treatment if there is a treatment delay due to non-DLT related reasons).

- Among the first 6 patients receiving irinotecan liposome injection 85 mg/m^2, (i.e. 85 mg/m^2 cohort) if ≤2 patients experience a DLT, another 6 patients will be enrolled. Otherwise, enrollment into the 70 mg/m^2 cohort will be initiated.
- Among the first 12 patients receiving irinotecan liposome injection 85 mg/m^2, if ≤ 2 patients experience a DLT, this cohort will be expanded by 12 additional patients. Otherwise, the enrollment of the 85 mg/m^2 cohort will be stopped and the enrollment of the 70 mg/m^2 cohort will be initiated.

The same design will be followed for the 70 mg/m^2 cohort.

Part 2 will be randomized and assess the efficacy of irinotecan liposome injection versus IV topotecan. Approximately 450 eligible patients will be randomized in a 1:1 ratio between the experimental arm (Arm A: based on the findings of Part 1 (85 mg/m^2 or 70 mg/m^2 of irinotecan liposome injection)) and the control arm (Arm B: IV topotecan). Patients will be randomized to the treatment arms using the Interactive Web Response System (IWRS) at a central location. Randomization will be stratified, based on the following factors:

- Region (North America vs Asia vs Other)
- Platinum sensitivity (sensitive vs resistant). Progression within 90 days from the completion of first-line platinum therapy is considered "platinum resistant" and the others "platinum sensitive".
- Performance status (ECOG 0 vs 1)
- Prior immunotherapy (yes vs no)

Only region (North America vs Asia vs Other) and platinum sensitivity (sensitive vs resistant) will be used for the

CCI

[redacted]

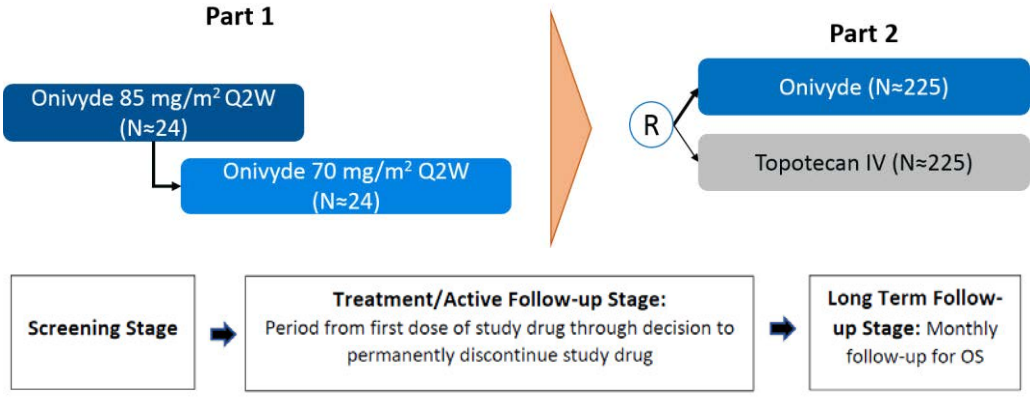
stratified efficacy analysis.

Tumor responses will be measured and recorded every 6 weeks (+/- 1 week), using the RECIST guideline (V1.1) (Parts 1 and 2) and Response Assessment in Neuro-Oncology (RANO) criteria for central nervous system (CNS lesions) in Part 2. Patients who discontinue study treatment, for reasons other than disease progression, should continue to be followed-up until radiological documentation of progressive disease with the same schedule of tumor assessments (every 6 weeks ± 1 week) until radiological documentation of progressive disease or until the start of new anti-neoplastic therapy. All patients will be followed at least monthly for survival status until death, loss to follow-up or study closure, whichever comes first.

A quality of life assessment will be performed using the EORTC-QLQ-C30, EORTC-QLQ-LC13, PGI-C, PRI-S and EQ-5D-5L (Part 2 only) questionnaires. All patients reported outcome will be administered at baseline and prior to dosing at 6 week intervals following start of treatment, at treatment discontinuation and at the 30-day follow-up visit. Adverse events (AEs) will be evaluated according to the National Cancer Institute's (NCI) Common Terminology Criteria for Adverse Events version 5.0 (CTCAE Version 5.0). For summary of AEs, events will be coded using the latest Medical Dictionary for Regulatory Activities (MedDRA) dictionary version at the start of the study.

Each part will consist of three stages: screening stage, treatment/active follow-up stage and long-term follow-up stage. Once consented, patients will enter a screening stage. Upon first dose of study treatment in Part 1 or randomization (Part 2), patients will enter the treatment/active follow-up stage. Once a decision is made to permanently discontinue the patient from study treatment, a 30-day follow-up visit will occur and the patient will enter the long-term follow-up stage. It is intended that all patients in Part 1 and Part 2 will be treated until progressive disease or unacceptable toxicity. After permanent discontinuation of study treatment, patients will return to the study site for a 30-day follow up visit.

Figure 1: Study Scheme



Note: ONIVYDE is also known as irinotecan liposome injection, liposomal irinotecan or nal-IRI

CCI

4.2. Schedule of Events

Schedule of events can be found in Section 7.1 and 7.2 of the protocol.

4.3. Changes to Analysis from Protocol

No changes to analysis from protocol.

5. Planned Analyses

The following analyses will be performed for this study:

Part 1:

- Final Analysis: after the last patient of Part 1 finished their follow-up visit and the database was cleaned and got locked.

Part 2:

Analyses for Data Monitoring Committee (DMC) meetings.

- The first DMC will occur after the 30th patient is randomized and treated for at least one cycle or permanently discontinues study drug, whichever occurs first.

Interim Analysis:

- An Interim Analysis for OS futility is planned at approximately the 29% information time, after at least 100 OS events have occurred.
- A supplementary interim analysis is planned on data collected 24 weeks after the last patient has been randomized based on the ITT Population.

Primary Analysis is planned when at least 350 OS events have occurred in part 2 .

5.1. Data Monitoring Committee (DMC)

An independent DMC will be established to monitor data in Part 2 of the study to evaluate the planned interim analyses and to make recommendation to the sponsor based on the results of the interim analyses over continuation or stoppage of the study. At the time of the interim analysis for the OS futility analysis, ORR by blinded independent central review (BICR) tumor assessments will be analyzed descriptively. The independent DMC will notify the sponsor if pre-specified criteria for ORR are met, as detailed in the DMC charter. All the criteria will be pre-specified in the DMC charter documents.

CCI

[redacted]

5.2. INTERIM ANALYSIS

An interim OS analysis for futility will be conducted when approximately 29% of the planned final number of OS events (i.e., 100 of 350 OS events) has been observed in the Part 2 intent-to-treat (P2ITT) population. At the time of the interim analysis for the OS futility analysis, ORR by blinded independent central review (BICR) tumor assessments will be analyzed descriptively in the first 200 randomized patients in part 2 (i.e. first 100 patients from each arm).

5.2.1. OS FUTILITY ANALYSIS

To control type I and type II errors, the planned interim analyses for OS will utilize an alpha and (non-binding) beta spending function approach with Hwang-Shih-Decani spending and γ parameter equal to -4.5 and -1 for each error type, respectively (Hwang, 1990). Non-binding for the futility indicates that the boundary will be constructed in such a way that it can be overruled if desired without inflating the type I or II error. Nominal alpha of 0.0001 will be spent at the supplemental interim analysis requested by FDA. Since the boundary is dependent on the number of OS events, at the time of final analysis, the actual boundary used will be re-calculated, incorporating the spending function as defined, based on the number of actual OS events and to reflect the nominal alpha spent at the supplemental interim analysis. The $P_{boundary}$ will be used as the criteria for the formal statistical inference.

The interim and primary analysis specifications per the plan are provided in the table below:

Table 2: Type I (α) and Type II (β) Error Spending for the Planned OS Analyses

		Futility				Efficacy			
Analysis	**D**	**$Z_{boundary}$**	**$p_{boundary}$**	**β spend**	**HR_{crit}**	**$Z_{boundary}$**	**$p_{boundary}$**	**α spend**	**HR_{crit}**
Interim analysis	100	-0.276	0.609	0.025	1.057	-3.181	0.001	0.001	0.529
Primary	350	1.976	0.025	0.13	0.810	-1.967	0.025	0.025	0.810

D=# of OS events at analysis. $Z_{boundary}$ is the critical test statistic value at which futility (<Z) or efficacy (>Z) would be concluded. $P_{boundary}$ is the critical one-sided p-value threshold for the comparison (>p for futility, <p for efficacy). HR_{crit} is the observed hazard ratio threshold (>HR for futility, <HR for efficacy).

The results of the interim analyses will be based on the unblinded treatment groups (open-label study). The list of outputs provided with the full set of output templates (planned for the primary analysis) will highlight which of these outputs will also be provided for the interim analysis. Derivations and definitions for the interim analysis will be based on those required for the primary analysis contained in this analysis plan, unless deviations are stated within the text.

5.2.2. ORR ANALYSIS

ORR by blinded independent central review (BICR) tumor assessments will be analyzed descriptively in the first 200 randomized patients in part 2 (i.e. first 100 patients from each arm). An estimate of the ORR and its 95% CI

CCI

will be calculated for each treatment arm. An estimate of Duration of Response (DOR) with 95% CI will be reported for each arm to support the ORR analysis. Time to objective response with 95% CI will also be reported for each arm. The independent DMC will review the efficacy and safety data and notify the sponsor if the criteria for ORR are met. All the criteria will be pre-specified in the DMC charter documents.

Sensitivity analysis will be performed for ORR by BICR according to RECIST v1.1 guideline (E.A. Eisenhauera 2009) and RANO-BM criteria for CNS lesions in the per protocol population (PP) for each treatment arm. Supplementary analysis for ORR using BICR assessments will also be performed by combining patients at the 70mg/m^2 dose level in part 1 with part 2 Irinotecan patients. Another supplementary analysis will be done for ORR using the investigator assessment response data in the first 100 part 2 patients randomized into Irinotecan liposome injection arm and also in this patient cohort combined with patients at 70mg/m^2 dose level in part 1.

An additional sensitivity analysis will be performed for confirmed response for each treatment arm by BICR. Best overall response needs to be derived based on confirmed consecutive CR and/or PR per the BIRC's assessment at least 4 weeks later with the criteria listing in table 3 of RECIST v1.1 (E.A. Eisenhauera 2009) in Appendix 6. Another sensitivity analysis will be performed for confirmed response by the investigator based on confirmed consecutive CR and/or PR using the same imaging technique at least 4 weeks later.

Safety Analysis for Interim Analysis

The safety analysis will be performed as per section 20 on the first 200 patients in the safety population and also in the safety population for all part 2 patients enrolled by the data cut-off date for this interim analysis.

Subgroup analysis for ORR

Subgroup analysis outlined in section 8.5 will be performed for this interim analysis for ORR by BICR and investigator assessment for the first 100 patients enrolled in each treatment arm in part 2.

5.3. Supplemental Interim Analysis

A supplemental interim analysis will be conducted on data collected 24 weeks after the last patient has been randomized based on the ITT population. All analysis that occurred for the interim OS analysis, per section 5.2.2, will be done for the supplemental interim analysis but on all the part 2 randomized patients rather than the first 200 randomized patients.

Nominal alpha of 0.0001 will be spent for the descriptive OS analysis at this IA.

5.4. Primary Analysis

All planned analyses identified in this SAP will be performed by IQVIA Biostatistics following Sponsor Authorization of this Statistical Analysis Plan, Database Lock, Sponsor Authorization of Analysis Sets. The primary analysis is planned when at least 350 OS events have occurred in Part 2.

Analysis of exploratory biomarker endpoint and PK analysis will be performed and reported separately.

6. Analysis Sets

Agreement and authorization of patients included/ excluded from each analysis set will be conducted prior to the database lock.

6.1. All Patients Enrolled Population

The all patients enrolled (ENR) set will contain all patients who provide informed consent for this study.

6.2. Part 1 Safety Population

Patients enrolled and treated with at least one dose of irinotecan liposome injection will comprise the Part 1 Safety Population.

6.3. Part 1 PK Population

All irinotecan liposome injection treated Part 1 patients who received at least one dose and had at least one plasma concentration will comprise the Part 1 PK Population.

6.4. Part 2 Safety Population

Patients randomized and treated with at least one dose of study medication will comprise the Part 2 Safety Population. Patients will be analyzed according to the study medication received.

6.5. Part 2 Intent-to-Treat Population (ITT)

Patients randomized in Part 2 will comprise the Intent-to-Treat Population. Patients will be analyzed according to the study medication planned.

6.6. Part 2 Per Protocol Population (PP)

Patients from the Part 2 Intent-to-Treat Population without any major protocol deviations following manual medical review will comprise the Per Protocol Population. Prior to each database lock a data review meeting will be held in order to finalize the PP population. More information will be documented in the protocol deviation plan.

Protocol deviations are collected through CTMS system, critical and major protocol violation that could exclude a

CCI [redacted]

patient from the population are, but not limited to the following:

- Patients who did not satisfy inclusion and exclusion criteria
- Patients received incorrect treatment arm
- Screening tumor scan and RECIST 1.1 assessment was not performed.
- Post baseline tumor scan and RECIST 1.1 assessment was not performed.

6.7. Part 2 PK Population

All irinotecan liposome injection treated Part 2 patients who received at least one dose and had at least one plasma concentration and no major protocol deviations affecting PK variables will comprise the Part 2 PK Population.

7. General Considerations

7.1. Reference Start Date and Study Day

Study Day will be calculated from the reference start date and will be used to show start/ stop day of assessments and events.

The following reference dates will be used in this study:

1) Informed consent date for background data
2) First dose date for safety analysis and Part 1 efficacy. The first dose of study medication is taken from the *exposure page* of the CRF with the start date and time of infusion, (day 1 is the day of the first dose of study medication). For tumor assessments in Part 2 the first dose date is also used.
3) Randomization date for Part 2 efficacy analysis

- If the date of the event is on or after the reference date then:

  Study Day = (date of event – reference date) + 1.

- If the date of the event is prior to the reference date then:

  Study Day = (date of event – reference date).

CCI [redacted]

7.2. BASELINE

Unless otherwise specified, baseline is defined as the last non-missing measurement taken prior to reference start date (including unscheduled assessments). For the PRO endpoints, baseline is defined as the last observed measurement prior to the first drug administration. In the case where the last non-missing measurement and the reference start date coincide, that measurement will be considered pre-baseline, but Adverse Events (AEs) and medications commencing on the reference start date will be considered post-baseline.

7.3. RETESTS, UNSCHEDULED VISITS AND EARLY TERMINATION DATA

In general, for by-visit summaries, data recorded at the nominal visit will be presented. Unscheduled measurements will not be included in by-visit summaries but will be used for the laboratory shift tables.

Early termination data will be mapped to the next available visit number for by-visit summaries.

Listings will include scheduled, unscheduled, retest and early discontinuation data.

7.4. WINDOWING CONVENTIONS

For summaries of PRO data, assessments will be assigned to calculated visit windows. The time windows will be exhaustive so that data recorded at any time point have the potential to be summarized. Inclusion within the visit window will be based on the actual date and not the intended date of the visit. The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls half way between the two visits. The criteria that will be used to define the PRO assessment visit are displayed in the table below.

Table 3: Visit Windowing

Visit	Window
Baseline	Study day after First Dose Date ≤1
Week 6	1 < Study day after First Dose Date ≤ 64
Week 12	64 < Study day after First Dose Date ≤ 106
Week 18	106 < Study day after First Dose Date ≤ 148
Week 24	148 < Study day after First Dose Date ≤ 190
Week 30	190 < Study day after First Dose Date ≤ 232

If there is more than one value per patient within an assessment window then the closest to the planned study day value should be summarized, or the earlier in the event the values are equidistant from the planned study day. The value at a given time point will be missing if no assessment was reported within the specified assessment window around the planned study day.

CCI [redacted]

7.5. Statistical Tests

Part 1 of the study will assess the safety and tolerability of irinotecan liposome injection monotherapy. No formal inferential analyses will determine safety and tolerability.

Part 2 of the study, the randomized open-label portion of the study, where the primary hypothesis is to test whether OS is increased in patients treated with irinotecan liposome injection. Comparisons between the two treatment groups will be done at a level of significance controlled at the one-sided 0.025 level. The secondary endpoints will also be tested at the one-sided 0.025 level

Other comparisons will be done at the significance level of (5%); confidence intervals will be 95% and all tests will be two-sided.

7.6. Common Calculations

For quantitative measurements, change from baseline will be calculated as:

- Test Value at Visit X – Baseline Value

The percent change from baseline will be calculated as:

- (Test Value at Visit X – Baseline Value)/Baseline Value

7.7. Software Version

All analyses will be conducted using SAS version 9.4.

8. Statistical Considerations

8.1. Adjustments for Covariates and Factors to be Included in Analyses

The following covariates and factors are used as strata in the analyses.

- Region: North America, Asia, Other
- Platinum Sensitivity: Sensitive, Resistant

8.2. Multicenter Studies

This study will be conducted by multiple investigators at multiple centers internationally, around 137 sites will participate. Randomization to treatment arms is stratified by factors including region.

CCI

[redacted]

When specified, statistical analysis will be adjusted for geographic region.

8.3. MISSING DATA

Missing safety data will not be imputed.

Missing efficacy data will be handled as described in section 19 of this analysis plan.

8.4. MULTIPLE COMPARISONS/ MULTIPLICITY

Key secondary endpoints in Part 2, which are: PFS, ORR, PRO endpoints (including change from baseline for dyspnea and change from baseline for cough).

Key secondary endpoints will be tested no more than once. To control the overall Type I error rate for the comparison between irinotecan liposome injection and topotecan for the primary and secondary endpoints, a hierarchical approach will be applied to the statistical testing of the secondary endpoints. The statistical inference for the first secondary endpoint of PFS will only be performed if the primary endpoint (OS), is statistically significant. The second secondary endpoint of ORR will only be tested if PFS is statistically significant. Similarly, the PRO endpoints will only be tested if ORR is statistically significant. Any parameter which is not statistically significant will be regarded as descriptive and exploratory.

Where treatment comparisons are by stratified analysis, stratification factors will be region and platinum sensitivity with classification according to the randomization designation.

If all OS, PFS and ORR are found significant, then the dyspnea and cough symptoms will be tested using the Benjamini-Hochberg correction (Benjamini&Hochberg, 1995) for a one-sided α level testing of two planned comparisons (two symptoms from the EORTC-QLQ, dyspnea and cough). Adjusted p-values will be reported, using SAS PROC MULTTEST with FDR option or equivalent algorithm. Any parameter which is not statistically significant will be regarded as descriptive and exploratory.

The Benjamini&Hochberg procedure will be applied to control for multiplicity of the statistical tests. The details of the procedure are as follows:

- P-values of the 2 tests will be sorted in ascending order:
    - Let p_1, p_2 be the raw p-values obtained when testing the 2 comparisons related to the two EORTC-QLQ symptoms
    - Let $p_{(1)}<p_{(2)}$ be the corresponding ordered p-values. The FDR-adjusted p-values $\tilde{p}_{(i)}$ are obtained as follows where m=2 in our particular testing scheme:
        -

$$\tilde{p}_{(i)} = \begin{cases} p_{(m)} & \text{for } i = m \\ \min\left(\tilde{p}_{(i+1)}, \frac{m}{i} p_{(i)}\right) & \text{for } i = m-1, \ldots, 1 \end{cases}$$

CCI

- Using PROC MULTTEST with option FDR in SAS the adjusted p-values will be obtained. These adjusted p-values will then be compared to the level of type-1 error used for testing OS as defined by the alpha-spending function.

8.5. EXAMINATION OF SUBGROUPS

Subgroup analyses will be conducted as stated in the subgroup analysis sections for Part 2 only. It should be noted that the study was not designed to detect treatment differences with high statistical power within subgroups Results will be presented if all the subcategories will be equal to or more than 5 patients.

The following subgroups will be assessed and described within the subgroup analysis sections:

- Gender:
  - Female
  - Male
- Age (years):
  - <65
  - ≥65
- Race in 2 categories:
  - Black/African American
  - All other races combined (American Indian/Alaska Native, Asian, Native Hawaiian/Other Pacific Islanders, or White)
- Region:
  - North America
  - Asia
  - Other
- Corrected Platinum Sensitivity
  - Sensitive
  - Resistant
- Brain metastases disease at baseline
  - Yes
  - No
- Prior immunotherapy
  - Yes
  - No

The following subgroup will also be assessed as a sensitivity analysis:

- Original Platinum Sensitivity:
  - Sensitive
  - Resistant

CCI [redacted]

9. Output Presentations

Appendix 1 shows conventions for presentation of data in outputs.

The templates provided with this SAP describe the presentations for this study and therefore the format and content of the summary tables, figures and listings to be provided by IQVIA Biostatistics.

10. Disposition and Withdrawals

All patients who provide informed consent will be accounted for in this study.

For Part 1 and Part 2 of the study, the number and percentage of patients screened, randomized (Part 2 only), treated and who discontinued from study and treatment, as well as their reasons for discontinuation will be summarized by part and by treatment group. The number of patients in long term follow-up will also be included. Patients who discontinued study or treatment due to COVID-19 will also be summarized.

Patient disposition and withdrawals will be presented for the all patient enrolled set.

11. Protocol Deviations

All instances of deviations from the protocol will be documented in the Clinical Trial Management System (CTMS) log which will include any programmable deviations. A deviation occurs when an investigator site, or study patient, does not adhere to protocol stipulated requirements. Deviations will be assessed and categorized as either Major or Minor. The Protocol Deviation log must be approved by Ipsen prior to database lock.

Both Major and Minor protocol deviations in the CTMS log for patients in the safety population will be summarized, for part 1 and part 2 separately, displaying the number and percentage of patients with any protocol deviation as well as the categories for the deviations. All deviations resulting in subjects excluded from the PP population will also be summarized separately.

All protocol deviations will be listed. A separate listing will be prepared for patients with protocol deviations related to the COVID-19 pandemic. A separate listing will also be prepared for COVID-19 related study disruption by subject and by investigational site as recorded in the CRF.

12. Demographic and other Baseline Characteristics

Demographic data and baseline characteristics will be presented for the Safety Part 1 and the Intent-to-Treat for the Part 2 of the study separately. The different subgroup variables will be summarized for Part 2 within the same table.

No statistical testing will be carried out for demographic or baseline characteristics.

The following demographic and other baseline characteristics will be reported for this study:

CCI [redacted]

Demographics:

- Age (years) - calculated relative to date of consent
- Sex
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islanders, White, Other, Not Reported)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported/Unknown)
- Weight (kg) at screening
- Height (cm)
- BMI (kg/m^2)
- Baseline ECOG status
- Region
- UGT1A1*28 allele status (homozygous TA7/7 – Yes or No)

Disease History:

- Time since diagnosis (weeks)
- Time since most recent progression date (weeks)
- Disease Stage at Diagnosis (limited, extensive)
- Platinum Sensitivity (sensitive, resistant)
- Prior Immunotherapy (yes, no)
- Prior Chemotherapy (yes, no)
- Symptomatic from Small Cell Lung Cancer (dyspnea, cough, fatigue, pain, weight lost in the last 60 days, other, no)
- Disease status (locally advanced, metastatic)
- Disease Site per Organ (patients may have more than one)

Patient History

- Prior Medical Condition (yes, no)
- Past Surgical Procedure (yes, no)
- Past Radiotherapy for Small Cell Lung Cancer (yes, no)
- Past Cancer Treatment Regimens for Small Cell Lung Cancer (yes, no)
- Tobacco use (never, current, former)
- Time since start of tobacco use (months)
- Time since stop of tobacco use (months)
- Duration of tobacco use (months)

The following patient characteristics will be summarized by platinum sensitivity (platinum sensitive and platinum resistant) in all part 2 patients for each treatment group.

- Age (years) - calculated relative to date of consent
- Sex
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islanders, White, Other, Not Reported)
- Disease status (locally advanced, metastatic)
- ECOG (0, 1)

CCI [redacted]

- On Treatment at cut-off (Y, N)

12.1. DERIVATIONS

- BMI (kg/ m^2) = weight (kg)/ height $(m)^2$
- Time since diagnosis (weeks) = ((date of informed consent – initial diagnosis date) + 1)/7
- Time since most recent progression date (weeks) = ((date of informed consent – date of recent progression) + 1)/7
- Time since start of tobacco use (months) = ((informed consent date- tobacco start date) + 1)/30.4375
- Time since stop of tobacco use (months) =
  - if stop date is prior to consent date then = ((informed consent date – tobacco stop date) + 1)/30.4375
  - if stop date is after consent date then = (tobacco stop date – informed consent date)/30.4375
- Duration of tobacco use (months) = ((tobacco stop date – tobacco start date) + 1)/30.4375

Imputation for partial tobacco use dates and recent progression is described in **Error! Reference source not found.**.

13. SURGICAL AND MEDICAL HISTORY

Surgical and Medical History information will be presented for the Safety Population for Part 1 and Part 2 separately.

- Prior Medical Condition will be coded using the most current MedDRA dictionary.
  - Data captured on the *Pertinent Prior Medical Conditions* page of the CRF will be presented by SOC (System Organ Class) and PT (Preferred Term).
  - A separate summary will be provided on the cancer related prior medication conditions.
  - A by patient listing will also be provided.
- Past Surgeries and Medical or Surgical Procedures will be coded using the most current MedDRA dictionary.
  - Data captured on the *Past Surgeries* and *Medical or Surgical Procedures* page of the CRF will be presented by SOC (System Organ Class) and PT (Preferred Term).
  - A by patient listing will also be provided.
- Past Radiotherapy
  - The number of sites of radiotherapy, the total dose received, time since last radiotherapy (derived as the date of first dose of study drug – end date of last radiotherapy), treatment intent (adjuvant, neo-adjuvant, palliative, curative) and best response will be summarized descriptively
  - A by patient listing will also be provided.

- Previous Small Cell Lung Cancer Treatment Regimens
  - Cancer therapy agents captured on the *Previous Small Cell Lung Cancer Treatment Regimens* page of the CRF will be presented by regimen name.
  - A by patient listing will also be provided.

14. CONCOMITANT PROCEDURES

Concomitant Procedures will be presented for the Safety Population Part 1 and Part 2.

- Concomitant Procedures will be coded using the most current MedDRA dictionary.
  - Concomitant Procedures are those which started prior to or at Screening and are ongoing at the date of Screening.
  - Data captured on the *Concomitant Procedures* page of the CRF will be presented by SOC (System Organ Class) and PT (Preferred Term).
  - Concomitant Procedures will be presented in a by patient listing.

15. CONCOMITANT RADIOTHERAPY

Concomitant Radiotherapy will be presented for the Safety Population Part 1 and Part 2.

- The number of sites of radiotherapy, the total dose received, the treatment indication and the best response after concomitant radiotherapy will be summarized in a table.
  - Concomitant Radiotherapy are those which started prior to or at Screening and are ongoing at the date of Screening.
  - Concomitant Radiotherapy will be presented in a by patient listing.

16. PRIOR AND CONCOMITANT MEDICATIONS

Medications will be presented for the Safety Population Part 1 and Part 2 and coded using the most current WHOdrug dictionary.

See Appendix 2 for handling of partial dates for medications, in the case where it is not possible to define a medication as prior, concomitant, or post treatment, the medication will be classified by the worst case; i.e. concomitant.

- 'Prior' medications are medications which started and stopped prior to the first dose of study medication.
- 'Concomitant' medications are medications which:
  - started prior to, on or after the first dose of study medication and started no later than the last dose of study drug taken,
  - AND ended on or after the date of first dose of study medication or were ongoing at the end of the study.

CCI

- 'Post' medications are medications which started after the last dose of study medication.

The number and percentage of patients taking medications will be summarized by treatment group for the prior and concomitant medications using drug class (ATC level 2, i.e. therapeutic group) and preferred term. Medications will be sorted alphabetically by drug class and by preferred term within drug class. If a patient has more than one qualifying record for a medication, the patient will be counted only once.

Post treatment will be included in listings. Post-treatment anti-cancer procedures will be monitored, flagged and listed.

17. EXTENT OF DISEASE AT BASELINE

Number of patients with target lesions only, non-target lesions only, and both target and non-target lesions will be summarized. Individual disease site for target or non-target lesions will be summarized.

18. STUDY MEDICATION EXPOSURE

For Part 1 and Part 2 the following will be summarized based on the Safety Population.

The extent of exposure will be summarized for Irinotecan liposome and Topotecan (for Part 2 only) as follows:

- Number and percentage of patients beginning 1, 2, 3, 4, 5, 6,..., > x cycles
- Number of cycles started as a continuous variable
- Total Actual Dose received defined as actual dose (mg), summed across all administrations
- Total planned dose is defined as the cumulative planned dose in each cycle (regardless of dose skipped or treatment discontinuation). If a patient discontinues the treatment mid cycle the planned dose in following dosing visits in that cycle will be imputed with the last available value within the cycle.
- Relative Actual Dose = [Total actual dose (mg) divided by Total intended dose (mg)] X 100, where total intended dose is defined as the planned dose as prescribed which is summed from first to last intake date irrespective of missed, interrupted, or unplanned reduced doses. Up to three dose reductions of irinotecan liposome injection or up to two dose reductions of topotecan per patient are permitted due to toxicities and will be considered in the planned dose.
- Total Exposure (weeks) derived as (last dose date – first dose date +14)/7 for patients in the Irinotecan arm or (last dose date – first dose date + 21)/7 for patients in the Topotecan arm.
- Relative total dose intensity (%) will be summarized as actual total dose intensity divided by planned total dose intensity * 100.
    - The actual total dose intensity (mg/m^2/week) = (cumulative actual dose in mg/m^2) / (treatment duration), where the treatment duration is derived as (last dose date – first dose date +14)/7 for patients in the Irinotecan arm or (last dose date – first dose date + 21)/7 for patients in the Topotecan arm.

CCI
[redacted]

- The planned total dose intensity (mg/m²/week) = (cumulative planned dose in mg/m²) / (treatment duration), and treatment duration (weeks) = 6 weeks * the number of cycles that a patient receives a dose at. If a patient discontinues the treatment mid cycle then the planned dose and the BSA in following dosing visits in that cycle will be imputed with the last available value within the cycle.

Treatment delays and dose modification

Definitions:

A dose reduction is defined for Irinotecan Liposome, as an actual dose that is reduced due to hematological toxicities such as neutropenia, leukopenia, or thrombocytopenia Grade 3 or 4 Neutropenic fever and all nonhematological toxicities (except asthenia and anorexia), Grade 3 or 4. Up to three dose reductions are allowed. Dose reduction of topotecan is allowed at investigator discretion as clinically indicated after the initial full starting dose of 1.5 mg/m2, up to two dose reductions in topotecan are allowed. Dose reductions are recorded in the CRF by the investigator.

A dose delayed/interrupted/withdrawn will be documented by the investigator and recorded in the CRF.

For Part 1 and Part 2 of the study, the frequency and percentage of the following will be summarized:

- Patients with at least one dose reduction overall and by cycle
- Patients with at least one dose reduction due to an adverse event
- Total number of dose reductions per patient
- Number of dose reductions overall and by cycle
- Patients with dose delayed, interrupted and withdrawn overall and by cycle.

19. EFFICACY OUTCOMES

19.1. PRIMARY EFFICACY

This section only covers Part 2 primary efficacy, as for Part 1 the primary objective is Safety.

19.1.1. PRIMARY EFFICACY VARIABLE& DERIVATION

OS is defined as the time from randomization date to the date of death from any cause. In the absence of confirmation of death, survival time will be censored at the last date the patient is known to be alive. The date last known to be alive will be taken from the *Overall Survival* CRF form, but checks will also be made of laboratory sample dates, adverse event start and stop dates, concomitant medication start and stop dates, as well as normal visit/follow-up dates in order to obtain the last recorded date that a patient was known to be still alive. Patients with no post-randomization information will be censored on the date of randomization. OS calculations (months) will be the following:

CCI [redacted]

[Death/Censoring Date – Randomization Date + 1]/30.4375

19.1.2. Primary Analysis of Primary Efficacy Variable

The null hypothesis is that the OS is the same for the two treatments. The alternative hypothesis is that Irinotecan Liposome Injection (ONIVYDE ®) has a longer OS compared to Topotecan IV.

The primary efficacy analysis will be performed for the Part 2 Intent-to-Treat Population.

The primary analysis will be performed using a stratified log-rank test (stratified by region (North America vs. Asia vs. Other) and platinum sensitivity (sensitive vs. resistant) comparing the OS between two treatment groups with level of significance controlled at the one-sided 0.025 level. Kaplan-Meier methods will be used to estimate median OS (with 95% confidence intervals) and to display OS time graphically. A stratified Cox proportional hazards model, (stratified by region and platinum sensitivity) will be used to estimate the treatment hazard ratio and its corresponding 95% confidence interval.

The tabular summary for OS will present the frequency distribution of OS status (died or censored), frequency distribution of the reasons for censoring at the time of the data cut-off (alive, withdrawn from study, lost to follow-up), Kaplan-Meier estimates of median survival time, quartiles and Kaplan-Meier estimates of OS rate for months 3, 6, 9, 12 or more.

A figure displaying the Kaplan-Meier survival curves for each treatment will be presented.

19.1.3. Sensitivity Analysis of Primary Efficacy Variable

No sensitivity analysis planned.

19.1.4. Analyses of Subgroups

To assess the robustness of results, subgroup analyses of overall survival will be performed to examine differences in the effects of treatment in different segments of the study population. The factors explored will include those identified in Section 8.5.

A separate Cox regression analysis on overall survival will be conducted for each factor using a model containing treatment group, the factor, and the treatment group-by-factor interaction. For categorical factors, the hazard ratio will be estimated within each subgroup from the model (e.g., separately for males and females in the assessment of sex) and a test of homogeneity will be conducted (the test of the treatment group-by-factor interaction) to assess the consistency of the treatment group effect among subgroups.

Kaplan-Meier plots of estimated survival function will be created for each combination of treatment arm and subgroup. For each pairwise treatment arm comparison, a summary of the subgroup analyses will be presented in forest plots, showing hazard ratio estimates and unadjusted 95% CIs for each subgroup. These analyses will be performed on the Part 2 Intent-to-Treat population.

CCI

19.2. SECONDARY EFFICACY

The secondary efficacy analyses will be performed on the Safety Population for Part 1 patients and on the Intent-to-Treat Population for Part 2 patients. For Part 1 patients, only descriptive statistics will be provided.

19.2.1. SECONDARY EFFICACY VARIABLES & DERIVATIONS

19.2.1.1. Progression-free survival

The determination of PFS will be based on BICR tumor assessment using RECIST version 1.1, with PFS based on investigator tumor assessments as a sensitivity analysis. No separate programming based on RECIST for the responses will be performed. All PFS calculations (months) will be the following:

[Progression/Death/Censoring Date – Randomization Date + 1]/(30.4375)

For **Part 1** the progression-free survival time is the time from first dose of study drug to the first documented objective disease progression (PD) using RECIST Version 1.1 or death due to any cause, whichever occurs first.

For **Part 2** the progression-free survival is the time from randomization in Part 2 to the first documented objective disease progression (PD) using RECIST Version 1.1 (or RANO-BM criteria for CNS lesions) or death due to any cause, whichever occurs first. Determination of PFS will be per local radiology review and/or investigator assessment.

Censorship: In general, patients are censored at the date of the last objective disease assessment that verified lack of disease progression if they are last known to be alive, on-treatment or within 30 days following treatment discontinuation and progression-free. Censoring is explicitly described below and in table 4.

- Patients with inadequate/incomplete baseline disease assessment will be censored at the date of randomization.
- Patients not reassessed after randomization will be censored at the date of randomization unless death occurred within the time window for the first two assessment visits, provided the patient is still on-study treatment or <= 30 days after study treatment discontinuation
- Patients with at least one on-study disease assessment who discontinue treatment without documented disease progression or death are censored at the date of the last objective disease assessment documenting no progression.

There are two exceptions.

- If objective progression or death is documented <=30 days after study treatment discontinuation then progression or death is an event
- If a new anti-cancer treatment is started prior to objective progression and <= 60 days after study treatment discontinuation, then censorship is at the date of the last objective disease assessment that verified lack of disease progression prior to the new treatment

Patients with documentation of progression or death after an unacceptable long interval (i.e. 2 or more missed or indeterminate/not evaluable assessments) since the last tumor assessment will be censored at the time of last

CCI [redacted]

objective assessment prior to the missed/indeterminate/not evaluable assessments.

Table 4: Evaluation of Documented Disease Progression and Censoring

Situation	Date of	Outcome
Inadequate baseline tumor assessment	Randomization	Censored
No post-baseline assessment prior to first scheduled assessment		Censored
Death or documented PD after ≥ 2 consecutively missed and/or not evaluable scheduled tumor assessments	Last tumor assessment documenting no PD (prior to the missed/not evaluable assessments)	Censored
New anticancer treatment prior to PD or death	Last tumor assessment documenting no PD prior to new anticancer treatment	Censored
Study treatment discontinuation due to toxicity, undocumented progression, or other reason and no documented PD or death prior to treatment discontinuation	Last tumor assessment documenting no PD prior to study discontinuation	Censored
Alive, on-treatment, and no documented PD	Last tumor assessment documenting no PD	Censored
Documented disease progression within 2 consecutive scheduled tumor assessments of baseline or the last non-PD tumor assessment	First tumor assessment documenting PD	Progression (Event)
Death without documented PD and within 2 consecutive scheduled tumor assessments of baseline or the last non-PD tumor assessment	Death	Death (Event)

Note: Tumor assessments are planned every 6 weeks, with a window of +/-7 days for each assessment. Thus, a patient will be considered to have missed ≥2 consecutive assessments if the time between two evaluable assessments is > 14 weeks (98 days). Follow-up scans can be performed up to 4 weeks from last dose date according to the protocol.

19.2.1.2. Objective Response Rate (ORR)

Best Overall Response (BOR) is recorded from randomization for Part 2 or date of first treatment dose for Part 1 until documented disease progression or death. ORR is defined as the proportion of patients with a BOR characterized as either a Complete Response (CR) or Partial Response (PR) as defined according to RECIST V1.1 guideline in Part 2 relative to the total number of patients in the ITT population. Per RECIST V1.1 guidelines,

CCI [redacted]

patients with documented PR or CR will not need confirmatory tumor assessments to document tumor response. Patients without baseline or post-baseline tumor assessments will be considered non-evaluable for this analysis.

Patients with a BOR of CR or PR will be considered a responder for the ORR analysis, while patients with a BOR of SD, PD, Not Evaluable or Not Done will be considered a non-responder.

The ORR will be determined for the Part 1 and the Part 2. For Part 2 all patients will be reported using the RECIST 1.1 criteria for the ORR. For patients who have CNS disease at baseline the RANO criteria will be used for the CNS lesions.

The determination of ORR in part 2 will be based on BICR tumor assessment using RECIST version 1.1, with ORR based on investigator tumor assessments as a sensitivity analysis.

The duration of objective response will also be summarized. It is defined as follows: Duration (weeks)= ((Date of PD/death/censoring) minus date of OR (CR or PR whichever occurs first))+1)/7. If progression or death has not been documented, a patient's duration of OR will be censored at the date of the last valid tumor assessment. Patients with a new anti-cancer therapy after the first response but prior to documented PD or death will be censored at the last tumor assessment prior to new anti-cancer therapy.

The time to objective response will also be summarized. This is the time from randomization to the first objective tumor response observed for patient who achieved a CR or PR and is defined as follows: Time to OR (weeks)= ((Date of OR/censoring) minus date of randomization+1)/7. If objective response has not occurred, a patient's time to OR will be censored at the date of the last valid tumor assessment. Patients with a new anti-cancer therapy prior to OR will be censored at the last tumor assessment prior to new anti-cancer therapy.

19.2.1.3. Overall Survival Part 1

The OS for the Part 1 which is defined as the time from first dose of study drug to the date of death from any cause. In the absence or confirmation of death, survival time will be censored at the last date the patient is known to be alive. The date last known alive will be taken from the *Overall Survival* CRF form, but checks will also be made of laboratory sample dates, adverse event start and stop dates, concomitant medication start and stop dates, as well as normal visit/follow-up dates in order to obtain the last recorded date that a patient was known to be still alive. Patients with no post-first dose of study drug information will be censored on the date of first dose of study drug. OS calculations (months) will be the following:

[Death/Censor Date – First Study Drug Dose Date + 1]/30.4375

19.2.1.4. Patients with Symptom of cough and dyspnea

The EORTC QLQ-Lung Cancer 13 (LC13) is a 13-item, lung cancer-specific module to be used in conjunction with the EORTC QLQ-C30. The LC13 covers typical lung cancer-associated symptoms such as cough and dyspnea. Score will range from 0 to 100 once transformed. A high score represents a high level of symptomatology/problems. The instrument and scoring algorithm for EORTC QLQ C30 can be found in appendix 4 and for EORTC QLQ LC13 can be found in appendix 5.

The change from baseline to week 12 in EORTC QLQ-C30/LC13 dyspnea and in EORTC-QLQ-LC13 cough will be computed.

CCI [redacted]

19.2.2. Analysis of Secondary Efficacy Variables

19.2.2.1. Analysis of Progression Free Survival

For Parts 1 and Part 2 Kaplan-Meier methods will be used to estimate median and quartiles PFS (with 95% confidence intervals) and estimates of PFS at 3, 6, 9 and 12 months and to display PFS time graphically.

The tabular summary for PFS will present the frequency distribution of PFS status (event or censored) and the frequency distribution of the reasons for censoring at the time of the data cut-off.

For Part 2 the difference in PFS between treatments will be evaluated using a stratified log-rank test (stratified by region (North America vs. Asia vs. Other) and corrected platinum sensitivity (sensitive vs. resistant)). A stratified Cox proportional hazards model will be used to estimate the PFS hazard ratio and its corresponding 95% confidence interval. The Cox proportional hazard model will include the randomization stratification factors (region and platinum sensitivity).

For the Part 2, the following sensitivity analysis will be performed:

- Early discontinuation sensitivity analysis: Patients without documented PD who had subsequent therapy or were discontinued from treatment due to non-progression reasons are considered as PD at the time of these events
- Missing data sensitivity analysis: date of PD is backdated to the expected date of the first missed tumor assessment if one or more tumor assessments were missing immediately preceding PD

19.2.2.2. Analysis of Objective Response Rate

An estimate of the ORR and its 95% CI will be calculated. The difference in ORR between treatment groups will be compared using Cochran-Mantel-Haenszel method, incorporating analysis stratification factors (region and platinum sensitivity).

Analyses will be provided for Part 1 and Part 2. For Part 1, no treatment comparisons will be done, only descriptive summaries will be presented.

The duration of objective response will also be compared using a log-rank test stratified by region and platinum sensitivity.

19.2.2.3. Overall Survival for Part 1

The Part 1 OS analysis will be based on the Part 1 Safety Population. Kaplan-Meier methods will be used to estimate median OS (with 95% confidence intervals) and to display OS time graphically.

The tabular summary for OS will present the frequency distribution of OS status (died or censored), frequency distribution of the reasons for censoring at the time of the data cut (alive, withdrawn from study, lost to follow-up), Kaplan-Meier estimates of median survival time and Kaplan-Meier estimates of OS rate for months 3, 6, 9, 12.

A figure displaying the Kaplan-Meier survival curves will be presented.

CCI

19.2.2.4. Summary of Tumor Response

The following tumor response categories will be summarized by platinum sensitivity (platinum sensitive and platinum resistant) for all patients in part 2 split by treatment group. Tumor response according to RECIST 1.1 and RANO criteria will be summarized separately for both the Investigator response in the CRF and the independent reviewer response.

- Best Overall Response (BOR) (CR, PR, SD, PD, NE)
- ORR (Yes, No)
- Disease Control Rate (DCR) (Yes, No) where DCR is defined as proportion of patients with CR, PR or SD for at least 12 weeks
- Duration of response

19.2.2.5. Patients with Symptom of Cough and Dyspnea

A table with descriptive statistics for the change from baseline in dyspnea and cough and corresponding 95% confidence intervals (CI), will be presented by treatment group for each post-baseline assessment.

Change from baseline to Week 12 in the patient's perceived dyspnea and cough scores will be analysed by means of multiple imputation under the missing at random (MAR) assumption in the primary analysis and by means of pattern-mixture model under the missing not at random (MNAR) assumption in the sensitivity analysis.

Multiple Imputation (MI)

The analysis will include observed values while adhering to the initial randomized treatment as well as observations after discontinuation of the randomized treatment.

It is acknowledged that patients will discontinue treatment at different time points during the study and that this is an important time with regards to PRO data collection. In order to include the 30 days safety follow-up visits in the analysis, an analysis visit variable will be derived for each patient and assessment, so that the 30 days safety follow-up visits will be re-mapped and will be assigned a sequential number.

Patients having missing baseline, Week 6, or Week 12 scores will have their missing score imputed under the MAR assumption using the MI methodology as follows.

We distinguish between monotone and non-monotone missing values. Non-monotone missing values are values missing intermittently, where a patient may miss some PRO assessments but has PRO assessments for the same score later on. Monotone missing values are such that once a value is missing for a given score, no subsequent values for this score are available. Any given patient may have a combination of non-monotone and monotone missing values.

The imputation will be performed using the following sequence:

Step 1: Non-monotone missing data will be imputed first. Non-monotone missing values are assumed to be MAR and will be multiply-imputed using a Markov Chain Monte Carlo (MCMC) method of Proc MI in SAS. The imputation model will include the treatment group, stratification factors (region, platinum sensitivity), PRO score at baseline and at all other timepoints (e.g., Week 6 and Week 12), age, sex, prior chemotherapy, and disease status (locally advanced vs. metastatic).

Step 2: The imputation of baseline data. The imputation will be done using a predictive mean matching multiple

imputation model and a method of Fully Conditional Specification as implemented in the SAS Procedure MI (FCS statement). The imputation model will include the treatment group, stratification factors (region, platinum sensitivity, performance status, prior immunotherapy), PRO score at baseline, age, sex, prior chemotherapy, and disease status (locally advanced vs. metastatic).

Step 3: The imputation of monotone missing data. The imputation will be done only one time-point at a time, in chronological order, using a predictive mean matching multiple imputation model. The imputation model will include the treatment group, stratification factors (region, platinum sensitivity, performance status, prior immunotherapy), PRO score at baseline, the PRO score at the corresponding timepoint, age, sex, prior chemotherapy, and disease status (locally advanced vs. metastatic).

The number of closest observations used to sample an imputed value by the predictive mean matching method will be 5 (SAS default setting). One hundred (100) multiply-imputed datasets will be generated. The random seed used in the PROC MI procedure will be 12345.

Each imputed dataset will be analysed using an Analysis of Covariance (ANCOVA) model (not repeated measures, and with no random effects) with the following covariates using the SAS Proc Mixed procedure: treatment group, stratification factors (region, platinum sensitivity), and PRO score at baseline.

The results from multiple imputed datasets will be combined using Rubin's rule as implemented in the SAS Procedure MIANALYZE.

The same model will be considered for dyspnea and cough.

Sensitivity Analysis (Pattern Mixture Model (PMM))

A sensitivity analysis will be performed by means of pattern-mixture model by reasons of discontinuation using multiple imputation as described by O'Kelly and Ratitch.

Similar to the primary analysis, non-monotone missing data will be imputed first, followed by the imputation of baseline and then monotone missing data. Similar imputation models as described above will be considered for the non-monotone missing data and baseline missing data.

To impute monotone missing values, we define patterns depending on reason and timing of missingness as follows:

- Pattern 1: missing values before or at Week 12 after study treatment discontinuation due to adverse events (AEs) or progression but prior to death;
- Pattern 2: missing values before or at Week 12 with missingness that does not satisfy conditions of Pattern 1.

The following assumptions will be made for the missing data in each pattern:

For pattern 1, imputations will be performed for both treatment arms using a reference-group approach, where the reference group is defined as patients from the same treatment group who discontinued treatment but have post-discontinuation data corresponding to Week 12.

For pattern 2, data will be assumed to be MAR in both treatment arms.

One hundred (100) multiply-imputed datasets will be generated. The random seed used in the PROC MI

CCI [redacted]

procedure will be 12345.

Each imputed dataset will be analysed using an ANCOVA model (not repeated measures, and with no random effects) with the following covariates using the SAS Mixed procedure: treatment group, stratification factors (region, platinum sensitivity), and PRO score at baseline.

The results from multiple imputed datasets will be combined using Rubin's rule as implemented in the SAS Procedure MIANALYZE.

The treatment differences will be estimated from the final model with LS-means differences. The LS mean treatment difference, 95% CI, and p-value will be presented.

The above analysis will be performed only if sufficient numbers of patients have PRO outcomes collected after study drug discontinuation. If this is not the case, the PMM using sequential modelling with multiple imputation and delta-adjustment (described below) will be performed.

A delta-adjustment analysis assumes that patients who discontinued the study had outcomes that were worse than otherwise similar patients that remained in the study.

The difference (adjustment) in outcomes between dropouts and those who remain will be implemented as a shift in location (that is, a shift in mean of a patient's imputed PRO score).

To impute monotone missing values, we define patterns depending on reason and timing of missingness as follows:

- Pattern 1: missing values before or at Week 12 after study treatment discontinuation due to adverse events (AEs) but prior to progression/death;
- Pattern 2: missing values before or at Week 12 after progression but prior to death;
- Pattern 3: missing values before or at Week 12 with missingness that does not satisfy conditions of Patterns 1 to 2.

The following assumptions will be made for the missing data in each pattern:

For pattern 1, a clinically plausible assumption is that toxicity will produce some effects on patients' PROs. Therefore, it is assumed that patients who discontinue the study treatment due to AEs at a given timepoint would, on average, have their unobserved score right after an AE that led to treatment discontinuation worse by some amount δ compared to the observed scores of patients at that timepoint. This assumption applies to both treatment arms.

For pattern 2, a conservative assumption is that PRO scores will deteriorate continuously after disease progression. It will be assumed that unobserved scores after disease progression would be worse by some amount δ at each visit after progression compared to patients with available data (and no progression). This assumption will be applied to both treatment arms.

For pattern 3, data will be assumed to be MAR in both treatment arms.

The imputation of monotone missing data will be done sequentially for each scheduled PRO assessment visit, k=k1,...,K (where k1 corresponds to the first visit up to Week 12 with any monotone missing data, and K corresponds to Week 12) as follows:

1. Impute all monotone missing values at visit k using an MAR-based multiple imputation predictive mean

CCI [redacted]

matching model including the effects for baseline covariates as listed for the imputation model of non-monotone missing data above and PRO values at each schedule assessment time point up to (k-1). The imputation model will be estimated from all study patients with available data at visit k, regardless of treatment discontinuation, including those imputed in step (a) above.

2. Adjust the monotone imputed values at visit k by amount δ in the direction of score worsening for the following cases:
   - For pattern 1, for patients for whom visit k is the first scheduled PRO assessment time point after study treatment discontinuation.
   - All values in pattern 2, that is all monotone imputed values where visit k is after patient's disease progression.

The δ adjustment will be performed for both treatment arms. The adjustment will be implemented using MNAR statement in Proc MI, using timepoint-specific flags to designate values requiring δ adjustment.

The above steps (a)-(b) are performed for each visit k, before proceeding with the imputations of the next visit (k+1).

For pattern 1, this imputation strategy matches the single event of toxicity leading to the study treatment discontinuation with a single worsening (δ adjustment) in the PROs right after the study treatment discontinuation; the worsening imposed at this timepoint affects subsequent timepoints via the correlations between visits that are estimated in the imputation model. Since the imputation model is estimated from the data of patients who had their PROs collected, the assumption of this analysis is that the future values after the one-time δ adjustment would resemble the outcomes of patients who had a similar trajectory prior to the patient's missing value.

For pattern 2, this imputation strategy assumes that PRO scores will deteriorate (by some amount δ) even further after progression in both treatment arms. For this pattern, the δ adjustment will be applied at each visit, in addition to using previously δ-adjusted values as predictors of subsequent PROs. It explicitly enforces an assumption that progression is associated with worsening that persists throughout subsequent timepoints compared to patients without progression and does not rely exclusively on the estimated correlations between timepoints embedded in the imputation model in order to propagate the effect of δ through time.

The δ values we will impose (worsening) for each PRO score for patterns 1 and 2 correspond to the half of the established clinically meaningful threshold for change (e.g., 5 points).

A total of 100 multiply-imputed datasets will be created for each PRO. The random number generator seed for the imputation of non-monotone missing values using MCMC will be 5414, and the random seed for imputation of monotone missing values will be 5414+k, for k=1, 2, ... for each sequential visit with monotone missing data.

An ANCOVA model (not repeated measures, and with no random effects) with the same covariates as in the primary analysis will be performed at Week 12 using the SAS Mixed procedure.

The SAS MIAnalyze procedure will be used to combine the results of these analyses for the imputations. For a more detailed description of the implementation MNAR imputation, see Ratitch B and O'Kelly M7.

The treatment differences will be estimated from the final model with LS-means differences. The LS mean treatment difference, 95% CI, and p-value will be presented

19.3. EXPLORATORY EFFICACY

19.3.1. EXPLORATORY EFFICACY VARIABLES & DERIVATIONS

19.3.1.1. QTcF following treatment with irinotecan liposome injection

For Part 1 patients the QTc will be collected and derived using Fridericia's correction method as follow:

$$QTc = \frac{QT}{\sqrt[3]{RR}}$$

19.3.1.2. Time to Treatment Failure (TTF)

Treatment failure, assessed in Part 2 only, is defined as treatment discontinuation for any reason such as disease progression, death or toxicity. Time to treatment failure is computed as:

TTF (months)= (earliest date of [PD, death, study treatment discontinuation] – date of randomization + 1)/30.4375.

The treatment-failure date for patients who discontinue for reasons other than RECIST Version 1.1 progression will be the date of the last study drug administration. Patients who have not discontinued treatment prior to the data cut-off date will be censored at the date of last tumor assessment documenting no objective progression.

19.3.1.3. EORTC-QLQ-C30 and LC13

The following scores will be derived as per the PRO SAP.

EORTC QLQ-C30

- Global Health Status
- Functional scales:
  - Physical functioning
  - Role functioning
  - Emotional functioning
  - Cognitive functioning
  - Social functioning
- Symptom scales/items:
  - Fatigue
  - Nausea and vomiting
  - Pain

CCI [redacted]

- Dyspnea
- Insomnia
- Appetite loss
- Constipation
- Diarrhea
- Financial difficulties

EORTC QLQ-LC13

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

19.3.1.4. Patient Global Impression of Change (PGI-C) and Severity (PGI-S)

The following responses will be analysed as per the PRO SAP.

PGI-S:

- None
- Mild
- Moderate
- Severe
- Very Severe

PGI-C:

- Very much improved
- Much improved
- Minimally improved

CCI [redacted]

- No change
- Minimally worse
- Much worse
- Very much worse

19.3.1.5. EQ-5D-5L

The following responses will be analysed per the PRO SAP.

- Mobility
  - No problems in walking
  - Slight problems in walking
  - Moderate problems in walking
  - Severe problems in walking
  - Unable to walk
- Self-Care
  - No problems washing or dressing
  - Slight problems washing or dressing
  - Moderate problems washing or dressing
  - Severe problems washing or dressing
  - Unable to wash or dress
- Usual activities
  - No problems doing usual activities
  - Slight problems doing usual activities
  - Moderate problems doing usual activities
  - Severe problems doing usual activities
  - Unable to do usual activities
- Pain/Discomfort
  - No pain or discomfort
  - Slight pain or discomfort
  - Moderate pain or discomfort
  - Severe pain or discomfort
  - Extreme pain or discomfort

CCI [redacted]

- Anxiety/Depression
    - Not anxious or depressed
    - Slightly anxious or depressed
    - Moderately anxious or depressed
    - Severely anxious or depressed
    - Extremely anxious or depressed

19.3.1.6. Time to CNS progression (Part 2)

The time to CNS progression is defined as the time from randomization to the development of CNS progression as defined by the RANO-BM working group response assessment criteria, by BICR and by investigator assessment. The overall assessment will be taken from the CRF page and will not be programmed.

[Progression date from the RANO-BM working group criteria – randomization date + 1]/30.4375

Patients who do not have progressive disease according to RANO-BM criteria, will be censored to the last valid RANO-BM evaluation.

19.3.2. Analysis of Exploratory Efficacy Variables

Exploratory analyses will be performed on the Part 1 Safety Population for Part 1 analyses, and Intent-to-Treat Population for Part 2 analyses.

19.3.2.1. QTcF analysis (Part 1 only)

Descriptive summaries of observed and change from baseline values will be presented for QTc interval using Fridericia's correction method.

19.3.2.2. Time to Treatment Failure (TTF)

For Part 2, time to treatment failure (TTF) will be compared between the 2 treatment arms using stratified logrank tests. Kaplan-Meier analyses will be performed on each treatment group to obtain the median TTF with the 95% CI. A stratified Cox proportional hazards regression will be used to estimate the hazard ratios and their corresponding 95% confidence intervals.

The summary for TTF will present the frequency distribution of TTF status (treatment failure, death or censored), frequency distribution of the reasons for censoring at the time of the data cut (alive and not failed treatment, withdrawn from study, lost to follow-up), Kaplan-Meier estimates of median TTF, and two-sided p-values from the logrank tests. Summaries and analyses will be presented for the Part 2 Intent-to-Treat population.

A figure displaying the Kaplan-Meier TTF curves for each treatment will be presented.

The analysis will be performed on Part 1 and Part 2. For Part 1, only descriptive summary will be presented, no comparisons between treatment groups as there is only one treatment group.

CCI [redacted]

19.3.2.3. EORTC-QLQ-C30 and LC13

Descriptive summaries of observed and change from baseline values will be presented each function.

19.3.2.4. Patient Global Impression of Change (PGI-C) and Severity (PGI-S)

PGI-S:

Frequencies and percentages will be presented for each of the responses.

PGI-C:

Frequencies and percentages will be presented for each of the responses.

19.3.2.5. EQ-5D-5L

Frequency and percentages will be presented for each of the responses for each item.

The VAS score on how good or bad health is today will be summarized descriptively.

19.3.2.6. Time to CNS progression

For Part 2, time to CNS progression will be compared between the 2 treatment arms using stratified logrank tests (stratified for region and platinum resistant). Kaplan-Meier analyses will be performed on each treatment group to obtain the median Time to CNS progression with the 95% CI. A stratified Cox proportional hazards regression will be used to estimate the hazard ratios and their corresponding 95% confidence intervals.

The summary for Time to CNS progression will present the frequency distribution of Time to CNS progression status (CNS progression, death or censored), frequency distribution of the reasons for censoring at the time of the data cut, Kaplan-Meier estimates of median time to CNS progression, and two-sided p-values from the logrank tests. Summaries and analyses will be presented for the Part 2 Intent-to-Treat population.

A figure displaying the Kaplan-Meier Time to CNS progression curves for each treatment will be presented.

20. SAFETY OUTCOMES

All outputs for safety outcomes will be based on the Safety Analysis Set.

There will be no statistical comparisons between the treatment groups for safety data, unless otherwise specified with the relevant section.

All safety outcomes will be presented separately for the Part 1 and Part 2 of the study and by UGT1A1*28 allele status (homozygous TA7/7) for relevant AE summaries.

CCI [redacted]

20.1. Adverse Events

Adverse Events (AEs) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) central coding dictionary, using current version.

Adverse events will be considered treatment-emergent if:

- The event occurs for the first time after the start of study treatment and on or before 30 days after final dose of study treatment and was not observed prior to start of study treatment
- The event was observed prior to the start of study treatment but increased in NCI CTCAE v5.0 grade during study treatment.
- AEs that start on the same day as the first dose of study drug will be considered treatment emergent.

See Appendix 2 for handling of partial dates for AEs. In the case where it is not possible to define an AE as treatment emergent or not, the AE will be classified by the worst case; i.e. treatment emergent.

An overall summary of number of patients within the categories described in the sub-sections below, will be provided as specified in the templates. Separate summaries will be presented for Part 1 and Part 2.

Listings will include TEAEs and Non-TEAEs.

20.1.1. All TEAEs

Incidence of TEAEs will be presented by System Organ Class (SOC) and Preferred Term (PT). Patients with multiple adverse events within a system organ class or preferred term will only be counted once under each category. A separate table for all events occurring in at least 10% of patients in the Irinotecan Liposome Injection treatment will also be presented.

A separate table by preferred term only will also be presented.

20.1.2. All TEAEs By Maximum CTCAE Grade

Tabular summaries will be presented by maximum CTCAE grade. Patients will be counted once at each AE level (SOC and PT), using the highest CTCAE grade event for classification.

20.1.3. All TEAEs of Grade 3 or more

A summary of all TEAEs of Grade 3 or higher by SOC and PT will be prepared.

A separate table will be produced for TEAEs with NCI-CTCAE Grade 3 or higher.

20.1.4. All Related TEAEs

Related TEAE are events recorded as such by the Investigator in the CRF. If a patient report's the same AE more than once within that SOC and PT, the AE with the worst case relationship to study medication will be used in the corresponding relationship summaries. A separate table for all events occurring in at least 10% of patients in the Irinotecan Liposome Injection treatment will also be presented.

A separate table will be produced for Related TEAEs with NCI-CTCAE Grade 3 or higher.

A separate table will be produced for the Irinotecan liposome injection arm in Part 2 for Related TEAEs with terms including combined specified preferred terms.

20.1.5. All Serious TEAEs

Serious adverse events (SAEs) are those events recorded as "Serious" on the *Adverse Events* page of the CRF. A summary of serious TEAEs by SOC and PT will be prepared. A separate table for all events occurring in at least 5% of patients in the Irinotecan Liposome Injection treatment will also be presented.

20.1.6. All Serious Related TEAEs

Serious adverse events (SAEs) are those events recorded as "Serious" on the *Adverse Events* page of the CRF and who recorded "Related" on the *Adverse Events* page. A summary of serious Related TEAEs by SOC and PT will be prepared.

20.1.7. TEAEs Leading to Dose Adjustment

TEAEs leading to dose delayed/interrupted/reduced will be identified in the action taken on the CRF. A separate table for all events occurring in at least 5% of patients in the Irinotecan Liposome Injection treatment will also be presented.

Summaries by SOC and PT will be prepared.

20.1.8. TEAEs Leading to Treatment Discontinuation

TEAEs leading to permanent discontinuation of study medication will be identified by using the "discontinued from study treatment" in action taken on the CRF. A separate table for all events occurring in at least 2 patients in the Irinotecan Liposome Injection treatment will also be presented.

Summaries by SOC and PT will be prepared.

CCI [redacted]

20.1.9. TEAEs Leading to Death

TEAEs leading to Death are those events which are recorded as "Patient Died" on the *Adverse Events Outcome* page of the CRF. A summary of TEAEs leading to death by SOC and PT will be prepared.

20.1.10. Infusion Related TEAEs

Infusion related TEAEs are those events which answered "yes" to the infusion related check box on the CRF.

Summaries by SOC and PT will be prepared.

20.1.11. Dose Limiting Toxicities (DLTs) AEs

For Part 1 only, DLT AEs which are indicated by the Investigator on the CRF will be summarized and listed.

20.1.12. Related TEAEs Grade 3 or Higher by Platinum Sensitivity

Related TEAEs of Grade 3 or higher will be summarized by platinum sensitivity (platinum sensitive and platinum resistant) in all part 2 patients split by treatment group. A separate table for all events occurring in at least 10% of patients in the Irinotecan Liposome Injection treatment will also be presented.

20.1.13. Subgroups of Adverse Events

The following summaries will also be created for patients that are UGT1A1 homozygous TA7/7:

- Overall summary of adverse events
- All treatment emergent adverse events
- All treatment emergent adverse events by maximum NCI-CTCAE grade (all grades, grade 1, grade 2, grade 3, grade 4 and grade 5)
- All related treatment emergent adverse events
- All treatment emergent adverse events of NCI-CTCAE grade 3 or higher
- All related treatment emergent adverse events of NCI-CTCAE grade 3 or higher
- All treatment emergent adverse events leading to treatment discontinuation
- All serious treatment emergent adverse events
- All serious related treatment emergent adverse events
- All treatment emergent adverse events leading to death

20.2. DEATHS

If any patients die during the study as recorded on the *Death Details* page of the CRF, the information will be presented in a summary table and a data listing.

20.3. LABORATORY EVALUATIONS

Laboratory evaluations are collected by Q2 central laboratory.

Scheduled clinical safety laboratory parameters will be summarized. These include:

- **Hematology**: hemoglobin, hematocrit, platelet count, Red Blood Cells (RBC), White Blood Cells (WBC) with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils and other cells)
- **Chemistry**: electrolyte (sodium, potassium, chloride and bicarbonate), Blood Urea Nitrogen (BUN), serum creatinine, glucose, bilirubin, AST, ALT, alkaline phosphatase, lactate dehydrogenase, uric acid, total protein, albumin, calcium, magnesium and phosphate

Presentations will use SI Units.

Quantitative laboratory measurements reported as "< X", i.e. below the lower limit of quantification (BLQ), or "> X", i.e. above the upper limit of quantification (ULQ), will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "< X" or "> X" in the listings.

The following summaries will be provided for laboratory data:

- Actual and change from baseline by visit (for quantitative measurements)
- Shift from baseline to end of study according to Common Toxicity (CTC) grading system (see Appendix 3). Shifts from baseline to worst post-dose (i.e. highest grade) in CTCAE Grade of laboratory results will also be summarized.
- Listing of CTCAE grade 3 and higher values. In this listing, all observations for a lab parameter will be displayed for a patient who has any post-baseline value with CTCAE grade greater than or equal to 3 for the parameter.
- For specified laboratory results, the percentage of patients in the irinotecan liposome injection arm in Part 2 with worsening laboratory test results from baseline to >=20% post baseline result (for all patients with baseline and post baseline value) for a) all post baseline grades and b) for all post baseline grades where the worst grade is Grade 3 or 4 according to Common Toxicity (CTC) grading system (see Appendix 3).

20.4. ECG EVALUATIONS

Results from the central ECG (Electrocardiogram) Reading Centre will be included in the reporting of this study for Part 1 only.

The following ECG parameters will be reported for the Part 1 only and be collected at Cycle 1 day 1(pre-dose, end of infusion, 2 hours after infusion), day 2 (24 hours after end of infusion), day 8 and day 15 (pre-dose).

- Heart Rate (bpm)
- PR Interval (msec)
- RR Interval (msec)
- QRS Interval (msec)
- QT Interval (msec)
- Overall assessment of ECG (Investigator's judgment):
  - Normal
  - Abnormal, Not Clinically Significant (ANCS)
  - Abnormal, Clinically Significant (ACS)

The following summaries will be provided for ECG data:

- Actual and change from baseline by visit (for quantitative measurements)
- Listing of patients meeting markedly abnormal criteria
- The number and percentage of patients with "Normal", "Abnormal, not Clinically Significant" and "Abnormal, Clinically Significant" ECG results will be summarized by treatment at each visit.

20.4.1. ECG Markedly Abnormal Criteria

Markedly abnormal quantitative ECG measurements will be identified in accordance with the following predefined markedly abnormal criteria:

- Absolute values for QTcF Interval classified as:
  - > 500 msec
  - > 480 msec and ≤ 500 msec
  - > 450 msec and ≤ 480 msec
  - Change from Baseline > 30 and ≤ 60 msec
  - Change from Baseline > 60 msec

20.5. Vital Signs

The following Vital Signs measurements will be reported for this study:

- Resting Systolic Blood Pressure (mmHg)
- Resting Diastolic Blood Pressure (mmHg)
- Respiratory Rate (resp/min)

CCI

[redacted]

- Pulse (bpm)
- Temperature (^{0}C)
- Height (cm)
- Weight (kg)
- BMI (kg/m^{2})

The following summaries will be provided for vital signs data:

- Actual and change from baseline by visit
- Incidence of out of normal range during treatment (up to 30 days post last dose date) as presented in Table 5.
- Listing of patients with markedly abnormal values

Table 5: Vital Signs Normal Range

Parameter	Units	Normal Range
Systolic Blood Pressure	mmHg	>90 - <180
Diastolic Blood Pressure	mmHg	>50 - <105
Pulse Rate	Beats/minute	>50 - <120
Respiratory Rate	Resp/minute	>12 – <25

20.5.1. VITAL SIGNS MARKEDLY ABNORMAL CRITERIA

Markedly abnormal quantitative Vital Signs measurements will be identified in accordance with the following predefined markedly abnormal criteria as presented in the table below:

Table 6: Markedly Abnormal Criteria in Vital Signs

Variable	Unit	Low	High
Systolic Blood Pressure (SBP)	mmHg	≤ 90 mmHg AND change from baseline ≤ -20 mmHg	≥ 180 mmHg AND change from baseline ≥ 20 mmHg
Diastolic Blood Pressure (DBP)	mmHg	≤ 50 mmHg AND change from ≤ -15 mmHg	≥ 105 mmHg AND change from baseline ≥ 15 mmHg
Heart rate	Bpm	≤ 50 bpm AND change from baseline ≤ -15 bpm	≥ 120 bpm AND change from baseline ≥ 15 bpm

CCI

[redacted]

Variable	Unit	Low	High
Body temperature	˚C	NA	≥ 38.3 ˚C AND change from baseline ≥ 1.1 ˚C
Weight	Kg	percentage change from baseline ≤ -7.0 %	percentage change from baseline ≥ 7.0 %

20.6. PHYSICAL EXAMINATION

Results of physical examinations for each body system will be listed by treatment group.

20.7. UGT1A1*28 AND OTHER GENOTYPES

Individual listings of UGT1A1*28 allele status (homozygous TA7/7) as well as summary statistics split by treatment arm will be presented for each Part of the study.

Safety summaries will also be presented by UGT1A1*28 allele status (homozygous TA7/7). The impact of UGT1A1*28 status will also be explored as part of the pharmacokinetic covariate assessment (see section 20.8). Association will also be described between UGT1A1*28 and other UGT1A1 genotypes, SN-38 concentration (Irinotecan liposome injection treated patients only) and safety.

20.8. PHARMACOKINETICS

The PK and PK/PD relationship will be performed for the Onivyde treatment arm only. All PK summaries will be performed on the PK populations (Part 1 and part 2).

20.8.1. LISTINGS AND SUMMARY STATISTICS OF CONCENTRATIONS

Descriptive statistics (N, number with BLQ, arithmetic mean, SD, median, minimum, maximum, arithmetic coefficient of variation (CV) (%), geometric mean and geometric CV(%) will be used to summarize Irinotecan and SN-38 plasma concentration data, by dose level, at each planned time point for part 1 and part 2. To compute descriptive statistics, all BLQ values must be replaced by missing and excluded from the calculation of descriptive statistics. Concentration values reported as below the limit of quantificiation (BLQ) will be listed with the lower limit of quantification in parentheses. The descriptive statistics should be displayed by visit/time point, only if at least 2/3 (2 out of every 3 values) of the data are available and above the limit of quantification. Otherwise, only minimum and maximum are reported.Values excluded from the summary statistics of concentrations will be flagged with an asterisk. Spaghetti plots of individual observed concentrations will be included in a separate

Population PK report.

20.8.2. Pharmacokinetic Data Analysis

Pharmacokinetics of total irinotecan and SN-38 will be quantified from the concentrations from plasma samples using nonlinear mixed effect modeling (NONMEM). A Population PK model previously developed on historical data will be applied to this study in order to characterize the PK in 2L SCLC patient population, to assess interindividual variability, and to attempt to explain part of it by testing the impact of some covariates. Covariates evaluated would include, but not be limited to, age, gender, race, and UGT1A1*28 genotype. Primary and derived PK parameters will be computed from the population PK model (e.g. maximum observed concentration (C_{max}), average observed concentration ($C_{average}$), area under the curve (AUC), minimum concentration (C_{min})). Full description of this Population PK analysis will be captured separately in a data analysis plan, and results will be reported in a standalone report.

20.8.3. Pharmacokinetics / Pharmacodynamics Relationship

Graphical exploration will be performed to investigate any relationship between PK and pharmacodynamic endpoints (safety e.g. neutropenia and diarrhea; and efficacy, e.g. ORR, OS, PFS). If a trend is shown, PK/PD modelling will be performed and this will be described in a separate Data Analysis Plan and reported in a standalone report.

21. REFERENCES

Hwang, I. K., Shih, W. J., and DeCani, J. S. (1990), "Group Sequential Designs Using a Family of Type I Error Probability Spending Functions," *Statistics in Medicine*, 9, 1439–1445.

O'Kelly, M., & Ratitch, B. (2017). Clinical Trial with Missing Data: A Guide for Practitioners. John Wiley and Sons, Ltd.

E.A. Eisenhauera,* , P. Therasseb , J. Bogaertsc , L.H. Schwartzd , D. Sargente , R. Fordf , J. Danceyg , S. Arbuckh , S. Gwytheri , M. Mooneyg , L. Rubinsteing , L. Shankarg , L. Doddg , R. Kaplanj , D. Lacombec , J. Verweijk "New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)" *EUROPEAN JOURNAL OF CANCER* 45 (2009) 228 – 247

APPENDIX 1. REPORTING CONVENTIONS

Page Layout:

The Section 14 tables and Appendix 16 listings should be in landscape orientation by default.

The output in Section 14 and Appendix 16 will be in RTF file format using Courier New font with 8 point size.

Statistical conventions:

The patient percentages (%) should be rounded to one decimal place.

Percentages for values in the tables that are less than < 0.1 should be presented as "<0.1".

If "%" is part of the column heading, do not repeat the "%" sign in the body of the table.

If a value is zero (0), then do not use 0% and leave the corresponding percentage blank.

The format for range should always be "Min, Max".

If there are missing data, then a missing row will be added to keep track of all patients. If there are no missing data, then delete the missing row. Percentages will not be presented on the missing category row.

Standard Deviation should be abbreviated as "SD", and Standard Error should be abbreviated as "SE"; it is presented within parenthesis next to the mean value, without any +/- sign. The Standard Deviation or Standard Error should have one additional decimal point beyond that of the mean (for example, if the mean has one decimal point, SD/SE should have two decimal points). Mean and median should have one additional decimal point beyond that of the data being summarized.

"N" will represent the entire treatment group, while "n" will represent a subset of the treatment group. For tables with population designated as a row heading, "N" should be used (i.e. tables where all the participant data is not available for every variable within a treatment group). As a guideline, if the number is used in denominator that it should be presented as "N". If the number is used in numerator it should be presented as "n".

P-values will be presented with 4 decimals.

APPENDIX 2. PARTIAL DATE CONVENTIONS

Imputed dates will NOT be presented in the listings.

Algorithm for Treatment Emergence of Adverse Events:

START DATE	STOP DATE	ACTION
Known	Known	If start date < study med start date, then not TEAE If start date >= study med start date, then TEAE
	Partial	If start date < study med start date, then not TEAE If start date >= study med start date, then TEAE
	Missing	If start date < study med start date, then not TEAE If start date >= study med start date, then TEAE
Partial, but known components show that it cannot be on or after study med start date	Known	Not TEAE
	Partial	Not TEAE
	Missing	Not TEAE
Partial, could be on or after study med start date	Known	If stop date < study med start date, then not TEAE If stop date >= study med start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If stop date >= study med start date, then TEAE
	Missing	Assumed TEAE

CCI [redacted]

START DATE	STOP DATE	ACTION
Missing	Known	If stop date < study med start date, then not TEAE If stop date >= study med start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, then not TEAE If stop date >= study med start date, then TEAE
	Missing	Assumed TEAE

Algorithm for Prior / Concomitant Medications:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study med start date, assign as prior If stop date >= study med start date and start date <= end of treatment, assign as concomitant If stop date >= study med start date and start date > end of treatment, assign as post study
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date >= study med start date and start date <= end of treatment, assign as concomitant If stop date >= study med start date and start date > end of treatment, assign as post treatment
	Missing	If stop date is missing could never be assumed a prior medication If start date <= end of treatment, assign as concomitant If start date > end of treatment, assign as post treatment

CCI

START DATE	STOP DATE	ACTION
Partial	Known	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date >= study med start date and start date <= end of treatment, assign as concomitant If stop date >= study med start date and start date > end of treatment, assign as post treatment
	Partial	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date >= study med start date and start date <= end of treatment, assign as concomitant If stop date >= study med start date and start date > end of treatment, assign as post treatment
	Missing	Impute start date as earliest possible date (i.e. first day of month if day unknown or 1st January if day and month are unknown), then: If stop date is missing could never be assumed a prior medication If start date <= end of treatment, assign as concomitant If start date > end of treatment, assign as post treatment
Missing	Known	If stop date < study med start date, assign as prior If stop date >= study med start date, assign as concomitant Cannot be assigned as 'post treatment'
	Partial	Impute stop date as latest possible date (i.e. last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study med start date, assign as prior If stop date >= study med start date, assign as concomitant Cannot be assigned as 'post treatment'
	Missing	Assign as concomitant

CCI [redacted]

Algorithm for Other Partial Dates:

START DATE/STOP DATE	ACTION
Missing day	Replace by 1
Missing month	Replace by January
Missing year	Leave as missing

CCI

[redacted]

APPENDIX 3. CTCAE GRADING FOR NUMERIC LAB VALUES

From

https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf

Parameter	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin decreased	6.2 mmol/L - < LLN	4.9 mmol/L - <6.2 mmol/L	<4.9 mmol/L	
Hemoglobin increase	↑ in >0 – 2 g/dL above ULN or > BL if BL>ULN	↑ in >2 – 4 g/dL above ULN or > BL if BL>(ULN+2)	↑ in >4 g/dL above ULN or > BL if BL>(ULN+4)	
Lymphocyte decrease	$0.8x10^9$/L – <LLN	$0.5x10^9$/L – <$0.8x10^9$/L	$0.2x10^9$/L – <$0.5x10^9$/L	<$0.2x10^9$/L
Neutrophils decrease	$1.5x10^9$/L – <LLN	$1.0x10^9$/L – <$1.5x10^9$/L	$0.5x10^9$/L – <$1.0x10^9$/L	<$0.5x10^9$/L
Platelets decrease	$75.0x10^9$/L – <LLN	$50.0x10^9$/L – <$75.0x10^9$/L	$25.0x10^9$/L – <$50.0x10^9$/L	<$25.0x10^9$/L
WBC decrease	$3.0x10^9$/L – <LLN	$2.0x10^9$/L – <$3.0x10^9$/L	$1.0x10^9$/L – <$2.0x10^9$/L	<$1.0x10^9$/L
ALT	>1 ULN – 3xULN if baseline was normal; 1.5 – 3.0xbaseline if baseline was abnormal	>3.0 ULN – 5.0xULN if baseline was normal; 3.0 – 5.0xbaseline if baseline was abnormal	>5.0 ULN – 20.0xULN if baseline was normal; 5.0 – 20.0xbaseline if baseline was abnormal	>20xULN if baseline was normal; >20.0xbaseline if baseline was abnormal
Alkaline Phosphatase	>1 ULN – 2.5xULN if baseline was normal; 2.0 – 2.5xbaseline if baseline was abnormal	>2.5 ULN – 5.0xULN if baseline was normal; 2.5 – 5.0xbaseline if baseline was abnormal	>5.0 ULN – 20.0xULN if baseline was normal; 5.0 – 20.0xbaseline if baseline was abnormal	>20xULN if baseline was normal; >20.0xbaseline if baseline was abnormal

CCI

[redacted]

Parameter	Grade 1	Grade 2	Grade 3	Grade 4
AST	>1 ULN – 3xULN if baseline was normal; 1.5 – 3.0xbaseline if baseline was abnormal	>3.0 ULN – 5.0xULN if baseline was normal; 3.0 – 5.0xbaseline if baseline was abnormal	>5.0 ULN – 20.0xULN if baseline was normal; 5.0 – 20.0xbaseline if baseline was abnormal	>20xULN if baseline was normal; >20.0xbaseline if baseline was abnormal
Total Bilirubin	>1 ULN – 1.5xULN if baseline was normal; >1.0 – 1.5xbaseline if baseline was abnormal	>1.5 ULN – 3.0xULN if baseline was normal; 1.5 – 3.0xbaseline if baseline was abnormal	>3.0 ULN – 10.0xULN if baseline was normal; 3.0 – 10.0xbaseline if baseline was abnormal	>10xULN if baseline was normal; >10.0xbaseline if baseline was abnormal
Creatinine	>1xBL – 1.5xBL or >1xULN – 1.5xULN	>1.5xBL – 3xBL or >1.5xULN – 3xULN	>3xBL – 6xBL or >3xULN – 6xULN	>6xBL or >6xULN
Calcium increase	>ULN – <2.9 mmol/L	> 2.9 mmol/L– 3.1 mmol/L	> 3.1 mmol/L– 3.4 mmol/L	>3.4 mmol/L
Calcium decrease	2.0 mmol/L – <LLN	1.75 mmol/L – <2.0 mmol/L	1.5 mmol/L – <1.75 mmol/L	<1.5 mmol/L
Albumin	3 g/dL - <LLN	2 g/dL - < 3 g/dL	< 2 g/dL	
Magnesium decrease	0.5 mmol/L - <LLN	0.4 mmol/L - <0.5 mmol/L	0.3 mmol/L - <0.4 mmol/L	<0.3 mmol/L
Magnesium increase	>ULN – 1.23 mmol/L		>1.23 mmol/L – 3.30 mmol/L	>3.30 mmol/L
Potassium decrease	3.0 mmol/L - <LLN		2.5 mmol/L <3.0 mmol/L	<2.5 mmol/L

CCI

[redacted]

Parameter	Grade 1	Grade 2	Grade 3	Grade 4
Potassium increase	>ULN – 5.5 mmol/L	>5.5 mmol/L – 6.0 mmol/L	>6.0 mmol/L – 7.0 mmol/L	>7.0 mmol/L
Glucose	3.0 mmol/L - <LLN	2.2 mmol/L - <LLN	1.7 mmol/L - <2.2 mmol/L	<1.7 mmol/L
Sodium decrease	130 mmol/L - <LLN		120 mmol/L - <130 mmol/L	<120 mmol/L
Sodium increase	>ULN – 150 mmol/L	>150 mmol/L – 155 mmol/L	>155 mmol/L – 160 mmol/L	>160 mmol/L
Phosphate	0.8 mmol/L - <LLUN	0.6 mmol/L - <0.8 mmol/L	0.3 mmol/L - <0.6 mmol/L	<0.3 mmol/L
Uric Acid	>ULN – 0.59 mmol/L		>0.59 mmol/L	

CCI [redacted]

APPENDIX 4: EORTC QLQ-C30

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in your initials: ____
Your birthdate (Day, Month, Year): ____
Today's date (Day, Month, Year): 31 ____

	Not at All	A Little	Quite a Bit	Very Much
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2. Do you have any trouble taking a long walk?	1	2	3	4
3. Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4. Do you need to stay in bed or a chair during the day?	1	2	3	4
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
6. Were you limited in doing either your work or other daily activities?	1	2	3	4
7. Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8. Were you short of breath?	1	2	3	4
9. Have you had pain?	1	2	3	4
10. Did you need to rest?	1	2	3	4
11. Have you had trouble sleeping?	1	2	3	4
12. Have you felt weak?	1	2	3	4
13. Have you lacked appetite?	1	2	3	4
14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4
16. Have you been constipated?	1	2	3	4

Please go on to the next page

CCI [redacted]

ENGLISH

During the past week:	Not at All	A Little	Quite a Bit	Very Much
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

CCI [redacted]

Table 7. EORTC QLQ-C30 Scoring Guide (Fayers et al 2001)

	Abbreviation	Number of items	Item range*	Version 3.0 Item numbers
Global health status / QoL				
Global health status/QoL (revised)†	QL2	2	6	29, 30
Functional domains				
Physical functioning (revised)†	PF2	5	3	1 to 5
Role functioning (revised)†	RF2	2	3	6, 7
Emotional functioning	EF	4	3	21 to 24
Cognitive functioning	CF	2	3	20, 25
Social functioning	SF	2	3	26, 27
Symptom domains / items				
Fatigue	FA	3	3	10, 12, 18
Nausea and vomiting	NV	2	3	14, 15
Pain	PA	2	3	9, 19
Dyspnea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite loss	AP	1	3	13
Constipation	CO	1	3	16
Diarrhea	DI	1	3	17
Financial difficulties	FI	1	3	28

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

Scoring algorithm

The EORTC-QLQ-C30 is composed of both multi-item domains and single-item measures. These include five functional domains, three symptom domains, a global health status / QoL domain, and six single items. Each of the multi-item domains includes a different set of items - no item occurs in more than one domain.

The principle for scoring these domains is the same in all cases:

1. Estimate the average of the items that contribute to the domain; this is the raw score.

2. Use a linear transformation to standardize the raw score, so that scores range from 0 to 100; a higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms.

The technical details are provided below.

If items I_1, I_2, ... I_n are included in a domain, the procedure is as follows:

Calculate the raw score (RS) as follows: $RS=(I_1+I_2+ \ldots +I_n)/n$

Apply the linear transformation as follows:

Functional domains: Domain score = $1 - ((RS - 1)/range) \times 100$

Symptom domains / items: Domain score = $(RS - 1)/range \times 100$

Global health status/QoL: Domain score = $(RS - 1)/range \times 100$

where Range is the difference between the maximum possible value of RS and the minimum possible value. The EORTC-QLQ-C30 has been designed so that all items in any domain take the same range of values. Therefore, the range of RS equals the range of the item values (see **Error! Reference source not found.** PRO Variables).

If at least half the components of a domain are present, then the domain score will be calculated using the average of all items answered as the raw score; otherwise the score will be set to missing. For single measures, it the item is missing the domain score is set to missing.

CCI

[redacted]

APPENDIX 5: EORTC QLQ-LC13

EORTC QLQ - LC13

Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week :	Not at All	A Little	Quite a Bit	Very Much
31. How much did you cough?	1	2	3	4
32. Did you cough up blood?	1	2	3	4
33. Were you short of breath when you rested?	1	2	3	4
34. Were you short of breath when you walked?	1	2	3	4
35. Were you short of breath when you climbed stairs?	1	2	3	4
36. Have you had a sore mouth or tongue?	1	2	3	4
37. Have you had trouble swallowing?	1	2	3	4
38. Have you had tingling hands or feet?	1	2	3	4
39. Have you had hair loss?	1	2	3	4
40. Have you had pain in your chest?	1	2	3	4
41. Have you had pain in your arm or shoulder?	1	2	3	4
42. Have you had pain in other parts of your body?	1	2	3	4
If yes, where _____				
43. Did you take any medicine for pain?				
1 No 2 Yes				
If yes, how much did it help?	1	2	3	4

CCI [redacted]

Table 8. EORTC-QLQ-LC13 Lung Cancer Module Scoring Guide

	Domain	Number of items	Item range*	Item numbers
Symptom domains / items				
Dyspnea†	LCDY	3	3	33,34,35
Coughing	LCCO	1	3	31
Hemoptysis	LCHA	1	3	32
Sore mouth	LCSM	1	3	36
Dysphagia	LCDS	1	3	37
Peripheral neuropathy	LCPN	1	3	38
Alopecia	LCHR	1	3	39
Pain in chest	LCPC	1	3	40
Pain in arm or shoulder	LCPA	1	3	41
Pain in other parts	LCPO	1	3	42

* "Item range" is the difference between the possible maximum and the minimum response to individual items.

† The dyspnea domain should only be used if all three items have been answered. Some respondents ignore question 35 because they never climb stairs; in this case, the score for the dyspnea domain would be biased if it were based upon the other two items. Hence if item 35 is missing then items 33 and 34 should be used as single-item measures.

In addition, a new domain will be calculated, Dyspnea Alternate (LCDYALT) which combines the three EORTC-QLQ-LC13 items and the Item 8 from EORTC-QLQ-C30:

LCDYALT= {[(Q8+Q33+Q34+Q35)/4]-1}/3 * 100

The scoring approach for the EORTC-QLQ-LC13 is identical to that for the symptom domains / single items of the EORTC-QLQ-C30. A raw score is calculated as the average of the items that contribute to the domain and next a linear transformation is used to standardize the raw score (RS), so that scores range from 0 to 100. The linear transformation is as follows:

Domain score = (RS –1)/range×100

where RS=raw score and range=3 (see Table 8).

If at least half the components of a domain are present, then the domain score will be calculated using the average of all items answered as the raw score; otherwise the score will be set to missing. For single measures, it the item is missing the domain score is set to missing.

CCI [redacted]

APPENDIX 6: CONFIRMED BEST OVERALL RESPONSE

Table 3 – Best overall response when confirmation of CR and PR required.

Overall response First time point	Overall response Subsequent time point	BEST overall response
CR	CR	CR
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.
a If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.