Integrated Analysis Plan

Study Number:	Debio 1143-SCCHN-301/MS202359_0006
Clinical Study Protocol Title:	A randomized, double-blind placebo-controlled, Phase 3 study of Debio 1143 in combination with platinum-based chemotherapy and standard fractionation intensity-modulated radiotherapy in patients with locally advanced squamous cell carcinoma of the head and neck, suitable for definitive chemoradiotherapy (TrilynX)
Study Phase:	Phase III
Merck Compound:	Xevinapant (also known as Debio 1143)
Protocol Version:	29 February 2024 / Version 12.0
Integrated Analysis Plan Author:	**Function** Global Biostatistics — **Coordinating Author** PPD **Function** Biostatistics, PPD — **Author(s) / Data Analyst(s)** PPD
Integrated Analysis Plan Date and Version:	8 March 2024 / Version 5.0
Integrated Analysis Plan Reviewers:	**Function** — **Name** CMS Biostatistics — PPD Biostatistics, PPD — CCI GCD GPS GRA PPD GEVD DMMI GDO/COL

CCI

Approval Page

Integrated Analysis Plan: Debio 1143-SCCHN-301/MS202359_0006

A randomized, double-blind placebo-controlled, Phase 3 study of Debio 1143 in combination with platinum-based chemotherapy and standard fractionation intensity-modulated radiotherapy in patients with locally advanced squamous cell carcinoma of the head and neck, suitable for definitive chemoradiotherapy (TrilynX)

Approval of the IAP by all Merck Data Analysis Responsible has to be documented within EDMS via eSignature. With the approval, the Merck responsible for each of the analysis also takes responsibility that all reviewers' comments are addressed adequately.

By using eSignature, the signature will appear at the end of the document.

1 Table of Contents

Approval Page 2
1 Table of Contents....3
1.1 List of Tables 7
2 List of Abbreviations and Definition of Terms 8
3 Modification History 11
4 Purpose of the Integrated Analysis Plan....11
5 Objectives and Endpoints 12
6 Overview of Planned Analyses....20
6.1 Safety Monitoring Analysis....21
6.2 Interim Analyses....22
6.3 Primary Analysis 23
6.4 Further OS Analyses....23
7 Changes to the Planned Analyses in the Clinical Study Protocol 24
8 Analysis Sets and Subgroups....25
8.1 Definition of Analysis Sets....25
8.2 Subgroup Definition and Parameterization 27
8.2.1 Randomization stratification factors (from IWRS) 27
8.2.2 Other subgroups....28
8.2.3 Analyses in Asian Participants 29
9 General Specifications for Data Analyses 30
9.1 Data handling after Cut-off Date 30
9.2 Definition of Baseline and Change from Baseline 30
9.3 First Dose Date 31
9.4 Last Dose Date....31
9.5 Definition of On-treatment Period....31
9.6 Study Day / Study Treatment Day....31
9.7 Definition of Duration and 'Time Since' Variables 31
9.8 Conversion Factors 32
9.9 Date of Last Contact 32
9.10 Time Window 32
9.11 Time on Study....32

9.12 Imputation of Missing Data 33
9.12.1 Partial and Missing Dates 33
9.12.2 Tumor assessments 34
9.12.3 Tumor Evaluation and Adjudication 34
9.13 Unscheduled Assessments or Retests 35
9.14 Scoring of Health-Related Quality of Life (HRQL) Data 35
9.15 Derivation of Best Overall Response 35
10 Study Participants 36
10.1 Disposition of Participants and Discontinuations 36
10.1.1 Screen Failures and Eligibility Criteria 36
10.1.2 Participant Disposition 36
10.1.3 Time on Study 37
10.2 Protocol Deviations 37
11 Demographics and Other Baseline Characteristics 39
11.1 Demographics 39
11.2 Medical and Surgical History 39
11.3 Other Baseline Characteristics 40
12 Prior, Concomitant and Subsequent Medications/Procedures 41
12.1 Concomitant 5-HT3 Receptor Antagonist 41
CCI [redacted]
13 Study Intervention: Compliance and Exposure 42
14 Efficacy Analyses 44
14.1 Primary Efficacy Endpoint: Event-free Survival (EFS) 44
14.1.1 Primary Analysis of the Primary Efficacy Endpoint EFS 48
14.1.2 Sensitivity Analyses of the Primary Efficacy Endpoint EFS 50
14.1.3 Supplementary Analyses of the Primary Efficacy Endpoint EFS 50
14.1.4 Subgroup Analyses of the Primary Efficacy Endpoint EFS 51
14.2 Secondary Efficacy Endpoints and Analyses 52
14.2.1 Overall Survival 52
14.2.2 Progression-Free Survival 53
14.2.3 Locoregional Control 55

14.2.4 Objective Response 57
14.2.4.1 Complete Response 58
14.2.5 Duration of Response 59
14.2.6 Proportion of Participants with Radical Salvage Surgery 59
14.2.7 Time to New Subsequent Systemic Cancer Treatment 60
CCI [redacted]
14.3.1 Proportion of Definitive Tracheostomy 60
14.3.2 Proportion of Enteral and Parenteral Feeding 60
14.3.3 Time to Distant Metastasis 61
CCI [redacted]
15 Safety Analyses 63
15.1 Adverse Events 63
15.1.1 All Adverse Events 64
15.1.2 Adverse Events Leading to Discontinuation of Study Treatment 67
15.2 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events 67
15.2.1 Deaths 67
15.2.2 Serious Adverse Events 67
15.2.3 Other Significant Adverse Events 68
15.2.3.1 Treatment-emergent AESIs 68
CCI [redacted]
[redacted]
15.2.3.4 Legacy Definition of AESIs 68
15.3 Clinical Laboratory Evaluation 69
15.4 Vital Signs 70
15.5 Other Safety or Tolerability Evaluations 70
15.5.1 ECG 70
15.5.2 ECOG Performance Status 71
15.5.3 Transfusions 71
15.5.4 Ear Examination and Audiogram 71
15.5.5 Dental Examination 71
15.5.6 Nutritional Status 72

C
CI
CCI
CCI
16.2.2 Instruments Scoring 75
16.2.2.1 EORTC Questionnaires 75
CCI
16.2.3 Key Baseline Covariates for HRQL Analysis 76
16.2.4 Analysis of HRQL Patterns 76
16.2.4.1 Descriptive Analysis of Health-Related Quality of Life Missing Data 76
16.2.5 Analysis of Health-Related Quality of Life Data 77
16.2.5.1 Descriptive Health-Related Quality of Life Analysis 77
16.2.5.2 Primary Health-Related Quality of Life Analysis 77
16.2.5.3 Exploratory Health-Related Quality of Life Analysis 77
16.2.6 Health State Utilities Assessments Reporting Guidelines 78
17 References 79
18 Appendices 80
Appendix 1 – Health Related Quality of Life Questionnaires Scoring Manuals 80
Appendix 2 – Legacy Definition of Adverse Events of Special Interest 84
Appendix 3 - NCI-CTCAE grading for Laboratory Parameters 86
Appendix 4 – Definition of Western Europe 88
CCI
Appendix 6 – Programming Specifications 95
CCI
Appendix 8 - Country-specific Amendment for China 102

1.1 List of Tables

Table 1: Objectives, endpoints and statistical considerations....13
Table 2: Planned Efficacy Analyses and Expected Timing....20
Table 3: Operating Characteristics and Power Estimations....21
Table 4: Censoring rules for EFS....45
Table 5: Estimand Attributes EFS....46
Table 6: Sensitivity analyses for EFS....50
Table 7: Supplementary analysis for EFS....51
Table 8: Estimand Attributes OS....53
Table 9: Censoring rules for PFS....54
Table 10: Analyses for PFS....54
Table 11: Censoring rules for LRC....56
Table 12: Analyses for LRC....56
Table 13: Analyses for DoR....59
Table 14: Analyses for TDM....62
Table 15: Scoring method for EORTC questionnaires....75
Table 16: EORTC QLQ-C30 definition of subscores....80
Table 17: EORTC QLQ-H&N35 definition of subscores....82
Table 18: NCI-CTC gradable parameters....86
Table 19: NCI-CTC non-gradable parameters....87
Table 20: Treatment-Emergent AESIs (Potential Tier 1 Safety Topics)....89
CCI

2 List of Abbreviations and Definition of Terms

A-P	Anatomopathological
ADaM	Analysis Data Model
AE	Adverse Event
AECI	Adverse Events of Clinical Importance
AESI	Adverse Event of Special Interest
AJCC	American Joint Committee on Cancer
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
ALP	Alkaline Phosphatase
ATC	Anatomical Therapeutic Chemical Classification
BIRC	Blinded Independent Review Committee
BMI	Body Mass Index
BSA	Body Surface Area
CI	Confidence Interval
CCI [redacted]	[redacted]
[redacted]	[redacted]
cIPD	Clinically Important Protocol Deviations
CR	Complete Response
CRR	Complete Response Rate
CRO	Contract Research Organization
CRT	ChemoRadioTherapy
CSR	Clinical Study Report
CT	ChemoTherapy
CxDy	Cycle x Day y
DFT	On-Study Duration of Feeding Tube Procedures
DoR	Duration of Response
ECG	Electrocardiogram
ECOG-PS	Eastern Cooperative Oncology Group Performance Status
eCRF	Electronic Case Report Form
eDISH	evaluation of Drug-Induced Serious Hepatotoxicity

EFS	Event-Free Survival
eGFR	estimated Glomerular Filtration Rate
END	Elective Neck Dissection
EORTC	European Organization for Research and Treatment of Cancer
EOT	End Of Treatment
EOS	End Of Study
GHS	Global Health Scale
HR	Hazard Ratio
HRQL	Health-Related Quality of Life
HRU	Health Resource Utilization
HSU	Health State Utilities
IAP	Integrated Analysis Plan
ICF	Informed Consent Form
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IDMC	Independent Data Monitoring Committee
IPD	Important Protocol Deviation
IMRT	Intensity Modulated Radio Therapy
IQWIG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Integrated Randomization Technology
ITT	Intent-to-treat
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LA-SCCHN	Locally Advanced Squamous Cell Carcinoma of the Head and Neck
LRC	Locoregional Control
MAR	Missing At Random
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Differences
NAT	New Anticancer Treatment
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OIS	Overall Imbalance Score
OR	Objective Response

ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PFS	Progression-Free Survival
CC [redacted]	[redacted]
PR	Partial Response
PT	Preferred Term
QoL	Quality of Life
QTcF	QT interval corrected using Fridericia's formula
RECIST	Response Evaluation Criteria in Solid Tumors
RS	Raw Score
RTF	Rich Text Format
SAE	Serious Adverse Event
SAF	Safety
SD	Standard Deviation
SD	Stable Disease
SDTM	Study Data Tabulation Model
SI	Standard International System of Units
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
TDM	Time to Distant Metastasis
TEAE	Treatment Emergent Adverse Event
TFTI	Time to Feeding Tube Insertion
TRAE	Treatment-Related Adverse Event
ULN	Upper Limit of Normal
VAS	Visual Analogue Scale
WHO-DD	World Health Organization Drug Dictionary
CCI [redacted]	[redacted]

3 Modification History

Unique Identifier for Version	Date of IAP Version	Author	Changes from the Previous Version
1.0	25 OCT 2022	PPD	n.a.
2.0	27 OCT 2022		Administrative changes (title page)
3.0	07 DEC 2023		Implementation of health authority feedback and adding clarifying languages for the cut-off date determination of the interim and primary analysis. Adding new analyses or details to existing analyses following country-specific and general submission requirements (feeding tube, cIPD, Time to CR, AESI, CCI).
4.0	08 DEC 2023		Administrative changes (Links in 15.2.3.1). Some editorial changes.
5.0	07 MAR 2024		Update of the sequence of analyses (IAP section 6) according to the Clinical Study Protocol Version 12.0

4 Purpose of the Integrated Analysis Plan

The purpose of this integrated analysis plan (IAP) is to document technical and detailed specifications for all the planned analyses of data collected for protocol Debio 1143-SCCHN-301 / MS202359_0006. Results of the analyses described in this IAP will be included in the Clinical Study Report (CSR), except when otherwise stated. Additionally, the planned analyses identified in this IAP may be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in the CSR but not identified in this prospective IAP will be clearly identified in the CSR. Candidate predictive biomarkers will be analyzed exploratory only and described in a separate analysis plan.

The IAP is based upon Section 10 (Statistical Considerations) of the study protocol dated 11-April-2023/version 10.0 and is prepared in compliance with the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) E9.

Details on outputs used to support review of the study by an external Independent Data Monitoring Committee (IDMC) are available in the IDMC TFL shell document and further detail planning included in the IDMC charter 13-Jun-2022 / version 3.0.

The wording used in this IAP is chosen to best match the respective wording in the study protocol template, the Clinical Study Report (CSR) template, CDISC requirements and special requirements for table layouts. Therefore, the following approach is used:

Generally, the term ‘participant’ will be used instead of ‘subject’ or ‘patient’. However, in tables and listings the term ‘subject’ will be used to match CDISC requirements, except for in-text tables where ‘participant’ will be used to match the CSR and protocol templates. Similarly, the term

'study intervention' and 'study treatment' will be used synonymously in this document, tables and listings will use 'treatment' for brevity reasons.

5 Objectives and Endpoints

The objectives and corresponding endpoints are presented in Table 1.

Table 1: Objectives, endpoints and statistical considerations

Objectives	Endpoints	Hypothesis	Primary Analysis Population	Timeframe/Time Point
Primary				
To demonstrate superior efficacy of xevinapant vs placebo when added to chemoradiotherapy (CRT) in locally advanced squamous cell carcinoma of the head and neck (LA-SCCHN)	Event-free survival (EFS) as assessed by the Blinded Independent Review Committee (BIRC) defined as the time from randomization to the first occurrence of any of the following events: • Death from any cause. • Progression: o Radiological, assessed per RECIST v1.1 or o Clinical, with or without RECIST v1.1-radiologically documented progression, assessed endoscopically. • Primary treatment failure before achieving a complete response (CR): requirement for radical salvage surgery that includes the primary tumor site, with documented viable tumor presenting anatomopathological findings even in the absence of formal RECIST v1.1 radiological progression. • Any radiological or clinical relapse after achieving a CR (locoregionally), including any event defined as locoregional treatment failure, even in the absence of formal radiological progressive disease confirmation: o Requirement for radical salvage surgery that includes the primary tumor site, regardless of anatomopathological findings or o Elective neck dissection or biopsy, with positive viable tumor cells on anatomopathological findings at 22 weeks or later after randomization. • Second cancers unless anatomopathological findings exclude squamous histology. Note: Investigator-assessed EFS will be used for supportive analysis.	Xevinapant prolongs EFS in previously untreated patients with LA-SCCHN, compared to matched placebo when administered in combination with platinum-based chemotherapy and intensity-modulated radiotherapy (IMRT).	Intent-to-treat	- From randomization to the earliest between any EFS event or end of study (EOS). - Rate at 9, 12, 18, 24, 30 and 36-months post-randomization.

Objectives	Endpoints	Hypothesis	Primary Analysis Population	Timeframe/Time Point
Secondary				
To assess the efficacy of xevinapant compared to placebo when added to CRT according to additional efficacy endpoints	Overall survival (OS) defined as the time from randomization to death due to any cause.	Xevinapant prolongs OS in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From randomization to the earliest between death or EOS - Rate at 12, 24, 36, 48 and 60 months post-randomization.
	Progression-free survival (PFS) according to RECIST v1.1 defined as the time from randomization to the first occurrence of progression (radiological or clinical, as assessed by the BIRC) or death from any cause. Note: Investigator-assessed PFS will be used for supportive analysis.	Xevinapant prolongs PFS in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From randomization to the earliest between PFS event or EOS. - Rate at 9, 12,18, 24, 36, 48 and 60 months post-randomization.
	Locoregional control (LRC) defined as the time from randomization to the first occurrence of progression at the site of the primary tumor or the locoregional lymph nodes, either according to RECIST v1.1 or based on clinical assessment (radiological or clinical, as assessed by the Investigator). Note: Will be investigated by BIRC as supportive analysis.	Xevinapant prolongs LRC in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From randomization to the earliest between PFS event or EOS. - Rate at 9, 12, 18, 24, 36, 48 and 60 months post-randomization.
	Objective response rate (ORR), defined as the proportion of subjects with CR or partial response by RECIST v1.1, as assessed by the BIRC. Note: Investigator-assessed response will be used for supportive analysis.	Xevinapant increases the ORR in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From randomization to the earliest between progression, start of a new anticancer therapy for this disease or participant EOS. - Rate at 9 and 12 months post-randomization.

Objectives	Endpoints	Hypothesis	Primary Analysis Population	Timeframe/Time Point
	CR rate (CRR) defined as the proportion of participants with CR by RECIST v1.1, as assessed by the BIRC. Note: Investigator-assessed response will be used for supportive analysis.	Xevinapant increases the CRR in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From randomization to the earliest between progression, start of a new anticancer therapy for this disease or subject EOS. - Rate at 9 and 12 months post-randomization.
	Duration of response (DoR) defined as the time from the first evidence of response (partial or complete, as assessed by the BIRC according to RECIST v1.1) to the first occurrence of progression (radiological or clinical, as assessed by the BIRC) or death from any cause. Note: Investigator-assessed DoR will be used for supportive analysis.	Xevinapant prolongs DoR in previously untreated participants with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	- From first evidence of response (partial or complete) to the earliest between DoR event or EOS. - Rate at 6, 12 and 24 months post first evidence of response.
	Proportion of participants with radical salvage surgery (excluding elective neck dissection without anatomopathological evidence of residual malignant cells).		Intent-to-treat	• From randomization to the earliest between radical salvage surgery or EOS. • Rate at 9, 12, 18, 24, 36, 48 and 60 months post-randomization.
	Time to subsequent systemic cancer treatment.		Intent-to-treat	- From randomization to EOS - Rate at 12, 18, 24, 36, 48 and 60 months post-randomization.

Objectives	Endpoints	Hypothesis	Primary Analysis Population	Timeframe/Time Point
To compare safety, tolerability and treatment compliance of xevinapant vs placebo, when added to CRT	Incidence and severity of adverse events (AE), serious adverse events and adverse events of special interest, changes in laboratory values, vital signs, and electrocardiograms according to NCI-CTCAE v 5.0.		Safety	• From signed informed consent to EOS.
	Extent of exposure of the different treatment agents (i.e., xevinapant or matched placebo, radiotherapy, chemotherapy) including: • Treatment duration. • Number of cycles. • Actual dose. • Dose intensity. • Relative dose intensity. • Incidence of treatment interruption. • Incidence of treatment reduction. • Incidence of treatment discontinuation.		Safety	• From first study treatment dose to EOT.
To compare the health-related quality of life of xevinapant vs placebo when added to CRT using patient-reported outcome questionnaires	Changes from baseline in: • Global Health Scale /Quality of Life and Fatigue using European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 questionnaire. • Swallowing and Pain using EORTC QLQ-HN35 questionnaire.	Xevinapant is non-inferior for Global Health Scale /Quality of Life, Fatigue, Swallowing and Pain at Month 12 post-randomization in previously untreated patients with LA-SCCHN compared to matched placebo when administered in combination with platinum-based chemotherapy and IMRT.	Intent-to-treat	• From randomization to EOS. • At baseline, Study Week 10, 20 and months 7, 12, 24, 36, 48 and 60 post-randomization.


CCI

CCI

CCI

6 Overview of Planned Analyses

This IAP covers the analyses based on the data cut-off dates for the interim analyses, the primary analysis, and the final analysis. Statistical analyses will be performed using cleaned eCRF data as well as external data. All data will be included up to a data cut-off date which is determined by the number of events for EFS/OS required for the analysis. Refer to Table 2 for the respective timing of the analyses.

Table 2: Planned Efficacy Analyses and Expected Timing

Analysis	Events (IF)	Assumed Median		Expected month[a] after FPI	Notes
		Arm A	Arm B		
EFS IA1	279 (65%)	23.3	17	30	
EFS IA2	365 (85%)			37	skipped if H_0^{EFS} could be rejected at EFS IA1
EFS PA	429 (100%)			44	skipped if H_0^{EFS} could be rejected at EFS IA1 or EFS IA2
OS IA1	154 (40%)	49.3	37	30	only if H_0^{EFS} could be rejected
OS IA2	216 (56%)			37	only if H_0^{EFS} could be rejected, but not H_0^{OS} at OS IA1
OS IA3	270 (70%)			44	only if H_0^{EFS} could be rejected, but not H_0^{OS} at OS IA1 or OS IA2
OS IA4	309 (80%)			50	only if H_0^{EFS} could be rejected, but not H_0^{OS} at OS IA1, IA2 or IA3
OS FA	386 (100%)			65	only if H_0^{EFS} could be rejected, but not H_0^{OS} at OS IA1, IA2, IA3 or IA4

Abbreviations: **EFS:** event-free survival, **FPI**: date of first participant randomized; **IA**: interim analysis; **IF**: information fraction; **OS:** overall survival; **PA**: primary analysis.

[a] Based on exponential distributions using given medians for both arms and a 22-month-recruitment-period.

Analyses will be combined by using the same data cut-off date for EFS IA1/OS IA1, EFS IA2/OS IA2 as well as for EFS PA/OS IA3, which will be triggered by the respective numbers of EFS events as mentioned above or maximum additional follow-up time (refer IAP section 6.2), while the significance boundaries are adjusted accordingly to control the family-wise error rate (FWER) at 2.5% one-sided. However, in case the null hypothesis for EFS can already be rejected at an EFS interim analysis (IA) and the Sponsor decides to unblind the study, confirmatory analysis of EFS is completed with the respective EFS IA, and any further OS IA will be triggered by the number of OS events. In case the subsequent interim efficacy analyses will be triggered by the number of OS events as given in Table 2, it will be ensured to have an addition of at least 39 OS events to the

event count of the previous interim analysis (about 10% increase in information fraction) but not later than 12 months after the previous interim analysis cut-off date.

A hierarchical testing procedure will be used to test first primary null hypothesis H_0^{EFS} in a group sequential design, which uses alpha-spending according to Lan-DeMets with O'Brian-Fleming-like boundaries planning for a total number of 429 EFS events at primary analysis. If and only if that can be rejected either on interim or primary analysis, the null hypothesis H_0^{OS} will be confirmatory tested in a group sequential design with interim analysis as planned in Table 2, which uses alpha-spending according to Lan-DeMets with O'Brien Fleming-like boundaries planning for a total number of 386 OS events at final analysis. Table 3 shows the operating characteristics and decision boundaries for the confirmatory interim and primary/final analyses as planned.

Table 3: Operating Characteristics and Power Estimations

		Boundaries (HR)		Boundaries (P-value)		Local	
Analysis	Events (IF)	Efficacy	Futility	Efficacy	Futility	Power[a]	Power[a]
EFS IA1	279 (65%)	0.737	0.886	0.005	0.163	53.3%	
EFS IA2	365 (85%)	0.793	-	0.013	-	25.8%	
EFS PA	429 (100%)	0.821	0.821	0.020	0.020	9.1%	88.2%
OS IA1	154 (40%)	0.583	-	0.0004	-	5.7%	
OS IA2	216 (56%)	0.684	-	0.003	-	19.4%	
OS IA3	270 (70%)	0.739	-	0.007	-	21.1%	
OS IA4	309 (80%)	0.767	-	0.010	-	13.5%	
OS FA	386 (100%)	0.813	0.813	0.021	0.021	19.9%	79.6%

Abbreviations: EFS: event-free survival; **HR**: hazard ratio; **IA**: interim analysis; **IF**: information fraction; **OS**: overall survival; **PA**: primary analysis.

[a] Power estimations are based on an assumed HR=0.73 for EFS and a HR=0.75 for OS and the prerequisite that the EFS null hypothesis can be rejected at either IA or PA. Futility-bounds are non-binding.

Note: Decisions will be based on p-values with respect to actual number of events. Expected HRs under the assumption of an exponential distribution are only given as further information.

6.1 Safety Monitoring Analysis

The IDMC will review quarterly descriptive summaries of accumulating participants' disposition and safety. The analyses will be conducted by unblinded supporting staff independent from the study team. The investigators/site personnel, CRO and sponsor team will remain blinded, except for a dedicated unblinded statistical team at the CRO that is preparing analyses for the IDMC.

The primary intent of these analyses is the patient safety monitoring. IDMC may recommend change or early stop of the study in case of safety issue.

Details on outputs used to support the review by the IDMC are available in the IDMC TFL shell document.

6.2 Interim Analyses

A first interim analysis is planned once 279 EFS events as assessed by the BIRC (65% of EFS events required for the primary analysis) across both arms are observed. The number of EFS events as assessed by the BIRC will be monitored throughout the study. The cutoff date will be determined prospectively based on an event projection as the date when approximately 279 events will be reached. In case it turns out at the time of the EFS IA1 database lock that the observed number of EFS events is >3.5% below the planned number (<270), the IA cut-off date will be revised prior to the conduct of the analysis, based on blinded information on the number of EFS events only, to include at least 270 events in the IA dataset. The second interim analysis will be triggered once 85% of the total number of events (365 events) are observed or one year after the first interim analysis, whichever occurs first. The corresponding cutoff date will be re-evaluated based on event projections.

The interim analyses will be conducted for the IDMC by unblinded supporting staff independent from the study team and will focus on primary and selected secondary efficacy endpoint analyses. The Investigators/site personnel, Sponsor team and CRO will remain blinded, except for a dedicated unblinded statistical team at the CRO that is preparing analyses for the IDMC.

The main intent of these analyses is to demonstrate that xevinapant + CRT prolongs EFS as assessed by the BIRC versus placebo + CRT. No early stopping for superiority of efficacy is planned. Provided that EFS interim analysis results are positive, the data might be used for regulatory and health technology assessment submission purposes, in case of which the CRO and Sponsor team may be unblinded. Productions of outputs related to the analyses of the exploratory endpoints will be done only in case of positive recommendation from the IDMC, as needed. The IDMC will first share recommendation(s) based on the results of the interim efficacy analyses to the Firewall Team and jointly discuss results, as needed, for final Sponsor decision of unblinding according to the protocol, keeping all individuals involved in the study conduct blinded to preserve its integrity. If the study is futile or if EFS is detrimental at the IA, IDMC may recommend early stop of the study providing an observed lack of efficacy and/or futility putting the patients at risk when continuing the study.

For the primary efficacy endpoint, a group-sequential design with α-spending according to Lan-DeMets with O'Brian-Fleming-like boundaries will be used to control the type I error rate. The observed number of EFS events as assessed by the BIRC at the time of the interim analysis will be used to recalculate the exact efficacy and futility boundaries to be used for the interim analysis as well as for the primary analysis (see also Table 3).

If the null hypothesis H_0^{EFS} of the primary efficacy endpoint EFS can be rejected, the null hypothesis H_0^{OS} of the secondary efficacy endpoint OS will be confirmatory tested by a group-sequential design with α-spending according to Lan-DeMets with O'Brian-Fleming-like boundaries to control for the type I error rate. The observed number of OS events at the time of the interim analysis will be used to recalculate the exact efficacy boundaries to be used for this interim analysis as well as for the further planned analyses (see also Table 3).

6.3 Primary Analysis

The primary analysis will be triggered once 429 EFS events across both arms (as assessed by the BIRC) are observed. This is an event driven analysis and monitoring of the number of events will be done throughout the study. The cutoff date will be determined prospectively based on an event projection as the date when approximately 429 EFS events will be reached. In case the observed number of EFS events is >4.5% below the planned number (<410), the PA cut-off date will be revised prior to the conduct of the analysis, based on blinded information on the number of EFS events only, to include at least 410 EFS events into the primary analysis.

The primary intent of this analysis is to demonstrate that xevinapant + CRT prolongs EFS as assessed by the BIRC versus placebo + CRT in case this has not been demonstrated at the interim analysis yet. The observed numbers of EFS events at the interim and primary analysis will be used to recalculate the exact efficacy and futility boundaries (see also Table 3).

If the null hypothesis H_0^{EFS} of the primary efficacy endpoint EFS can be rejected at the interim or primary analysis and the null hypothesis H_0^{OS} of the secondary efficacy endpoint OS was not rejected at the interim analysis, the null hypothesis H_0^{OS} of the secondary efficacy endpoint OS will be confirmatory tested by a group-sequential design with α-spending according to Lan-DeMets with O'Brian-Fleming-like boundaries to control for the type I error rate. The observed number of OS events at the time of the primary analysis will be used to recalculate the exact efficacy boundaries to be used for the primary analysis as well as for the final analysis (see also Table 3).

In case the null hypothesis for EFS can already be rejected at an EFS interim analysis (IA) and the Sponsor decides to unblind the study, confirmatory analysis of EFS is completed with the EFS IA, and subsequent OS IAs will be triggered by the number of OS events.

Other secondary efficacy endpoints and exploratory endpoints will be analyzed without any type I error adjustment. All provided p-values for these analyses will be exploratory.

6.4 Further OS Analyses

A fourth interim analysis for OS will be triggered once 309 OS events are observed across arms. in case there is an increase of at least 39 OS events compared to the event count of the previous interim analysis (~10% increase of information fraction) but not later than 12 months after the previous interim analysis cut-off date. Only endpoints affected by longer follow-up will be analyzed.

The final analysis will be triggered once 386 OS events are observed across both arms (assumed to be 65 months after the first participant was randomized) or the last on-study participant has either reached the 60-month post-randomization follow-up visit or prematurely discontinues, whichever happens first. Only endpoints affected by longer follow-up will be analyzed.

The primary intent of these analyses is to demonstrate that xevinapant + CRT prolongs OS versus placebo + CRT.

If the null hypothesis H_0^{EFS} of the primary efficacy endpoint EFS can be rejected at interim or primary analysis and the null hypothesis H_0^{OS} of the secondary efficacy endpoint OS was not rejected at the interim analysis nor the primary analysis, OS will be confirmatory tested by a group-sequential design with α-spending according to Lan-DeMets with O'Brian-Fleming-like boundaries to control for the type I error rate at the time of final analysis. The observed number of OS events at interim, primary and final analysis will be used to recalculate the efficacy boundaries for H_0^{OS} (see also Table 3).

7 Changes to the Planned Analyses in the Clinical Study Protocol

The primary HRQoL Analysis will focus on the change from baseline at end of combination therapy and the completion of investigational therapy rather than evaluating non-inferiority at Month 12 as stated in Section 3 of the study protocol. Data from assessments after completion of investigational therapy will be presented purely descriptively. The rationale for this approach is that the assessments during the on-treatment period are expected to reflect changes in HRQoL due to side effects of the study treatment, whereas a substantial proportion of missing data is expected at later timepoints such as Month 12, because PRO assessments are not continued after an EFS event.

No sensitivity/supportive analyses (except DoR by investigator assessment) will be performed for the secondary efficacy endpoint duration of response (DoR) due to the reason that the analysis of DoR is not based on the ITT and must be seen itself as a supportive analysis. For this reason, DoR will be analyzed by treatment arms but without any between-arm comparisons.

The list of timepoints for the analysis of the proportion of participants with enteral and parenteral feeding (Section 14.3.2 of this SAP) has been revised to match the schedule of assessments for nutritional status in the study protocol.

In case the difference between the ITT and safety sets is less than 5%, demographics and baseline characteristics will only be assessed and summarized by arm in the ITT set using summary tables and figures.

No changes to the planned analysis of the efficacy or safety endpoints are planned due to the impact of Coronavirus disease 2019 (COVID-19) outbreak. Sensitivity analyses may be considered depending on the observed COVID-19 findings.

Additional outputs (summary table and listing) will be generated for a description of the impact by COVID-19 on the study. The number and percentage of participants will be presented for the following findings due to COVID-19:

- Adverse Events (MedDRA SMQ "COVID-19")
- Important protocol deviations
- Missed Visits (including number of missed visits and detailed reason)
- Missed efficacy evaluations (including number of missed efficacy evaluations)
- Drug administration - missed doses

- Drug administration - dose adjustments
- Treatment discontinuation
- Study discontinuation
- Death

A frequency table will be produced for the SAF analysis set to present the number of participants with important protocol deviations related to COVID-19 (categorized by frequency of participants with an important protocol deviation overall as well as by category of protocol deviation and type of protocol deviation).

In addition, important COVID-19 protocol deviations will also be listed.

Outputs related to disposition and exposure will be amended to present reason of treatment/study discontinuations due to COVID-19 and treatment changes due to COVID-19 (if possible).

8 Analysis Sets and Subgroups

8.1 Definition of Analysis Sets

The following analysis sets will be used for the planned analysis of the global study population: the screened set, the intent-to-treat (ITT) set, and the safety (SAF) set.

Screened Set

The Screened set will include all participants having signed the Informed Consent Form (ICF). Unless otherwise specified, this set will be used for disposition (summary and eligibility participant data for the Consort disposition diagram).

Intent-to-Treat Set

The ITT set includes all randomized participants. Participants will be analyzed according to the randomized treatment (assigned arm) assignment following the intention-to-treat principle. The ITT set will be the primary set for all summaries and analyses of efficacy.

Safety Set

The SAF set includes all participants who received any dose of any of the study intervention (xevinapant/matched placebo, cisplatin/carboplatin, IMRT). Any participants who are randomized to the placebo + CRT group but who are incorrectly administered any dose of xevinapant will be analyzed in the xevinapant + CRT group. Any participant randomized to the xevinapant + CRT who incorrectly receives only placebo rather than xevinapant will be analyzed in the placebo group. The SAF set will be used for all safety analyses and baseline characteristics.

Pharmacokinetic Set

The PK set will include all the participants who have undergone at least one PK sample scheduled. The PK set will be used for all PK analyses.

	Analysis Set			
Analyses	**Screened Set**	**ITT Set**	**SAF Set**	**PK Set**
Participant Disposition	✓			
Demographics and Baseline Characteristics		✓		
Medical/Surgical History		✓		
Prior Medications		✓		
Concomitant/Subsequent Medications		✓		
Surgical Procedures		✓		
Concomitant 5-HT3 receptor antagonist medication		✓		
Compliance and Exposure			✓	
Efficacy: Primary Endpoint		✓		
Efficacy: Secondary Endpoints		✓		
Safety and Tolerability			✓	
Health-Related Quality of Life		✓		
Health Resource Utilization		✓		
Pharmacokinetics (PK)				✓

8.2 Subgroup Definition and Parameterization

Subgroup analyses will be performed on the primary efficacy endpoint EFS (BIRC assessed) as well as on the secondary efficacy endpoints OS, PFS (BIRC assessed), LRC (investigator assessed), DoR (BIRC assessed), CR (BIRC assessed), and for the exploratory endpoint of time to distant metastasis (investigator assessed) as defined below. Subgroup analyses on these efficacy endpoints, as well as the demographic and baseline data, and the safety data, might be performed for specific regions defined below among "Rest of the World", if the data is to be used for regulatory submission purposes based on the interim, primary, or final analysis. All subgroup analyses will be exploratory, no adjustment for multiplicity will be performed.

For the definition of subgroup level, data as documented in the eCRF will be taken. The category "missing" will not be included in any subgroup analysis.

In case of low number of participants within a category (about < 5% of the randomized population), categories will be pooled when meaningful.

The following subgroups based on baseline characteristics will be considered.

8.2.1 Randomization stratification factors (from IWRS)

The following subgroups are relevant for all endpoints listed in Section 8.2.

Region

- North America
- Western Europe (as defined in Appendix 4) (reference level)
- Rest of the World

Primary tumor site

- Larynx
- Other (Oropharynx and Hypopharynx) (reference level)

Lymph node involvement

- N0-1 (reference level)
- N2
- N3

T size

- T4
- Other (T1, T2, T3) (reference level)

8.2.2 Other subgroups

The following subgroups are only relevant for the primary efficacy endpoint EFS (BIRC assessed) and the secondary efficacy endpoint OS.

Age Group 1

- Age < 65 years (reference level)
- Age ≥ 65 years

Age Group 2

- Age < 70 years (reference level)
- Age ≥ 70 years

Sex

- Male (reference level)
- Female

Race

- Asian
- American Indian or Alaska Native
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White (reference level)
- Other

Ethnicity

- Hispanic or Latino
- Not Hispanic or Latino (reference level)

Note: Race or Ethnicity category “not reported” will be handled as “missing”

Primary tumor site (as recorded in the eCRF):

- Larynx (reference level)
- Hypopharynx
- Oropharynx

Eastern Cooperative Oncology Group Performance Status (ECOG-PS)

- 0 (reference level)
- ≥1

American Joint Committee on Cancer (AJCC) cancer staging

- III (reference level)
- IVA
- IVB

Alcohol

- current use
- no current use (reference level)

Smoking

- current use (reference level)
- history of use
- no history of use

8.2.3 Analyses in Asian Participants

In order to satisfy requirements of the Japanese Pharmaceuticals and Medical Devices Agency as well as the requirements for an ethnicity report, a subset of data summaries will be produced for each of the countries Japan, Korea, Taiwan compared to the respective complementary set by treatment group and East Asian (pooled including Japan, China, Taiwan and Korea). These will include summaries of:

- Participant disposition
- Participants randomized vs treated
- Analysis sets
- Demographic characteristics
- Disease history
- Subsequent anti-cancer drug treatments
- Primary Endpoint: Event free survival (EFS),
- Secondary Endpoints: overall survival (OS), progression-free survival (PFS), local regional control (LRC), objective response (OR), complete response (CR), duration of response (DoR), proportion of subjects with radical salvage surgery, time to subsequent systemic cancer treatments and HRQoL.
- Treatment exposure
- Adverse events (overall, related, grade ≥ 3, related grade ≥ 3, serious, leading to death, related leading to death)

In order to meet the requirement of regulatory consultation and submission to the Chinese National Medical Products Administration, more detailed information on the analyses in Chinese participants are described in Appendix 8.

9 General Specifications for Data Analyses

All statistical analyses will be conducted using SAS® for Windows, Release 9.4 or higher (SAS® Institute Inc., Cary, NC, USA). All outputs will be generated in Rich Text Format (RTF) from Microsoft Word and combined pdf file formats. Further programming specifications can be found in the mock shells.

Unless otherwise specified, summaries will be presented for each treatment arm (labelled xevinapant + CRT and Placebo + CRT) and overall (labelled Total, when appropriate). Unless stated otherwise, participant data listings will be sorted by treatment arm (xevinapant + CRT and Placebo + CRT), participant number, and assessment date (and time).

Unless otherwise specified, continuous variables will be summarized using descriptive statistics, i.e., the number of participants with non-missing values (n), mean, standard deviation, median, 25th percentile (Q1) and 75th percentile (Q3), minimum, and maximum.

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise stated the calculation of proportions will be based on the number of participants of the analysis set of interest. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

Descriptive statistics by nominal visit or time point, e.g., for laboratory measurements, will include only data from scheduled visits. Unscheduled visits will be included in the derivation of baseline or worst on-treatment values.

9.1 Data handling after Cut-off Date

Data after the data cut-off date do not undergo the cleaning process. The only exceptions are the date of death and the date last known to be alive from the eCRF.

Data obtained after a cut-off will not be displayed in any listings or used for summary statistics, e.g., laboratory values of samples taken after data cut-off, AE with onset date after data cutoff, etc. will not be included in any analysis or listing.

These rules will be applied to all analyses performed for the interim analyses and primary analysis. For the final analysis no cut-off date will be applied: the analysis will be performed only after all the data have been collected, fully cleaned and the database has been locked.

9.2 Definition of Baseline and Change from Baseline

For efficacy analyses, baseline will be defined as the last non-missing measurement prior to randomization. If no such value is available, the last measurement prior to the first study treatment will be used as the baseline measurement.

For safety analyses, the last non-missing measurement prior to the first study treatment will serve as the baseline measurement.

If an assessment that is planned to be performed before randomization, or study treatment per protocol is performed on the same day as the randomization or start of study treatment, respectively, but the assessment time is not available, it will be assumed that it was performed prior and will be considered for derivation of baseline.

If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be analyzed similar to an unscheduled post-dose measurement.

If both central and local labs are collected, the baseline will be derived based only on the central lab collected data.

Absolute and percent changes from baseline are defined as

- absolute change = visit value – baseline value
- percent change = 100 * (visit value – baseline value) / baseline value

9.3 First Dose Date

First dose date is defined as the day of **first** administration of any study intervention (xevinapant/matched placebo, chemotherapy (cisplatin and/or carboplatin) or IMRT treatment (background radiotherapy)).

It is expected that on C1D1 xevinapant or matched placebo is given first and followed by IMRT. First dose of chemotherapy is scheduled on C1D2, after xevinapant/matched placebo dose and before IMRT.

9.4 Last Dose Date

Last dose date is defined as the latest of last administration of any study intervention (xevinapant/matched placebo, chemotherapy (cisplatin and/or carboplatin) or IMRT).

9.5 Definition of On-treatment Period

The on-treatment period is defined as any time from first dose date until the last dose date + 30 days.

9.6 Study Day / Study Treatment Day

Day 1 is the day of randomization/start of study treatment, the day before is Day -1 (no Day 0 is defined). Study day / Study treatment day is defined relative to Day 1.

9.7 Definition of Duration and 'Time Since' Variables

Durations in days will be calculated by the difference of start and stop date + 1 (e.g., survival time (days) = date of death – date of randomization + 1) if not otherwise specified.

Treatment duration of each study intervention (xevinapant/matched placebo, chemotherapy (cisplatin and/or carboplatin) or IMRT) is calculated based on first dose date and last dose date of the respective study intervention component as ceiling (last dose date – first dose date + x) divided by 7 for display in weeks, where x is equal to 8 for xevinapant/matched placebo and chemotherapy and x is equal to 3 for IMRT.

The overall treatment duration is defined as max(xevinapant/matched placebo treatment duration, CT treatment duration, IMRT treatment duration).

The time since an event (e.g., time since first diagnosis) will be calculated as reference date minus date of event.

9.8 Conversion Factors

The following conversion factors will be used to convert days into months or years:

- 1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

9.9 Date of Last Contact

The date of last contact will be derived for participants not known to have died at the analysis cut-off using the latest complete date prior to or at the data cut-off date among the following:

Participants on study treatment/during EFS follow-up

- AE start dates
- Date of last study treatment
- Date of last tumor assessment (eCRF form: CT scan or MRI)

Participants during survival follow-up

- Last known to be alive date collected on the Survival Follow-up eCRF page

Only dates associated with actual examinations of the participant reported in the eCRF will be used in the derivation. Assessment dates after the cut-off date will not be applied to derive the last contact date.

9.10 Time Window

Time window approach will not be used in this study.

9.11 Time on Study

Time on study, which will be used for IDMC output only, is defined in relation to the last assessment date at the time of data cut-off (considering last assessment date – randomization date +1).

The last assessment date is the last available assessment that indicates the patient is alive or dead. It includes end of study visit, safety assessment and/or phone call. If a participant dies, the last assessment date is the date of death.

9.12 Imputation of Missing Data

Unless otherwise specified, all data will be evaluated as observed, and no imputation method for missing values will be used.

Missing statistics, e.g., when they cannot be calculated, should be presented as "nd". For example, if n=1, the measure of variability (standard deviation [SD]) cannot be computed and should be presented as "nd".

Imputation rules for missing efficacy data are described in Section 14.

9.12.1 Partial and Missing Dates

All missing or partial dates will be presented in the participant data listing as they are recorded on the eCRF.

Adverse Events

In case the AE start date is missing or incomplete, the date will be imputed for the calculation according to the worst case approach. If the AE start date is completely missing it will be imputed with the first dose date. For incomplete AE start dates the missing day will be imputed by the first day of the month and the missing month will be imputed as January. If the imputed AE start date is prior to the first dose date, the AE start date will be set to the first dose date. Exceptions are given where the partial AE start dates would allow to state otherwise. Incomplete AE stop dates will be imputed by the last day of the month (if day is missing only), if not resulting in a date later than the date of participant's death. In the latter case, the date of death will be used to impute the incomplete AE stop date. In all other cases, the AE incomplete stop date will not be imputed.

Prior, Concomitant, and Subsequent Medications

In case the prior, and concomitant, or subsequent medication start date is partially missing, the day will be imputed by the first day of the month and the missing month will be imputed as January. For an incomplete stop date, the missing day will be imputed with the last day of the month, and the missing month will be imputed with December.

Dates of subsequent anticancer treatment

Incomplete start dates for the subsequent anticancer treatment will be imputed as follows and will be used for determining censoring dates for efficacy related sensitivity analyses:

- If only day is missing, it will be imputed as the first day of the month unless it results in a date before the last dose date. In that case, the incomplete anticancer treatment start date will be imputed as the last dose date.

- If both day and month are missing, the incomplete anticancer treatment start date will be imputed as the last dose date if the subsequent anticancer treatment started on the same year as the last dose date; otherwise, day and month will be imputed as the 1st of January.

Incomplete subsequent anticancer treatment stop dates will not be imputed.

9.12.2 Tumor assessments

All investigation dates (e.g., CT scan, MRI) must be completed with day, month and year.

If there are multiple scan dates associated with an evaluation, i.e., radiological assessments occur over a series of days rather than the same day, the choice of date of assessment could impact the date of progression and/or date of response. If there are multiple scan dates associated with an evaluation, the earliest of the scan dates associated with the evaluation will be used as the date of assessment.

If one or more investigation dates for an evaluation are incomplete but other investigation dates are available, the incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the earliest of all investigation dates (e.g., CT scan, MRI).

If all measurement dates for an evaluation have no day recorded, the 1st of the month is used.

If the month is not completed, for any of the investigations for an evaluation, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both a previous and following assessments are not available, this assessment will not be used for any calculations.

9.12.3 Tumor Evaluation and Adjudication

Missing data from scheduled Tumor assessments not performed will not be imputed. Evaluable Tumor assessments are defined as those with overall response CR, PR, SD or PD for patients with measurable disease or CR, Non-CR/non-PD, or PD for patients with non-measurable disease only at baseline. Data from non-evaluable (NE) Tumor assessments will be used in the derivation of endpoints in accordance with Eisenhauer et al., (2009).

In case BIRC assessments are performed only the response assessments which are flagged as accepted after adjudication will be taken over to ADaM datasets and will be analyzed. In case of a missing adjudication flag for a response assessment and earliest image dates are equal, the assessment of reader 1 (the reader who completed baseline first) will be taken for analysis. Otherwise, the assessment for the reviewer who reviewed the record with the earliest image date will be analyzed. Analyses of tumor sizes will be based only on assessments of the reader whose assessment was accepted at baseline.

9.13 Unscheduled Assessments or Retests

Unscheduled assessments will not be used in the summary statistics per timepoint but will be used in the derivation of endpoints that are not related to a specific timepoint. All values collected will be displayed in the participant data listings.

9.14 Scoring of Health-Related Quality of Life (HRQL) Data

Unless otherwise specified, HRQL items will be scored using their published administration and most current scoring manual. For items with missing responses, the response will be managed as per the scoring manual.

9.15 Derivation of Best Overall Response

Best overall response will be assessed based on reported overall responses at different evaluation time points from randomization until documented disease progression in accordance to RECIST v1.1 as assessed by the BIRC or the investigator, respectively.

Only tumor assessments performed prior to the start of any subsequent anticancer treatment will be considered in the assessment of best overall response. If a tumor assessment was performed on the same day as the start of a new anticancer treatment, it will be assumed that the tumor assessment was performed prior to the start of the new anticancer treatment, therefore the tumor assessment will be included in the assessment of best overall response.

Best overall response (derivation):

- CR = at least one assessment of objective status of CR documented before progression.
- PR = at least one assessment of objective status of PR documented before progression (and not qualifying for CR).
- SD (applicable only to participants with measurable disease at baseline) = at least one SD assessment (or better) ≥ 6 weeks after randomization and before progression (and not qualifying for CR or PR).
- Non-CR/non-PD (applicable only to participants with non-measurable disease at baseline) = at least one non-CR/non-PD assessment (or better) ≥ 6 weeks after first dose date and before progression (and not qualifying for CR or PR).
- PD = progression latest at the time of the second scheduled post-baseline assessment, i.e., Day 1 of Month 7 plus 2 weeks (and not qualifying for CR, PR, SD or non-CR/non-PD). The condition PD in or out of the respective timeframe addresses a missing data situation in view of determining best overall response as PD or NE, respectively.
- NE: all other cases.

10 Study Participants

The subsections in this section include specifications for reporting participant disposition and study treatment/study discontinuations. Additionally, procedures for reporting protocol deviations are provided.

10.1 Disposition of Participants and Discontinuations

10.1.1 Screen Failures and Eligibility Criteria

Screen failures will be summarized overall based on all screened participants. The number of participants screened, and screen failures will be presented. In addition, the number and percentage of screen failures by reason will be displayed.

A participant data listing for all participants screened, along with whether the participant was a screening failure and if yes, the reason for the screen failure will be provided. Additionally, a second participant data listing will be provided showing whether the participant continued to meet eligibility to treatment prior to dosing, and if not, which eligibility to treatment was violated.

10.1.2 Participant Disposition

A participant is considered enrolled when randomized.

Participant disposition will be summarized by treatment arm and overall, for all participants based on Screened Set. The following will be summarized:

- Total number of participants screened (i.e., participants who gave informed consent).
- Number of randomized participants.
- Number of randomized participants who did not receive any study intervention.
- Number and percentage of participants who received at least one dose of any study intervention components.
- Number and percentage of treated participants who completed the treatment period.
- Number and percentage of treated participants who discontinued the treatment overall and by with the primary reason of discontinuation.
- Number and percentage of participants who continued into the EFS follow-up period.
- Number and percentage of participants who completed the EFS follow-up period until month 60 without EFS event.
- Number and percentage of participants who discontinued the EFS follow-up period overall and by primary reason for EFS follow-up period discontinuation.
- Number and percentage of participants who continued into the OS follow-up period.
- Number and percentage of participants who discontinued from study overall and by primary reason for study discontinuation.

Additionally, participant enrollment by country will be summarized by overall for all participants based on Screened set.

10.1.3 Time on Study

For IDMC only, time on study, as described in Section 9.11 will be summarized using descriptive statistics for continuous variables by treatment arm and overall based on the ITT set. In addition, Kaplan-Meier (KM) estimates (product-limit estimates) will be presented together with a summary of associated statistics as follows:

- Median time, Q1 and Q3 (whenever estimable) with their corresponding two-sided 95% Confidence Intervals (CIs) calculated according to Brookmeyer and Crowley (1982).
- The cumulative proportion of patients still in the study up to specific timepoints calculated using the reversed Kaplan-Meier methodology (Schemper & Smith 1996) where last assessment date is considered as an event and death is censored.

Graphical representation will be displayed with monotone increasing curves.

10.2 Protocol Deviations

Protocol deviations are categorized into:

- Non-important protocol deviations
- Important protocol deviations (IPDs)
- Clinically important protocol deviations (cIPDs)

IPDs are protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a participant's rights, safety, or well-being.

IPDs include:

- Participants enrolled and dosed on the study who did not satisfy enrolment criteria
- Participants who are not compliant with treatment: overdose, dose modifications not as per protocol, incorrect dose, etc.
- Participants who receive a prohibited concomitant medication
- Failure to collect data necessary to interpret primary endpoints
- Failure to collect necessary key safety data
- Deviation from Good Clinical Practice (GCP)
- Any other protocol deviation that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a participant's rights, safety, or well-being.

Important protocol deviations will be identified for all participants by either site monitoring, medical review processes or programming and confirmed prior to or at the Data Review Meeting at the latest.

IPDs are further classified into cIPDs which include, but are not limited to:

- Incorrect treatment group allocation, different to assignment at randomization.
- Incorrect stratification for randomization has been performed - incorrect information entered (1:1 randomized by: Region (North America vs Western Europe vs Rest of the world), Primary tumor site (larynx vs other), Lymph node involvement (N0-1 vs N2 vs N3), T size (T4 vs other).
- Histologically confirmed diagnosis in LA SCCHN patient (stage III, IVA or IVB according to the American Joint Committee on Cancer [AJCC]/TNM Staging System, 8th Ed.) suitable for definitive CRT, of at least one of the following sites: oropharynx, hypopharynx, and larynx is not confirmed.
- Stage IVC (metastatic disease) Exclusion criteria 2
- Impaired cardiovascular function or clinically significant diseases, including any of the following:
    - Ongoing or history of uncontrolled or symptomatic ischemic myocardiopathy within 6 months prior to randomization.
    - Known left ventricular ejection fraction < 50%, left ventricular hypertrophy, ventricular arrhythmias, bradycardia (heart rate < 50 bpm). •
    - History of myocardial infarction, or severe/unstable angina, within 6 months prior to randomization.
    - New York Heart Association grade ≥ 3 congestive heart failure.
    - Symptomatic pulmonary disease requiring continuous or intermittent oxygen supply. Exclusion criteria 17
    - Non-compensated or symptomatic liver cirrhosis (Child-Pugh score: B or C). Exclusion criteria 21
- History of another malignancy within the last 3 years prior to randomization, with the exception of completely resected non-melanoma cell skin cancer outside the head and neck area or completely resected stage I breast cancer, or completely resected in-situ nonmuscular invasive bladder, cervix and/or uterine carcinomas or T1a squamous cell carcinoma of the esophagus. Exclusion criteria 18
- Missed baseline CT scan or MRI. Inclusion criteria 5
- Prior treatment for LA SCCHN
- Xevinapant/matched placebo relative dose intensity below 67% compared to the prescribed dose not due to a medical or safety reason
- Did not receive 200 mg/m2 or more of the overall planned cisplatin dose not due to a medical or safety reason

- Did not receive 66.5 Gy or more of the planned dose of the IMRT not due to a medical or safety reason
- Use of prohibited concomitant medication
- Any criterion identified during the Blind Data Review Meeting prior to database lock which may impact the efficacy analysis.

All protocol deviations are documented in the Study Data Tabulation Model (SDTM) domain DV. All important protocol deviations (including cIPDs) are documented in the Analysis Data Model (ADaM) domain ADDV whether identified for all participants by either medical review processes or programming, and confirmed prior to or at the Blind Data Review Meeting at the latest.

The cIPDs as well as the IPDs will be summarized by treatment arm and overall, for all participants in the ITT.

11 Demographics and Other Baseline Characteristics

11.1 Demographics

Baseline demographics will be summarized by treatment arm and overall, for all participants in the ITT.

Baseline demographic data to be evaluated will include age at screening (years), age categories (18 to <55 years, 55 to <60 years, 60 to <65 years, 65 to <70 years, 70 to <75 years, >=75 years) sex, race, ethnicity, region, as well as the baseline values of height (cm), weight (kg), body mass index (BMI) [kg/m^2], and body surface area (BSA) [m^2]. In addition, use of tobacco (status, pack-years smoked, years smoked, < 10 pack years, >=10 pack years), alcohol (status, average of alcohol units per week, rehabilitation or treatment program needs), or any substance use rehabilitation or treatment program will be summarized.

Supporting participant data listing for demographic characteristics will be provided.

Specifications for computation:

- BSA [m^2] = 0.007184 * (Baseline weight (kg)^0.425)* (Baseline height (cm)^0.725)
- BMI [kg/m^2] = Baseline weight (kg) / [(Baseline height (cm)/100)^2]

11.2 Medical and Surgical History

Medical and surgical history will be coded using the last available version of Medical Dictionary for Regulatory Activities (MedDRA).

Medical and surgical history will be summarized by treatment arm and overall, for all participants in the ITT set.

Summary of medical and surgical history will include number and percentage of participants with a medical/surgical history condition overall, per System Organ Class (SOC) and per Preferred

Term (PT) within each SOC. A participant contributes only once to the count for a given medical/surgical history (SOC or PT). Medical and surgical history will be sorted by descending overall incidence in SOC and PT. In case of the same incidence, alphabetical order of SOC and PT will be applied.

Participant data listing will include condition/procedure, start and end dates, any treatment for condition and NCI-CTCAE grade (if ongoing).

11.3 Other Baseline Characteristics

Information on disease characteristics collected at baseline will be summarized overall and by treatment arm based on ITT. Summary statistics will be presented for:

- ECOG performance status: 0, 1, 2, 3 and 4.
- Feeding tube in place
- Primary tumor localization: larynx, hypopharynx and oropharynx (from eCRF), and larynx versus other (from Integrated Randomization Technology [IRT]).
- Time since initial diagnosis (months).
- AJCC cancer staging: III,IVa, IVb.
- Tumor size: T0, T1, T2, T3, T4a, and T4b (from eCRF), and T4 versus other (from IRT);
- Tumor grade: Gx, G1, G2, G3, G4, and Not Done
- Lymph node involvement: N0, N1, N2a, N2b, N2c and N3 (from eCRF), and N0-1 versus N2 versus N3 (from IRT);
- Metastasis: M0, M1.
- Number of target lesions: 1,2 3, …
- Number of nodal target lesions: 1, 2, 3, …
- Number of non-nodal target lesions: 1, 2, 3, …
- Number of non-target lesions: 1,2 3, …
- Sum of diameter of measurable disease at screening (mm).

Supporting participant data listing for baseline cancer characteristics will be provided.

Specifications for computation:

- Time since initial diagnosis (months) = (date of randomization - date of initial diagnosis) / 365.25 * 12.

12 Prior, Concomitant and Subsequent Medications/Procedures

Medications will be classified into 3 categories:

- Prior, defined as medications started before the first dose date regardless if continued at the time of starting study intervention.
- Concomitant, defined as medications taken any time during the on-treatment period.
- Subsequent, defined as medications started after the on-treatment period.

If the start and/or end date of a prior, concomitant, and subsequent medication is missing, it will be imputed as specified in Section 9.12.1.

The last available version of the World Health Organization Drug dictionary (WHO-DD) will be used for coding of medications.

Prior and concomitant medication as well as subsequent antineoplastic therapies will be summarized by treatment arm and overall, for all participants in the ITT set.

Prior and concomitant medication and subsequent antineoplastic therapies, each will be summarized by number and percentage of participants overall, per Anatomical Therapeutic Chemical (ATC) Classification level 2 and per preferred term within each ATC level 2. A participant contributes only once to the count for a given medication (ATC level 2 or preferred term). Medications will be sorted by descending overall incidence in ATC level 2 and preferred term. In case of equal incidence, alphabetical order of ATC level 2 or preferred term will be applied. In addition, subsequent antineoplastic therapies will be summarized by treatment arm and overall, according to types that are mapped by medical review.

In addition to medications, the surgical procedures summary will include the number and percentage of participants with a surgery, type of surgery, planned surgery, hospitalization, radical salvage surgery and pathological complete response together with the summary of anatomopathological findings (% viable cells) by treatment arm and overall, for the ITT set.

Prior and concomitant medication as well as subsequent antineoplastic therapies will be listed. Concomitant procedures, if any, will also be listed.

12.1 Concomitant 5-HT3 Receptor Antagonist

Concomitant 5-HT3 receptor antagonist medications (any medication coded with ATC Level 4 equal to “SEROTONIN (5HT3) ANTAGONISTS”) will be summarized overall and by treatment arm for all participants from SAF set who reported renal toxicity (grade 3 or 4) vs. participants who did not report renal toxicity. Renal toxicity grade 3 or 4 includes any participant with at least one post-baseline toxicity grade 3 or 4 for estimated Glomerular Filtration Rate (eGFR) or Creatinine. The following will be summarized:

- Number and percentage of patients with any 5-HT3 receptor antagonist medication.

- Number and percentage of patients by 5-HT3 receptor antagonist medication alone. PT will be used to summarize each medication.

- Number and percentage of patients with any combination of 5-HT3 receptor antagonist medications.

Supportive listings will be provided for patients with and without renal toxicity (grade 3 or 4) who took concomitant 5-HT3 receptor antagonist medication.

CCI

[redacted]

13 Study Intervention: Compliance and Exposure

The exposure to xevinapant/matched placebo, cisplatin/carboplatin, and IMRT will be summarized by cycle and overall, by actual treatment arm for all participants from SAF set. The following extent of exposure parameters will be summarized:

- Number of treatment cycles, defined as the number of cycles in the eCRF with non-zero dose administration of xevinapant/matched placebo, cisplatin/carboplatin. No cycle is defined for IMRT.

- Treatment duration (in weeks) is calculated per study intervention component as

    - Chemotherapy treatment duration = (last dose date - first dose date + 21)/7 weeks

        - Cisplatin treatment duration = (last dose date - first dose date + 21)/7 weeks

        - Carboplatin treatment duration = (last dose date - first dose date + 21)/7 weeks

    - IMRT treatment duration = (last dose date - first dose date + 3)/7 weeks

CCI

CCI

- Xevinapant / Placebo treatment duration = (last dose date - first dose date + 8)/7 weeks

Overall treatment duration is defined as the maximum of xevinapant / placebo treatment duration, chemotherapy treatment duration, and RT treatment duration.

- Total cumulative dose, per treatment (units' treatment specific), calculated as the sum of the dose in mg/m^2 for cisplatin and in Gy for IMRT administered during the treatment duration. The total cumulative dose for cisplatin will also be summarized into categories <100 mg/m^2, 100 to <200 mg/m^2, 200 to <300 mg/m^2, >=300 mg/m^2. The total cumulative dose for IMRT will also be categorized into <50, 50-<70, >=70 Gy.

- Dose intensity of xevinapant/matched placebo (mg/day) is defined as the total dose of xevinapant/matched placebo taken during a cycle, divided by the cycle duration in days. Overall dose intensity is calculated as the mean of the dose intensities of the individual cycles.

- Relative dose intensity of xevinapant/matched placebo, expressed in %, for any defined period is calculated as the dose intensity for that period divided by the planned dose intensity for that period * 100, where the planned dose intensity equals 2800/21 mg per day. Relative dose intensity will also be categorized into <80%, 80%-<100%, 100%. Additionally, the relative dose intensity will be computed during the combination period (first 3 treatment cycles only) and separately during the monotherapy period (cycles 4-6) and will be categorized into <80, 80%- <100%, 100%.

- Dose intensity of cisplatin in mg/m^2/week is defined as the total dose of cisplatin (mg/m^2) taken during a cycle, divided by the cycle duration in days and multiplied by 7. Overall dose intensity is derived as the mean of the dose intensities of the individual cycles.

- Dose intensity of carboplatin in mg·min/mL/week is defined as the total dose of carboplatin (mg·min/mL) taken during a cycle, divided by the cycle duration in days and multiplied by 7. Overall dose intensity is derived as the mean of the dose intensities of the individual cycles.

- Relative dose intensity of cisplatin/carboplatin, expressed in %, is calculated as the dose (mg/m^2 or mg·min/mL, respectively) received divided by the planned dose *100, where the planned dose equals 100 mg/m^2 per cycle for cisplatin and 5 mg·min/mL per cycle for carboplatin.

- Time to first dose reduction (days), defined as the date of the first dose to the date of the first dose reduction, or the date of the first dose interruption if participants had a dose interruption followed by a dose reduction.

- Number of participants receiving cisplatin or carboplatin at a specific cycle.

- For participants receiving cisplatin, the number and percentage of participants with pre-infusion hydration and post-infusion administered at a specific cycle.

In addition, for each cycle the number and percentage of participants who had missed, interrupted, reduced, or discontinued doses, together with the reason(s) will be summarized for each treatment component (xevinapant/matched placebo, and cisplatin/carboplatin). Additionally, the switch to carboplatin per participant will be presented for each cycle. For each modification action, each distinct reason will be reported in each specific category.

Supportive participant data listings will be presented.

14 Efficacy Analyses

The efficacy analyses will be performed based on the ITT set by treatment arm, except when otherwise stated.

14.1 Primary Efficacy Endpoint: Event-free Survival (EFS)

Event-Free Survival (EFS) time is defined as the time (in months) from randomization to the first occurrence of any of the following events as assessed by the BIRC: death due to any cause, disease progression (radiological or clinical), primary treatment failure, relapse (radiological or clinical) or occurrence of a new cancer (see Table 1). Death as reported on the eCRF page "Death" will also be considered in the derivation of EFS time.

EFS = (date of EFS event – randomization date +1) / 30.4375

Censoring rules for the primary and sensitivity analyses of EFS are described in Table 4 below. The last adequate tumor assessment is defined as the last tumor assessment result that is not "Not evaluable" or "Not assessed". If a participant is meeting a multiple censoring situation, the earliest censoring criterion will be applied.

Table 4: Censoring rules for EFS

Situation	Primary	Sensitivity #1	Sensitivity #2
No evaluable baseline or post-baseline assessment*	Right censored at the date of randomization	Right censored at the date of randomization	Right censored at the date of randomization
Death, PD or any event entering in the definition of EFS after 2 or more consecutive missing or inadequate tumor assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Not applicable
Initiation of a new anticancer treatment for the disease under investigation before any event entering in the definition of EFS	Right censored at their last adequate assessment before the initiation of the new anticancer treatment	Not applicable	Not applicable
No death, PD, nor any event entering in the definition of EFS	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment

Abbreviations: **EFS**: Event-free survival; **PD**: Progressive disease
*Any EFS event is considered as a post-baseline assessment.

Table 5: Estimand Attributes EFS

Objective	Estimand Attributes
Primary	

To demonstrate improvement in EFS with xevinapant compared to placebo when added to chemoradiotherapy	**Endpoint:** EFS as assessed by the BIRC defined as the time from randomization to the first occurrence of any of the following events (occurring within two scheduled tumor assessments after last evaluable assessment or randomization):
	• Death from any cause
	• Progression:
	  • Radiological, assessed per RECIST v1.1 or
	  • Clinical, with or without RECIST v1.1- radiologically documented progression, assessed endoscopically
	• Primary treatment failure before achieving a CR: requirement for radical salvage surgery that includes the primary tumor site, with documented viable tumor presenting anatomopathological findings even in the absence of formal RECIST v1.1 radiological progression.
	• Any radiological or clinical relapse after achieving a CR (loco-regionally), including any event defined as locoregional treatment failure, even in the absence of formal radiological progressive disease confirmation:
	  • Requirement for radical salvage surgery that includes the primary tumor site, regardless of anatomopathological findings
	  or
	  • Elective neck dissection or biopsy, with positive viable tumor cells on anatomopathological findings at 22 weeks or later after randomization.
	• Second cancers unless anatomopathological findings exclude squamous histology
	
	Population: Patients with LA-SCCHN Stage III, IVA or IVB suitable for definitive CRT
	Treatment: xevinapant and CRT followed by xevinapant vs placebo and CRT followed by placebo
	Intercurrent Event Strategy:
	• Death within two missing scheduled tumor assessments after last evaluable assessment or randomization will be considered as event (composite strategy)
	• Discontinuation of treatment: The endpoint will be analyzed regardless of whether or not treatment has been discontinued (treatment-policy strategy)
	• Start of subsequent anticancer therapy will be analyzed according to hypothetical strategy, i.e., tumor assessments after start of subsequent anticancer therapy will not be considered for analysis
	Population-Level Summary: Hazard Ratio

Objective	Estimand Attributes
Supplementary #1 (CSP Sensitivity #1)	**Endpoint:** EFS as assessed by the BIRC defined as the time from randomization to the first occurrence of any of the events (occurring within two scheduled tumor assessments after last evaluable assessment or randomization). **Population:** Patients with LA-SCCHN Stage III, IVA or IVB suitable for definitive CRT **Treatment:** xevinapant and CRT followed by xevinapant vs placebo and CRT followed by placebo **Intercurrent Event Strategy:** • Death within two missing scheduled tumor assessments after last evaluable assessment or randomization will be considered as event (composite strategy) • Discontinuation of treatment: The endpoint will be analyzed regardless of whether or not treatment has been discontinued (treatment-policy strategy) • Start of subsequent anticancer therapy: The endpoint will be analyzed regardless of the start of subsequent anticancer therapy **Population-Level Summary:** Hazard Ratio
Supplementary #2 (CSP Sensitivity #2)	**Endpoint:** EFS as assessed by the BIRC defined as the time from randomization to the first occurrence of any of the events. **Population:** Patients with LA-SCCHN Stage III, IVA or IVB suitable for definitive CRT **Treatment:** xevinapant and CRT followed by xevinapant vs placebo and CRT followed by placebo **Intercurrent Event Strategy:** • Discontinuation of treatment: The endpoint will be analyzed regardless of whether or not treatment has been discontinued (treatment-policy strategy) • Start of subsequent anticancer therapy: The endpoint will be analyzed regardless of the start of subsequent anticancer therapy. **Population-Level Summary:** Hazard Ratio

Objective	Estimand Attributes
Sensitivity #1 (CSP Sensitivity #3)	**Endpoint:** EFS as assessed by the Investigator defined as the time from randomization to the first occurrence of any of the events (occurring within two scheduled tumor assessments after last evaluable assessment or randomization): **Population:** Patients with LA-SCCHN Stage III, IVA or IVB suitable for definitive CRT **Treatment:** xevinapant and CRT followed by xevinapant vs placebo and CRT followed by placebo **Intercurrent Event Strategy**: • Death within two missing scheduled tumor assessments after last evaluable assessment or randomization will be considered as event (composite strategy) • Discontinuation of treatment: The endpoint will be analyzed regardless of whether or not treatment has been discontinued (treatment-policy strategy) • Start of subsequent anticancer therapy will be analyzed according to hypothetical strategy, i.e., tumor assessments after start of subsequent anticancer therapy will not be considered for analysis **Population-Level Summary:** Hazard Ratio

14.1.1 Primary Analysis of the Primary Efficacy Endpoint EFS

The primary efficacy analysis will assess the difference in EFS between treatment arms using a one-sided stratified log-rank test. The significance boundaries are adjusted accordingly to control the family-wise error rate (FWER) at 2.5% one-sided (see also IAP Table 3). The stratification factors are tumor size (T4 vs other) and lymph node involvement (N0-1 vs N2 vs N3). These factors were chosen among the stratification factors for randomization to avoid issues caused by very small sample stratum sizes.

A Cox-proportional hazard regression model adjusted for the randomization factors will be used on the EFS with primary censoring rule (Table 4) to derive the hazard ratio and 95% CI between the 2 randomized regimens. In addition, a two-sided repeated CI (RCI) for the hazard ratio will be reported (Jennison and Turnbull, 2000), to account for the group sequential design in the confirmatory analyses (interim and primary analysis, as applicable). Ties will be handled by replacing the proportional hazards model by the discrete logistic model. The stratification factors used for the randomization of participants will be inserted as covariates in the regression model. EFS with sensitivity censoring rules will be analyzed with the same approach.

A further estimate of the hazard ratio (including 95% confidence interval) will be calculated based on the Cox proportional hazards model stratified by the stratification factors tumor size (T4 vs other) and lymph node involvement (N0-1 vs N2 vs N3). Each stratum will define separate

baseline hazard functions. Ties will be handled by replacing the proportional hazards model by the discrete logistic model.

The null and alternative hypotheses can be described in terms of hazard ratios as:

$$H_0^{EFS}: \lambda_{(xevinapant + CRT)} / \lambda_{(placebo + CRT)} \geq 1 \text{ versus}$$

$$H_1^{EFS}: \lambda_{(xevinapant + CRT)} / \lambda_{(placebo + CRT)} < 1,$$

where $\lambda_{(xevinapant + CRT)}$, $\lambda_{(placebo + CRT)}$ represents the hazard of EFS for the (xevinapant + CRT) and the (placebo + CRT) group respectively.

The SAS code for the stratified log-rank test and the Cox-proportional hazard regression model is specified in Appendix 6

The proportional hazards assumption for the analysis of EFS will be examined using graphical methods by means of log-log-Survival plots (log(-log(S(t))) vs log(time) by treatment arm, plots=(lls) in PROC LIFETEST.)

Kaplan-Meier methodology will be used to estimate EFS as assessed by the BIRC (with primary censoring rule as defined in Table 4) in each treatment arm together with a summary of associated statistics: median time with corresponding two-sided 95% CIs, Q1 and Q3, min and max. Event-free rates at 9, 12, 18, 24, 30 and 36 months after randomization will be estimated with corresponding two-sided 95% CIs. Other rates may be displayed. For each interval the number of participants at risk at the start of the interval as well as the number of participants failed during the interval will be reported.

Kaplan-Meier estimates and median time are calculated with the PROC LIFETEST procedure in SAS. CIs for the median and the estimated survival rates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (2002) [CONFTYPE=LOGLOG option (default)] based on standard errors computed using Greenwood's formula.

The SAS code for the Kaplan-Meier estimates is specified in Appendix 6

Graphical representation of the survival curves as step functions will be provided per treatment arm and per stratification variable.

The number and percentage of participants with each event type (death due to any cause, disease progression (radiological or clinical), primary treatment failure, relapse (radiological or clinical) or occurrence of a new cancer) as assessed by the BIRC and censoring reasons will be presented by treatment arm for all 3 censoring scenarios as presented in Table 4.

The EFS time or censoring time and the reasons for censoring will also be listed for all 3 censoring scenarios as presented in Table 4.

A summary of concordance of EFS between BIRC and investigator assessment will be provided including status of "No Event" and which type of EFS event as well as number of cases where EFS was assessed at different timepoints by BIRC and investigator (shift tables per treatment arm).

Follow-up for EFS

In order to assess the follow-up for EFS, Kaplan-Meier estimates will be calculated using the censoring rules with reverse censoring indicator.

14.1.2 Sensitivity Analyses of the Primary Efficacy Endpoint EFS

To assess the robustness of the primary efficacy endpoint the following sensitivity analyses will be performed:

Table 6: Sensitivity analyses for EFS

Sensitivity analysis	Population	Censoring rules (Table 4)	Endpoint	Statistical methods
1	ITT	Primary	EFS as assessed by Investigator	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves

Frequency (number and percentage) of participants with each event type (death due to any cause, disease progression (radiological or clinical), primary treatment failure, relapse (radiological or clinical) or occurrence of a new cancer) as assessed by the investigator and censoring reasons will be presented by treatment arm.

The EFS time or censoring time and the reasons for censoring will also be listed.

A further sensitivity analysis to address possible biases due to the dynamic allocation process will be performed based on permutation testing for the primary endpoint EFS by BIRC. To this end, the stratified log-rank test as described in 14.1.1 for the primary endpoint EFS by BIRC will be repeated with rerandomized xevinapant and placebo arms. The rerandomized treatment arms are determined by reapplying the dynamic allocation approach using the actual randomization dates and randomization factors for all participants. The permutation test will be conducted with 10,000 replications.

14.1.3 Supplementary Analyses of the Primary Efficacy Endpoint EFS

The following supplementary analyses will be performed.

Table 7: Supplementary analysis for EFS

Supplementary analysis	Population	Censoring rules (Table 4)	Endpoint	Statistical methods
1	ITT	Sensitivity #1	EFS as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
2	ITT	Sensitivity #2	EFS as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves

Multivariable Cox Regression

Multivariable Cox regression analysis will be conducted for the primary endpoint EFS as assessed by the BIRC to assess and adjust the treatment effect for relevant baseline factors of potential prognostic impact. The subgroup variables defined in Sections 8.2.1 and 8.2.2 will be included in the model. A stepwise selection procedure will serve to identify explanatory variables of potential prognostic values additional to the randomization strata, which will be included in all models during the selection procedure. The stepwise selection method in SAS (PROC PHREG) will be used. The level of significance for an explanatory variable to enter the model is set to 0.15 (boundary for p-value of Score test for inclusion) and the significance level for removing it is set to 0.40 (boundary for p-value of Wald test for exclusion). Once the selection procedure is finalized, the treatment group will be added to the effect of treatment on EFS when adjusted for the selected explanatory variables.

The Cox's Proportional Hazard model is defined as: $h(t) = h(0;t) \exp(Xb)$, where $h(0;t)$ defines the Baseline Hazard function, X defines the vector of explanatory variables and b the unknown vector of regression parameters. The HR of all selected explanatory variables and of treatment effects will be reported including 2-sided 95% CIs. No interaction will be considered. Post-Baseline factors will not be considered for the model.

14.1.4 Subgroup Analyses of the Primary Efficacy Endpoint EFS

Subgroup analyses for the EFS as assessed by the BIRC and EFS as assessed by investigator (with primary censoring rule as defined in Table 4) will be performed by treatment arm for all subgroup levels defined in Sections 8.2.1 and 8.2.2. In the case of a low number of participants within a category (<5% of the randomized population), the categories may be pooled. All subgroup analyses are exploratory and will be performed unstratified. No adjustment for multiplicity will be performed.

Cox regression models will be fitted for the EFS time as dependent variable and with subgroup type, the treatment arm, and with and without the treatment by subgroup type interaction as explanatory variables. The unstratified HR and its corresponding 95% CI will be computed per subgroup level. A p-value for the interaction test (Likelihood Ratio test) comparing models with

and without the interaction term will be provided together with the hazard ratios and corresponding 95% CI of the interaction model parameter.

The HR and its corresponding 95% CI, overall and for each subgroup level will be presented in a forest plot.

14.2 Secondary Efficacy Endpoints and Analyses

This section describes the statistical methods for the analysis of the secondary efficacy endpoints. All secondary endpoint analysis will be performed on ITT. Confirmatory testing will only be done for OS. For all other secondary endpoints, no formal statistical hypotheses will be tested.

14.2.1 Overall Survival

Overall Survival (OS) time is defined as the time (in months) from randomization to death due to any cause.

$$OS = (\text{date of death} - \text{randomization date} + 1) / 30.4375$$

Participants without documented death at the time of analysis will be censored at the date of last known contact.

The treatment difference in OS between the 2 treatment arms will be assessed using a one-sided stratified log-rank test as described for the primary analysis of EFS in Section 14.1.1. The significance boundaries are adjusted accordingly to control the family-wise error rate (FWER) at 2.5% one-sided (see also Table 3). The stratification factors are tumor size (T4 vs other) and lymph node involvement (N0-1 vs N2 vs N3).

The null and alternative hypotheses can be described in terms of hazard ratios as:

$$H_0^{OS}: \lambda_{(\text{xevinapant + CRT})} / \lambda_{(\text{placebo + CRT})} \geq 1 \text{ versus}$$

$$H_1^{OS}: \lambda_{(\text{xevinapant + CRT})} / \lambda_{(\text{placebo + CRT})} < 1,$$

where $\lambda_{(\text{xevinapant + CRT})}$, $\lambda_{(\text{placebo + CRT})}$ represents the hazard of OS for the (xevinapant + CRT) and the (placebo + CRT) group respectively.

The treatment effect will be analyzed using a Cox Proportional Hazard model as described for the primary analysis of EFS in Section 14.1.1 including the RCI for the hazard ratio.

A further sensitivity analysis to address possible biases due to the dynamic allocation process will be performed based on permutation testing for the secondary endpoint OS as described for the primary endpoint EFS by BIRC in Section 14.1.2.

The Kaplan-Meier method (as described for the primary analysis of EFS in Section 14.1.1) will be used to estimate the OS in each treatment arm. Survival rates at 12, 24, 36, 48, and 60 months after randomization will be estimated with corresponding two-sided 95% CIs.

Graphical representation of the survival curves (Kaplan-Meier estimates) as step functions will be provided per treatment arm and per stratification variable.

The number and percentage of deaths will be presented by treatment arm. The OS time or censoring time will also be listed.

The same subgroup analyses as described for the EFS in Section 14.1.4 will be performed for OS.

Table 8: **Estimand Attributes OS**

Objective	Estimand Attributes
Secondary	
To demonstrate improvement in Overall Survival (OS) with xevinapant compared to placebo when added to chemoradiotherapy	**Endpoint:** OS defined as the time from date of randomization to death **Population:** Patients with LA-SCCHN Stage III, IVA or IVB suitable for definitive CRT **Treatment**: Xevinapant and CRT followed by xevinapant followed by subsequent cancer therapy vs. placebo and CRT followed by placebo followed by subsequent cancer therapy **Intercurrent Event Strategy:** The endpoint will be analyzed regardless of whether or not the following intercurrent events had occurred (treatment-policy strategy): • Treatment discontinuation • Start of subsequent anticancer therapy **Population-level Summary:** Hazard ratio

Follow-up for OS

In order to assess the follow-up for OS, Kaplan-Meier estimates will be calculated using the censoring rules with reverse censoring indicator.

14.2.2 Progression-Free Survival

Progression-Free Survival (PFS) time according to RECIST v1.1 is defined as the time (in months) from randomization to the first occurrence of disease progression (radiological or clinical, as assessed by the BIRC or the investigator, respectively) or death due to any cause.

PFS = (date of PD or death – randomization date +1) / 30.4375

Censoring rules for the primary and sensitivity analyses of PFS are described in Table 9 below. The last adequate tumor assessment is defined as the last tumor assessment result that is not "Not evaluable" or "Not assessed". If a participant is meeting a multiple censoring situation, the earliest censoring criterion will be applied.

Table 9: Censoring rules for PFS

Situation	Primary	Sensitivity #1	Sensitivity #2
No evaluable baseline or post-baseline assessment*	Right censored at the date of randomization	Right censored at the date of randomization	Right censored at the date of randomization
Death, PD or any event entering in the definition of PFS after 2 or more consecutive missing or inadequate tumor assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Not applicable
Initiation of a new anticancer treatment for the disease under investigation before any event entering in the definition of PFS	Right censored at their last adequate assessment before the initiation of the new anticancer treatment	Not applicable	Not applicable
No death, PD, nor any event entering in the definition of PFS	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment

Abbreviations: PD: Progressive disease; **PFS**: Progression-free survival.
*Any PFS event is considered as a post-baseline assessment.

The secondary endpoint analysis of PFS will consist of the analyses as presented in Table 10 below and will follow the same methods as described for EFS. The PFS rates will be estimated be at 9, 12, 18, 24, 36, 48 and 60 months post-randomization.

Table 10: Analyses for PFS

Analysis	Population	Censoring rules (Table 7)	Endpoint	Statistical methods
Main	ITT	Primary	PFS as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Sensitivity	ITT	Primary	PFS as assessed by the investigator	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves

Analysis	Population	Censoring rules (Table 7)	Endpoint	Statistical methods
Supplementary	ITT	Sensitivity #1	PFS as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Supplementary	ITT	Sensitivity #2	PFS as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Subgroups	ITT	Primary	PFS as assessed by the BIRC	same as EFS subgroup analysis (Section 14.1.4) only for subgroups defined in Section 8.2.1, and to be presented in a forest plot

The number and percentage of deaths due to any cause and disease progression (radiological or clinical) as assessed by the BIRC or the investigator respectively and censoring reasons will be presented by treatment arm for all 3 censoring scenarios as presented in Table 9.

The PFS time or censoring time and the reasons for censoring will also be listed for all 3 censoring scenarios as presented in Table 9.

14.2.3 Locoregional Control

Locoregional Control (LRC) time is defined as the time (in months) from randomization to the first occurrence of progression at the site of the primary tumor or the locoregional lymph nodes, either according to RECIST v1.1 or based on clinical assessment (radiological or clinical, as assessed by the investigator or the BIRC, respectively).

LRC = (date of PD (primary tumor or locoregional) – randomization date +1) / 30.4375

Censoring rules for the primary and sensitivity analyses of LRC are described in Table 11 below. The last adequate tumor assessment is defined as the last tumor assessment result that is not “Not evaluable” or “Not assessed”. If a participant is meeting a multiple censoring situation, the earliest censoring criterion will be applied.

Table 11: Censoring rules for LRC

Situation	Primary	Sensitivity #1	Sensitivity #2
No evaluable baseline or post-baseline assessment	Right censored at the date of randomization	Right censored at the date of randomization	Right censored at the date of randomization
PD after 2 or more consecutive missing or inadequate tumor assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Right censored at last adequate assessment before the consecutive missing/inadequate assessments	Not applicable
Initiation of a new anticancer treatment for the disease under investigation before any documented progressive disease	Right censored at their last adequate assessment before the initiation of the new anticancer treatment	Not applicable	Not applicable
No locoregional failure	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment	Right censored at the date of their last adequate assessment

Abbreviations: LRC: Locoregional control; **PD:** Progressive disease

The secondary endpoint analysis of LRC will consist of the analyses as presented in Table 12 below and will follow the same methods as described for EFS. The LRC rates will be estimated be at 9, 12, 18, 24, 36, 48 and 60 months post-randomization.

Table 12: Analyses for LRC

Analysis	Population	Censoring rules(Table 9)	Endpoint	Statistical methods
Main	ITT	Primary	LRC as assessed by the investigator	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Sensitivity	ITT	Primary	LRC as assessed by the BIRC	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves

Analysis	Population	Censoring rules(Table 9)	Endpoint	Statistical methods
Supplementary	ITT	Sensitivity #1	LRC as assessed by the investigator	same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Supplementary	ITT	Sensitivity #2	LRC as assessed by the investigator	Same as primary EFS analysis (Section 14.1.1), incl. Kaplan-Meier curves
Subgroups	ITT	Primary	LRC as assessed by the investigator	same as EFS subgroup analysis (Section 14.1.4) only for subgroups defined in Section 8.2.1, and to be presented in a forest plot

The number and percentage of disease progression (primary tumor or locoregional) as assessed by the BIRC or the investigator, respectively and censoring reasons will be presented by treatment arm for all 3 censoring scenarios as presented in Table 11.

The LRC time or censoring time and the reasons for censoring will also be listed for all 3 censoring scenarios as presented in Table 11.

14.2.4 Objective Response

Objective Response (OR) is defined as a best overall response (see Section 9.15 for the derivation) of CR or PR according to RECIST v1.1, as assessed by the BIRC or the investigator, respectively.

Each participant will have an objective response status (0: 'no OR'; 1: 'OR') at 9 and 12-months post-randomization.

Participants will be counted as non-responders in the assessment of OR, who

- do not have any evaluable post-baseline assessment,
- receive a new anticancer treatment prior to reaching CR or PR,
- only have non-measurable disease with a best overall response at non-CR/non-PD,
- progress, die, or drop out for any reason prior to reaching CR or PR.

The OR rate (ORR) is the proportion of participants with OR in the analysis set.

OR status as assessed by BIRC will be compared between the two treatment arms by using a logistic regression model adjusted for randomization stratification factors. The same stratification factors used for the randomization will be applied to the analysis. The number of participants with

OR will be summarized by treatment arm and presented along with the odds ratio and the corresponding 95% CI.

The SAS code for the logistic regression model is specified in Appendix 6

A sensitivity analysis will be performed based on the OR as assessed by the investigator.

The number and percentage of participants with a best overall response of CR, PR, SD, non-CR/non-PD, PD, and NE as assessed by the BIRC or the investigator, respectively will be summarized by treatment arm.

The best overall response of CR, PR, SD, non-CR/non-PD, PD, and NE as assessed by the BIRC or the investigator, respectively, will also be listed together with the OR status.

14.2.4.1 Complete Response

Complete Response (CR) is defined as a best overall response (see Section 9.15 for the derivation) of CR according to RECIST v1.1, as assessed by the BIRC or the investigator, respectively.

Each participant will have a complete response status (0: 'no CR'; 1: 'CR') at 9 and 12-months post-randomization.

Participants will be counted as non-responders in the assessment of CR, who

- do not have any evaluable post-baseline assessment,
- receive a new anticancer treatment prior to reaching CR,
- only have non-measurable disease with a best overall response at non-CR/non-PD,
- progress, die or drop out for any reason prior to reaching CR.

The CR rate (CRR) is the proportion of participants with CR in the analysis set.

The same analyses and summary statistics as described for the OR in Section 14.2.4 will be performed. Additional subgroup analyses will be done by treatment arm for all subgroup levels defined in Section 8.2.1.

Time to CR as assessed by BIRC or investigator is calculated for all patients who have a CR by BIRC or investigator, respectively, at any point during the study.

Time to CR is calculated for each responder as the number of days from the Date of Randomization until the Date of First Response. No censoring will be applied for this endpoint: patients who do not respond at any time during the study will be excluded from the analysis.

Summary statistics and an empirical cumulative distribution function plot for time to CR will be provided by treatment arm.

14.2.5 Duration of Response

DoR is defined as the time (in months) from the first evidence of response (partial or complete, as assessed by the BIRC or the investigator, respectively according to RECIST v1.1) to the first occurrence of progression (radiological or clinical, as assessed by the BIRC or the investigator, respectively) or death due to any cause. DoR will only be computed for participants with objective response (as defined in Section 14.2.4).

DoR = (date of PD or death – first CR/PR date +1) / 30.4375

Censoring rules for the analysis of DoR will follow the same as those applied to PFS described in Table 9. The DoR rates will be estimated at 6, 12 and 24 months post first evidence of response. DoR will be analyzed by treatment arms but without any between-arm comparisons.

Table 13: Analyses for DoR

Analysis	Population	Censoring rules (Table 7)	Endpoint	Statistical methods
Main	ITT	Primary	DoR as assessed by the BIRC	KM estimates (as described in Section 14.1.1), incl. Kaplan-Meier curves
Sensitivity	ITT	Primary	DoR as assessed by the investigator	KM estimates (as described in Section 14.1.1), incl. Kaplan-Meier curves
Subgroups	ITT	Primary	DoR as assessed by the BIRC	same as EFS subgroup analysis (Section 14.1.4) only for subgroups defined in Section 8.2.1, and to be presented in a forest plot

The number and percentage of CR, PR, death due to any cause and disease progression (radiological or clinical) as assessed by the BIRC or the investigator respectively and censoring reasons will be presented by treatment arm.

The DoR time or censoring time and the reasons for censoring will be listed.

14.2.6 Proportion of Participants with Radical Salvage Surgery

The proportion of radical salvage surgery is defined as the proportion of participants with radical salvage surgery, excluding elective neck dissection (END) without anatomopathological (A-P) evidence of residual malignant cells.

Participants experiencing END without A-P evidence of residual malignant cells, will be considered as **not** having experienced a radical savage surgery for the computation of the proportion of participants with radical salvage surgery.

Summary statistics will be provided for the number and proportion of participants with radical salvage surgery experienced before 9, 12, 24, 36, 48, and 60-months post-randomization. In addition, the proportions will be compared between the two treatment arms by using a stratified logistic regression model as described for the OR in Section 14.2.4. The number of participants with and without radical salvage surgery will be summarized by treatment arm and presented along with the odds ratio and the corresponding 95% CI.

If the number of events permits, time to radical salvage surgery will be analyzed using the same methods as described for the primary analysis of EFS in Section 14.1.1.

14.2.7 Time to New Subsequent Systemic Cancer Treatment

The time to new subsequent systemic cancer treatment (NAT) (in months) is defined as the time from randomization to first subsequent systemic cancer treatment for this disease (LA-SCCHN).

NAT = (date of new subsequent anticancer treatment – randomization date +1) / 30.4375

Participants without documented new subsequent systemic cancer treatment for this disease will be censored at the date of last known contact with evidence of no new systemic cancer treatment.

The secondary endpoint analysis of time to new subsequent systemic cancer treatment will follow the same methods as described for the primary analysis of EFS in Section 14.1.1. Rates will be estimated at 12, 18, 24, 36, 48, and 60 months post-randomization.

CCI [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]

[redacted]

CCI [redacted]

15 Safety Analyses

All safety analyses will be conducted on the SAF set and data will be presented by actual treatment arm (Xevinapant + CRT vs Placebo + CRT), unless otherwise specified.

15.1 Adverse Events

Please refer to Section 9.1.1 of the protocol for the definition of an AE. All AEs (including Adverse Event of Special Interest (AESIs) during the treatment period) will be collected from the time of ICF signature until the EOT visit. From the EOT visit until the EOS visit, only Serious Adverse Event (SAEs) and late-onset AESIs will be collected. AESI definitions used in corresponding analyses are provided in Section 15.2.3.

A treatment-emergent adverse event (TEAE) is any new sign or symptom, disease, other untoward medical event, or change of an existing condition in a participant that begins during the on-treatment period whether or not considered to be drug-related.

Late-onset adverse events are defined as adverse events with first onset of a not-ongoing AE episode after the on-treatment period, i.e., events that start earliest the day after on-treatment period without an incidence of an AE with the same preferred term in the same subject prior or during on-treatment period whether or not considered to be drug-related.

All AEs recorded during the course of the study will be coded with the last available version of MedDRA and will be graded according to NCI-CTCAE version 5.0, as described in Section 9.16 of the protocol.

A treatment-related AE (TRAE) is a TEAE where the Investigator determined it as having a "reasonable causal relationship" to study intervention (done for each study intervention component (xevinapant/matched placebo, CT or IMRT)).

A dose-modifying TEAE is any event which leads to a dose reduction of any of the study intervention components (xevinapant/matched placebo, CT or IMRT).

Missing AE start dates will be imputed as described in Section 9.12.1. Adverse events for which the start date could not be determined after imputation of missing date are considered to be treatment emergent.

For summaries by relationship, AEs with missing relationship will be counted as "reasonable causal relationship". Adverse events with missing NCI-CTCAE grade will not have NCI-CTCAE grade imputed.

All analyses described will be based on TEAEs if not otherwise specified. The AE listings will include all AEs (whether treatment-emergent or not). Non-TEAEs will be flagged in the listings.

All summaries by SOC and PT will display the number and percentage of participants reporting at least one event together with the corresponding number of events (when applicable) and will be grouped by treatment arm. A participant contributes only once to the count for a given TEAE (SOC or PT). If summarized by worst/maximum NCI-CTCAE grade, a participant contributes only once

for a given TEAE (SOC or PT) at the worst/maximum NCI-CTCAE grade. No number of events will be displayed for tables showing the worst/maximum NCI-CTCAE grades. Tables will be sorted by descending overall incidence in SOC and PT. In case of the same incidence, alphabetical order of SOC and PT will be applied.

All summaries by PT will display the number and percentage of participants reporting at least one event together with the corresponding number of events and will be grouped by treatment arm. A participant contributes only once to the count for a given TEAE.

3-tier Approach to Summarizing and Analyzing AEs

The 3-tier approach for summarizing and analyzing AEs in clinical studies will be followed. AEs in different tiers are analyzed using different types of statistical analyses.

Pre-specified AESIs and events of clinical importance (AECI) during the on-treatment period (see Table 20 in Appendix 5) will be analyzed as part of Tier 1 if they fulfill the Rule-of-4 (≥ 4 participants in any of the treatment arms). Otherwise, such AEs will be included in Tier 3. All of the pre-specified events (AESI or AECI as part of Tier 1/Tier 3) will be included in one table. AEs common enough per Rule-of-4 are presented as part of Tier 1, whereas the 95% CI will be omitted for those belonging to Tier 3.

All TEAEs will be classified as belonging to Tier 2 (≥ 4 participants in any of the treatment arms) or Tier 3 (< 4 participants in any of the treatment arms). Analyses methods are provided based on the Rule-of-4.

The Tier 1 and Tier 2 TEAEs will be assessed with a 95% CI for between-group comparisons. For the difference in incidence proportion (risk difference in percentage points), the CIs will be based on Miettinen & Nurminen method (Miettinen & Nurminen (1985)). The Tier 3 AEs will be assessed via summary statistics and risk differences without any CIs.

While analyses will be done for all Tier 2 TEAEs, only TEAEs with incidence proportion ≥ 10% on at least 1 arm will be presented graphically.

No multiplicity adjustment will be applied for Tier 1 and 2 TEAEs.

15.1.1 All Adverse Events

An overall summary of AEs by treatment arm, will show the number and percentage of participants (and the corresponding number of events, when applicable) who report:

- any TEAE
- any NCI-CTCAE grade ≥ 3 TEAE
- any NCI-CTCAE grade ≥ 4 TEAE
- *any TEAE with maximum NCI-CTCAE grade 3 (IDMC only)*
- *any TEAE with maximum NCI-CTCAE grade 4 (IDMC only)*
- any treatment-related TEAE (any study treatment and by study treatment components)

- any treatment-related NCI-CTCAE grade ≥ 3 TEAE (any study treatment and by study treatment components)
- any treatment-related NCI-CTCAE grade ≥ 4 TEAE (any study treatment and by study treatment components)
- *any treatment-related TEAE with maximum NCI-CTCAE grade 3 (any study treatment and by study treatment components) (IDMC only)*
- *any treatment-related TEAE with maximum NCI-CTCAE grade 4 (any study treatment and by study treatment components) (IDMC only)*
- any treatment-emergent AESI
- any late-onset AESI
- any NCI-CTCAE grade ≥ 3 late-onset AESI
- any NCI-CTCAE grade ≥ 4 late-onset AESI
- *any late-onset AESI with maximum NCI-CTCAE grade 3 (IDMC only)*
- *any late-onset AESI with maximum NCI-CTCAE grade 4 (IDMC only)*
- any treatment-related late-onset AESI (any study treatment and by study treatment components)
- any treatment-related NCI-CTCAE grade ≥ 3 late-onset AESI (any study treatment and by study treatment components)
- any treatment-related NCI-CTCAE grade ≥ 4 late-onset AESI (any study treatment and by study treatment components)
- *any treatment-related late-onset AESI with maximum NCI-CTCAE grade 3 (any study treatment and by study treatment components) (IDMC only)*
- *any treatment-related late-onset AESI with maximum NCI-CTCAE grade 4 (any study treatment and by study treatment components) (IDMC only)*
- any serious TEAE
- any treatment-related serious TEAE (any study treatment and by study treatment components)
- any TEAE resulting in dose reduction (any study treatment and by study treatment components)
- any TEAE resulting in temporarily treatment interruption (any study treatment and by study treatment components)
- any TEAE resulting in permanently treatment discontinuation (any study treatment and by study treatment components)
- any TEAE leading to death
- any treatment-related TEAE leading to death (any study treatment and by study treatment components)
- any late-onset AEs (starting more than 30 days after the last dose date).

In addition, the following tables summarizing the frequency of participants with TEAEs by SOC and PT will be produced:

- All TEAEs;
- All TEAEs by worst NCI-CTCAE grade (Any, ≥3, ≥4 and 5);
- *All TEAEs with maximum NCI-CTCAE grade 3 (IDMC only);*
- *All TEAEs with maximum NCI-CTCAE grade 4 (IDMC only);*

- All treatment-related TEAEs (any study treatment and by study treatment components);
- All treatment-related TEAEs by worst NCI-CTCAE grade (Any, ≥3, ≥4 and 5)(any study treatment and by study treatment components);
- *All treatment-related TEAEs with maximum NCI-CTCAE grade 3 any study treatment and by study treatment components) (IDMC only);*
- *All treatment-related TEAEs with maximum NCI-CTCAE grade 4 (any study treatment and by study treatment components) (IDMC only);*
- All non-serious TEAEs applying a frequency threshold of 5%;
- All TEAEs resulting in dose reduction (any study treatment and by study treatment components);
- All TEAEs resulting in temporarily treatment discontinuation (any study treatment and by study treatment components);
- All late-onset AEs (starting more than 30 days after the last dose date).

In addition, the following tables and figures summarizing the frequency of participants with TEAEs by PT will be produced:

- TEAEs with at least 10% overall incidence (table and figure);
- Treatment-related TEAEs with at least 10% overall incidence (any study treatment and by study treatment components, table and figure).

Furthermore, the following outputs will be provided for the 3-tier approach:

- Pre-specified Tier 1/3 TEAEs (displayed by term and PT including incidences, risk differences and corresponding 95% CIs for the difference in group incidence proportions, as applicable according to rule-of-4)
- Pre-specified Tier 1/3 TEAEs with NCI-CTCAE grade ⩾3 (displayed by term and PT including incidences, risk differences and corresponding 95% CIs for the difference in group incidence proportions, as applicable according to rule-of-4)
- Tier 2/3 TEAEs (displayed by SOC and PT including incidences, risk differences and corresponding 95% CIs for the difference in group incidence proportions, as applicable according to Rule-of-4)
- For Tier 1/2 TEAEs each, a Forest plot displaying incidence rates, differences in incidence rates as well as the 95% CIs for the difference will be provided.

Participant data listings will be provided for all AEs, treatment-related AEs and TEAEs considered as being dose-modifying events. Relationship and action taken to each study treatment (xevinapant/matched placebo, CT or IMRT) will be provided.

15.1.2 Adverse Events Leading to Discontinuation of Study Treatment

The frequency (number and percentage) of participants with each of the following will be presented for TEAEs leading to permanent discontinuation of each study treatment by study treatment arm, SOC and PT:

- All TEAEs resulting in treatment discontinuation (any study treatment and by study treatment components).

A listing of all TEAEs leading to treatment discontinuation will also be provided, including relevant information.

15.2 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

15.2.1 Deaths

The frequency (number and percentage) of participants with each of the following will be presented by treatment arm:

- *All TEAEs leading to death by SOC and PT (IDMC only);*
- Death (overall and by primary reason);
- Death within the on-treatment period (overall and by primary reason);
- Death within 126 days after first dose of study intervention (overall and by primary reason).

Listings will be provided presenting all deaths occurred in the study.

15.2.2 Serious Adverse Events

The frequency (number and percentage) of participants with each of the following will be presented for serious TEAEs by treatment arm, SOC and PT:

- All serious TEAEs
- All treatment-related serious TEAEs (any study treatment and by study treatment components)
- All serious TEAEs by worst NCI-CTCAE grade (Any, ≥3, ≥4 and 5)
- All serious treatment-related TEAEs by worst NCI-CTCAE grade (Any, ≥3, ≥4 and 5) (any study treatment and by study treatment components)

The listings of SAEs will also be provided with the relevant information.

15.2.3 Other Significant Adverse Events

The analyses for AESIs are defined distinguishing 3 main categories, which are defined in Sections 15.2.3.1, 15.2.3.2, and 15.2.3.3 and rely solely on investigator reported AEs. Those are intended as definition for main analyses of AESIs. The legacy definition of AESIs as provided in Section 15.2.3.4 defines the categories of 'AESIs During the On-Treatment Period' and 'Late-onset AESIs During the Follow-up Period', which are supposed to be used for the purpose of consistency with protocol definition. The outputs for analyses based on legacy definition of AESIs must be separated from the main AESI analysis outputs and clearly labeled as AESI per legacy definition.

15.2.3.1 Treatment-emergent AESIs

Treatment-emergent AESIs are defined as TEAEs (see Section 15.1), which match the criteria in Table 20 (see Appendix 5). The potential tier 1 events are defined as the treatment-emergent AESIs and corresponding displays are specified in Section 15.1.1.

CCI [redacted]

[redacted]

15.2.3.4 Legacy Definition of AESIs

All AESIs during the on-treatment period following the criteria as described in Appendix 2 will summarize the number and percentage of participants reporting at least one event together with

the corresponding number of events and will be grouped by treatment arm. The following will be displayed:

- Any AESI during treatment period
- Any AESI fulfilling the Aspartate Aminotransferase (AST)/ Alanine Aminotransferase (ALT) criteria
- Any AESI of Lipase > 5* Upper Limit of Normal (ULN)
- Any AESI of Amylase > 5* ULN
- Any AESI fulfilling the acute renal failure criteria
    - Any renal failure TEAE
    - Any renal failure (eGFR, CRCL, creatinine elevations)
- Any AESI fulfilling the QT interval corrected using Fridericia's formula (QTcF) prolongation/other Electrocardiogram (ECG) abnormality criteria

CCI

CCI

15.3 Clinical Laboratory Evaluation

Clinical laboratory results assessed by central and local laboratory will be graded according to NCI-CTCAE, Version 5.0 (see Appendix 3) and classified as low, normal, and high, or normal/abnormal according to the laboratory normal ranges.

All hematology, blood biochemistry, urinalysis, coagulation, and serology evaluations will be listed per participant for each assessment. Standard International System (SI) units will be used for listing and reporting. For continues laboratory results assessed by the central laboratory, actual value observed and change from baseline summary tables will be provided for each laboratory parameter by visit and treatment arm, together with the maximum on-treatment change. For categorical laboratory results frequency tables showing the number and percentage of participants will be provided for each laboratory parameter by visit and treatment arm.

For statistical and graphical summaries of the safety laboratory tests, values below or above the limit of detection (e.g., '<3' or '>500') are substituted with the lower/upper limit of detection (e.g., '<3' is substituted by 3, '>500' is substituted by 500). In the participant data listings, the values will be shown as collected, e.g., including the < or > sign.

CCI

CCI

Abnormalities classified according to NCI-CTCAE will be described using the maximum grade. For those parameters which are graded with two toxicities such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (e.g., hypokalemia) grades at baseline and post baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (e.g., hyperkalemia), and vice versa.

Shifts from baseline NCI-CTCAE grades to maximum on-treatment NCI-CTCAE grade will be presented by treatment arm for central and local laboratory results. Only scheduled visits will be included in the change from baseline summary table; however, unscheduled visits will be included in maximum post-baseline shift tables, liver function tests summaries, and participant data listings. Abnormal laboratory results will be presented in an additional listing.

In addition, a boxplot will be presented for each laboratory parameter assessed by the central laboratory by treatment arm, showing the distributions of change from baseline at each on-treatment study visit.

An eDISH (evaluation of Drug-Induced Serious Hepatotoxicity) plot showing the log/log display of correlation between on-treatment peak total bilirubin versus ALT/AST, both in multiples of ULN, with horizontal and vertical lines indicating Hy's law thresholds, i.e., ALT/AST = 3 × ULN and total bilirubin = 2 × ULN for both treatment arms will be provided for central and local laboratory results.

15.4 Vital Signs

Vital sign measurements will include blood pressure (mmHg), heart rate (bpm), temperature (°C), and respiratory rate (breaths/min). Height will only be measured at screening and weight will be measured as indicated in the schedule of events (Section 1.3 of the protocol). If temperature is collected in °F, values will be converted by using °C= (°F - 32) x 5/9. All vital signs measurements will be listed. Actual value observed and change from baseline summary table will be provided for each vital signs parameter by visit and treatment arm together with the maximum on-treatment increase/decrease. The maximum on-treatment increase/decrease from baseline will be calculated by vital sign measurement (temperature increase, weight increase/decrease, heart rate increase/decrease, systolic blood pressure increase/decrease, diastolic blood pressure increase/decrease and respiratory rate increase/decrease). Only scheduled visits will be included in the change from baseline summary table. All visits will be included in the participant data listing.

15.5 Other Safety or Tolerability Evaluations

15.5.1 ECG

A standard single 12-lead ECG will be recorded according to the schedule shown in Table 8-1 of the protocol. In case of abnormal ECG findings, triplicate ECG readings will be performed. Abnormal ECG results including heart rate (bpm), PR interval (msec), RR interval (msec), QRS duration (msec), QT interval (msec), QTcF interval (msec) and morphology rhythm, and classification of within normal limits, abnormal not clinically significant, and abnormal clinically significant with a description for abnormal findings will be presented in a participant data listing.

Actual value observed and change from baseline summary table will be provided for each ECG parameter by visit, time point and treatment arm together with the maximum on-treatment increase/decrease. The maximum on-treatment increase/decrease from baseline will be calculated by ECG measurement (heart rate increase/decrease, PR interval increase, RR interval increase, QRS duration increase, QT interval increase and and QTcF interval increase). In case of triplicate ECG readings, the mean of the triplicate ECG parameter values will be used for calculations. For delta QTcF the 90% 2-sided CI will be additionally included with a flag if upper bound > 10 msec.

Shifts from baseline to worst post-baseline values will be presented for the ECG overall evaluation. Additionally, the number and percentage of participants with noteworthy on-treatment QTcF values (<= 450 msec, >450 to 480 msec, >480 to 500 msec, >500 msec) at each timepoint, the number and percentage of participants with noteworthy on-treatment QTcF changes from baseline (<=30 msec, >30 to 60 msec; >60 msec) will be provided by treatment arm for each post-baseline timepoint. In addition, for each post-baseline timepoint the number and percentage of participants with any QTcF value > 500 msec and QTcF change from baseline > 60 msec during the on-treatment period will be displayed.

A box plot will be presented for each ECG parameter by treatment arm, showing the distributions of change from baseline at each on-treatment study visit and timepoint.

Only scheduled visits will be included in summary tables and figures. All visits will be included in the participant data listing.

15.5.2 ECOG Performance Status

All ECOG-PS results will be listed. The number and percentage of participants with each score will be presented by treatment arm at baseline and at each scheduled post-baseline visit. In addition shifts from baseline to worst post-baseline scores will be presented.

All visits will be included in the participant data listing.

15.5.3 Transfusions

The number of transfusions received per participant will be presented, together with the type of transfusion by treatment arm. A supporting participant data listing will be provided.

15.5.4 Ear Examination and Audiogram

For clinical ear examination assessment and audiogram the Investigator's interpretation will be presented in a frequency table by visit and treatment arm. A supporting participant data listing will be provided.

15.5.5 Dental Examination

Dental examination data will be listed at each time point.

15.5.6 Nutritional Status

Nutritional status data will be listed at each time point.

[redacted]

[redacted]

CCI [redacted] [redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted]

CCI [redacted] [redacted]

[redacted]

[redacted]

CCI

CCI

CCI

16.2.3 Key Baseline Covariates for HRQL Analysis

Key baseline covariates that should be included in all regression models for HRQL analysis are the following:

Randomization stratification factors (from IWRS):

- Primary tumor site: larynx versus other (oropharynx and hypopharynx);
- Lymph node involvement: N0-1 versus N2 versus N3;
- T size: T4 versus other;

Other covariates :

- ECOG-PS: 0 versus ≥1;
- Age (continuous);
- Gender;
- Baseline HRQL score (for the item/scale in question in the analysis).

Key baseline covariates include stratification factors for randomization.

16.2.4 Analysis of HRQL Patterns

16.2.4.1 Descriptive Analysis of Health-Related Quality of Life Missing Data

Number and percentage of subjects with expected visits and their completion and compliance rates will be calculated at each assessment timepoint and by treatment arm following standard EORTC procedure as defined by

- $\%\ Compliance = 100 \times \frac{\text{number of subjects with at least one item of the HRQL questionnaire available}}{\text{number of subjects for whom a HRQOL questionnaire is expected}}$
- $\%\ Completion = 100 \times \frac{\text{number of subjects with at least one item of the HRQL questionnaire available}}{\text{number of randomized/treated subjects}}$.

Number and percentage of non-expected visits will be reported by reasons (death, lost-to follow-up, withdrawal of consent, administrative reasons, other) along with the number and percentage of expected visits and the reasons for non-completion for each assessment timepoint and by treatment arm.

16.2.5 Analysis of Health-Related Quality of Life Data

16.2.5.1 Descriptive Health-Related Quality of Life Analysis

Baseline HRQL scores are collected during the screening period. Scores will be summarized by treatment arm using descriptive statistics for the CCI component, each QLQ-C30 and QLQ-HN35 scales and items. Supporting participant data listings will be provided.

16.2.5.2 Primary Health-Related Quality of Life Analysis

The endpoints of the primary HRQoL analysis will be the change from baseline at the end of combination therapy (any measurements until day 74) and at the completion of investigational therapy for EORTC QLQ-C30 subscale scores on GHS and fatigue and EORTC QLQ- HN35 subscale scores on swallowing and pain.

For the analysis of change from baseline until end of combination therapy (any measurements until day 74), HRQoL observations will be excluded after completion of combination therapy (while-on-treatment strategy). For the analysis until completion of the investigational therapy (EOT), observations will be excluded that are obtained in the EFS follow-up (while-not-in EFS follow-up) and also measurements after start of subsequent therapy (while not treated with subsequent anti-cancer therapy). Change from baseline in the primary PRO-endpoints will be analyzed by means of analysis of covariance (ANCOVA). The ANCOVA model will include the treatment arm, the randomization strata (from IWRS as given in Section 16.2.3) as classification factors and the PRO baseline score as covariate. The average effect of study interventions from randomization until end of combination therapy (any measurements until day 74) or treatment discontinuation or death, whichever comes first, and the average effect of study interventions from randomization until end of treatment or death, whichever comes first, will be compared between interventions using the difference in terms of least-squares means from the ANCOVA model. The difference in least squares (LS) means, corresponding standard error, 2-sided 95% confidence interval, 1-sided p-value for the difference between interventions will be presented without adjustment for multiplicity.

Supplementary analyses will be performed repeating ANCOVA but including the covariates as given in Section 16.2.3 as fixed factors.

16.2.5.3 Exploratory Health-Related Quality of Life Analysis

Treatment groups will be compared for the time points until end of treatment following the ANCOVA model above for all EORTC QLQ-C30 and HN35 domains in addition to the primary domains above and also for EQ5D-VAS.

The effect sizes including 95% confidence intervals in terms of differences in LS means will be displayed in a Forest plot accordingly.

CCI

CCI

Unadjusted mean PRO scores will be graphically displayed over time for all EORTC QLQ-C30 and HN35 domains and EQ5D-VAS. Time points with a low number of patients under observation (<10) will be excluded.

Musoro et al (2020) published "Minimally important differences for interpreting European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 scores in patients with head and neck cancer" in 2020 that will further guide analysis and interpretation of within and between group comparisons for all scales/items included in EORTC questionnaires.

CCI [redacted]

16.2.5.3.1 Additional Exploratory Analyses

The proportion of event-free participants who improved / maintained / deteriorated in each domains of for all EORTC QLQ-C30 and HN35 compared to baseline score after completion of the investigational therapy will be presented by treatment group. The threshold for maintenance should correspond to the anchor based MID for either improvement or deterioration of within treatment group differences as published by Musoro et al (2020). If not available a distribution-based threshold of 0.3 SD will be considered.

16.2.6 Health State Utilities Assessments Reporting Guidelines

CONSORT standards, PRO extension, should be followed to report HRQL analysis and results. Details to analyses health state utilities and health utility index will be described in a separate analysis plan for health economic purposes.

17 References

Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. *Biometrics* 1982; 38, 29–41. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. 2009. 'New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)', *Eur J Cancer*, 45 (2): 228-247.

Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. Chapman & Hall, CRC Press, London 2000:189-193.

Kalbfleisch JD, Prentice LP. The Statistical Analysis of Failure Time Data. *John Wiley and Sons*, 2nd edition, 2002.

Miettinen O, Nurminen M. Comparative analysis of two rates. *Stat Med* 1985;4:213-226

Musoro et al. 2020. 'Minimally important differences for interpreting European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 scores in patients with head and neck cancer', https://onlinelibrary.wiley.com/doi/full/10.1002/hed.26363

Schemper, M, and Smith, TL 1996, 'A note on quantifying follow-up in studies of failure time. *Controlled Clinical Trials*, 17 (4): 343-346.

18 Appendices

Appendix 1 – Health Related Quality of Life Questionnaires Scoring Manuals

EORTC QLQ-C30 -C30

The EORTC QLQ-C30 v.3 questionnaire contains 30 single items (Q1 – Q30). Q1 – Q28 range from 1 to 4, with 1 being the most favourable answer and 4 the worst case (1 = Not at all, 2 = A little, 3 = Quite a bit, 4 = Very much). Q29 and Q30 range from 1 (= Very poor) to 7 (= Excellent),) with 1 being the worst case and 7 the most favourable answer.

The subscores will include 5 functional scales, 3 symptom scales, a global health status / QoL scale, and 6 single items. Each of the multi-item scales includes a different set of items - no item occurs in more than one scale.

Table 16: EORTC QLQ-C30 definition of subscores

	Item Range	Single Items
Global health status / QoL		
Global health status / QoL (QL)	6	Q29, Q30
Functional scales		
Physical functioning (PF)	3	Q1 - Q5
Role functioning (RF)	3	Q6, Q7
Emotional functioning (EF)	3	Q21 - Q24
Cognitive functioning (CF)	3	Q20, Q25
Social functioning (SF)	3	Q26, Q27
Symptom scales / items		
Fatigue (FA)	3	Q10, Q12, Q18
Nausea and vomiting (NV)	3	Q14, Q15
Pain (PA)	3	Q9, Q19

	Item Range	Single Items
Dyspnoea (DY)	3	Q8
Insomnia (SL)	3	Q11
Appetite loss (AP)	3	Q13
Constipation (CO)	3	Q16
Diarrhoea (DI)	3	Q17
Financial difficulties (FI)	3	Q28

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.

All of the subscores range in score from 0 to 100. A high scale score represents a higher response level. Thus, a high score for a functional scale represents a high / healthy level of functioning; a high score for the global health status / QoL represents a high QoL, but a high score for a symptom scale / item represents a high level of symptomatology / problems.

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

Estimate the average of the items (I1, I2, ..In) that contribute to the scale; this is the raw score.

Raw score = RS = (I1 + I2 +…+ In) / n

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.

For functional scales: Score = {1 – [(RS - 1) / range] } x 100

For symptom scales and global health status / QoL: Score = {(RS - 1) / range} x 100

where range is the difference between the maximum possible value of RS and the minimum possible value of RS.

EORTC QLQ-H&N35

The EORTC QLQ-H&N35 questionnaire contains 35 single items (Q1 – Q35). Q1 – Q30 range from 1 to 4, with 1 being the most favourable answer and 4 the worst case (1 = Not at all, 2 = A little, 3 = Quite a bit, 4 = Very much). Q31 to Q32 are yes/no items.

The subscores will include 7 multi-item scales and 11 single items. Each of the multi-item scales includes a different set of items - no item occurs in more than one scale.

Table 17: EORTC QLQ-H&N35 definition of subscores

	Item Range	Single Items
Symptom scales / items		
Pain (HNPA)	3	Q1 - Q4
Swallowing (HNSW)	3	Q5 - Q8
Senses problems (HNSE)	3	Q13, Q14
Speech problems (HNSP)	3	Q16, Q23, Q24
Trouble with social eating (HNSO)	3	Q19 - Q22
Trouble with social contact (HNSC)	3	Q18, Q25 - Q28
Less sexuality (HNSX)	3	Q29, Q30
Teeth (HNTE)	3	Q9
Opening mouth (HNOM)	3	Q10
Dry mouth (HNDR)	3	Q11
Sticky saliva (HNSS)	3	Q12
Coughing (HNCO)	3	Q15
Felt ill (HNFI)	3	Q17
Pain killers (HNPK)	1	Q31
Nutritional supplements (HNNU)	1	Q32
Feeding tube (HNFE)	1	Q33
Weight loss (HNWL)	1	Q34
Weight gain (HNWG)	1	Q35

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

For all items and scales, high scores indicate more problems (i.e., there are no function scales in which high scores would mean better functioning). The scoring approach for the QLQ-H&N35 is identical in principle to that for the symptom scales/ / single items of the QLQ-C30.

CCI [redacted]

[redacted]

[redacted]

CCI [redacted]

CCI [redacted]

Appendix 2 – Legacy Definition of Adverse Events of Special Interest

AESIs During the On-Treatment Period

The following events will be considered AESIs during the treatment period regardless of relationship to study treatments (xevinapant/placebo, chemotherapy and IMRT):

- AST or ALT increases >3 × Upper Limit of Normal (ULN) associated with total bilirubin >2 × ULN and alkaline phosphatase (ALP) <2 × ULN.
- Lipase or amylase increases >5 × ULN.
- Acute renal failure - the composite term includes:
    - The MedDRA PTs: acute kidney injury, anuria, azotaemia, continuous haemodiafiltration, dialysis, haemodialysis, haemofiltration, nephropathy toxic, oliguria, peritoneal dialysis, prerenal failure, renal failure, renal impairment and subacute kidney injury, and/or
    - Grade ≥ 3 creatinine elevations (> 3.0 x baseline; > 3.0 - 6.0 x ULN) and/or Grade ≥ 3 eGFR/creatinine clearance (eGFR or CrCl < 29 ml/min/1.73 m2).
- Any QTcF prolongation >30 ms compared to baseline confirmed by the mean QTcF value of the triplicate readings or other ECG abnormalities such as Torsade de pointes, ventricular tachycardia, ventricular fibrillation, flutter or any other new cardiac abnormality at the ECG or during the physical examination considered as clinically significant.

CCI

[redacted]

CCI

CCI

CCI

CCI

CCI

Appendix 3 - NCI-CTCAE grading for Laboratory Parameters

Table 18: NCI-CTC gradable parameters

Laboratory Assessment	Parameters	Name in NCI-CTC	Direction(s) of abnormality
Hematology	Hemoglobin	Anemia/Hemoglobin increased	Low/High
	Leukocytes (WBC)	White blood cell decreased / Leukocytosis	Low/High
	Neutrophils	Neutrophil count decreased	Low
	Lymphocytes	Lymphocyte count decreased / increased	Low/High
	Eosinophils	Eosinophila [a]	High
	Platelets	Platelet count decreased	Low
Biochemistry	Albumin	Hypoalbuminemia	Low
	Alanine Aminotransferase (ALT)	Alanine Aminotransferase increased [a]	High
	Aspartate Aminotransferase (AST)	Aspartate Aminotransferase increased [a]	High
	Alkaline Phosphatase	Alkaline Phosphatase increased [a]	High
	Gamma Glutamyl Transferase (GGT)	GGT increased [a]	High
	Total Bilirubin	Blood bilirubin increased [a]	High
	Amylase	Serum amylase increased	High
	Lipase	Lipase increased	High
	Creatinine / eGFR / CrCl	Creatinine increased [a]	High
	Sodium	Hyponatremia / Hypernatremia	Low / High
	Potassium	Hypokalemia / Hyperkalemia	Low / High
	Calcium	Hypocalcemia / Hypercalcemia	Low / High
	Glucose	Hypoglycemia / Hyperglycemia	Low / High
	Magnesium	Hypomagnesemia/Hypermagnesemia	Low / High
Coagulation	Activated Partial Thromboplastin Time	Activated Partial Thromboplastin Time prolonged	High
	Prothrombin Intl. Normalized Ration (INR)	INR increased	High

Note: parameters with both Low and High directions of abnormality are going to be split. For example, Calcium is going to be split in Calcium Low and Calcium High.

[a] on treatment grading dependent on baseline grading

Table 19: NCI-CTC non-gradable parameters

Laboratory Assessment	Parameters
Hematology	Hematocrit
	Mean corpuscular volume (MCV)
	Mean corpuscular hemoglobin (MCH)
	Reticulocytes
	Reticulocytes/Erythrocytes
	Neutrophils/Leukocytes (%)
	Lymphocytes/Leukocytes (%)
	Basophils
	Basophils/Leukocytes (%)
	Eosinophils/Leukocytes (%)
	Monocytes
	Monocytes/Leukocytes (%)
Biochemistry	Total Protein
	Urea Nitrogen (BUN)
	Serum cystatin C
Coagulation	Prothrombin Time
	Prothrombin Time/Standard
Urinalysis	Blood
	Bacteria (Microscopic analysis)
	Casts (Microscopic analysis)
	Crystals (Microscopic analysis)
	Epithelial Cells (Microscopic analysis)
	Bilirubin
	Glucose in Urine
	Urobilinogen
	Ketones
	Nitrite
	pH
	Proteins in Urine
	Erythrocytes in Urine (Microscopic analysis)
	Leukocytes esterase
	Leukocytes in Urine (Microscopic analysis)
	Specific gravity

Appendix 4 – Definition of Western Europe

The following countries are part of Western Europe:

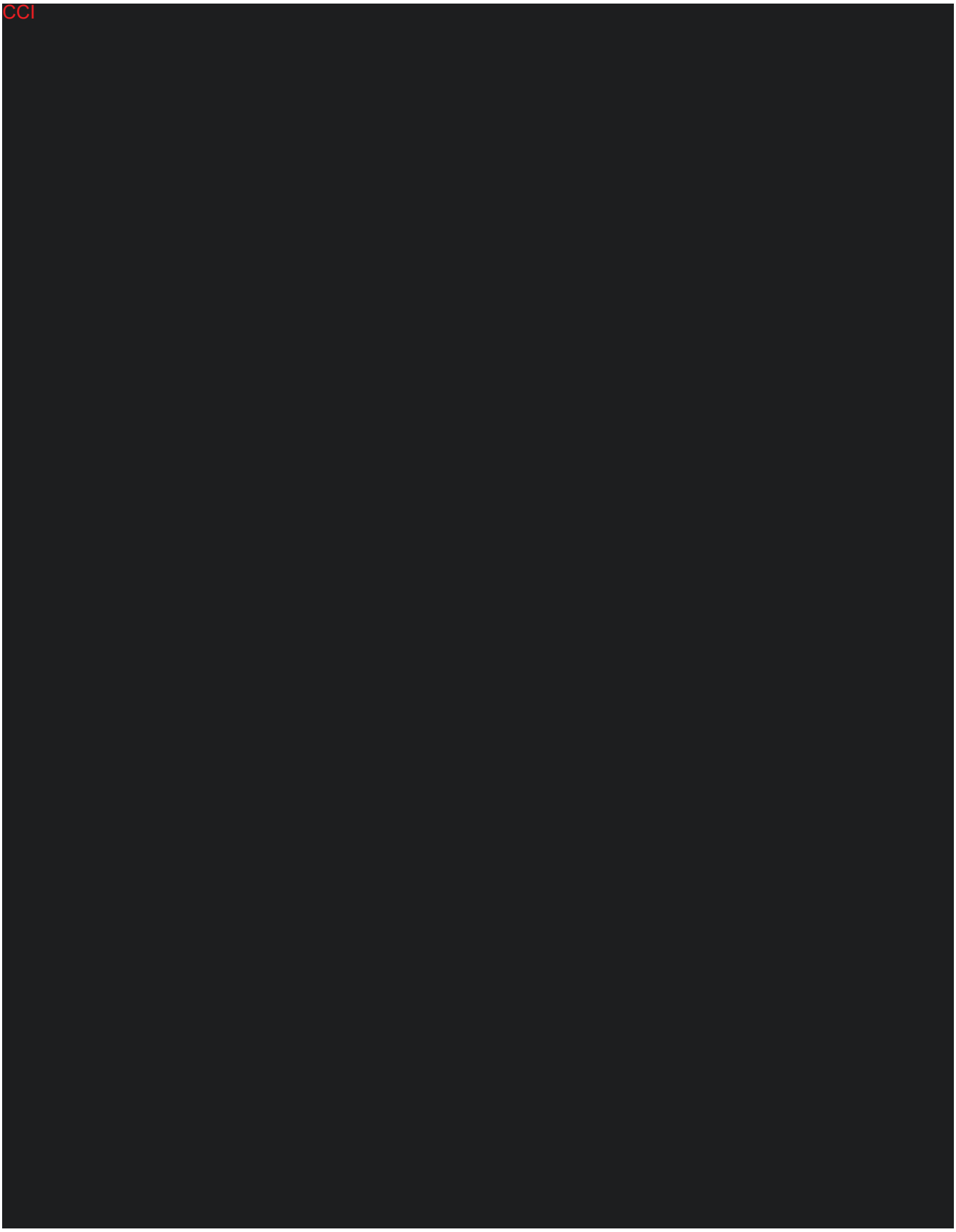
Austria
Belgium
Germany
Greece
Italy
France
Portugal
Spain
Switzerland
United Kingdom

CCI

[redacted]

CCI [redacted]

CCI [redacted]



CCI

CCI

CCI

[redacted]

CCI [redacted]

CCI [redacted]

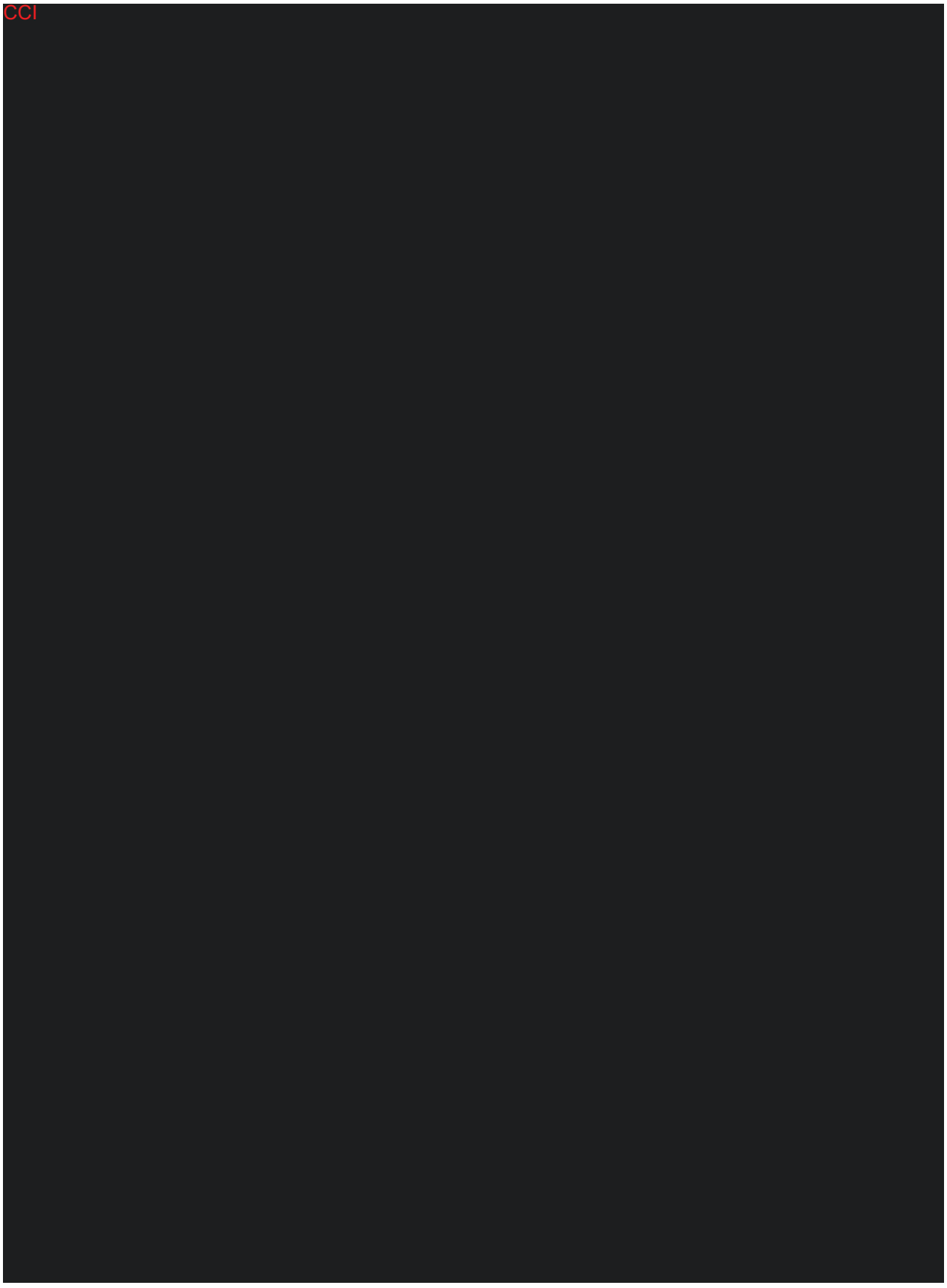
CCI

[redacted]

CCI [redacted]

CCI [redacted]

CCI



CCI

CCI

CCI

CCI

CCI

CCI

Appendix 6 – Programming Specifications

Stratified log-rank test:

```
proc lifetest data= ADTTE;
  time AVAL * CNSR(1,2,3,4);
  strata T N / group= TRTPN test= (logrank) trend;
  ods output TrendTests= logrank;
run;

CNSR ≠ 0 if censored, 0 otherwise
AVAL = EFS time in months
Randomization stratification factors: T = Tumor size, N = Lymph node involvement
TRTPN = planned treatment
```

Cox-Regression:

```
proc phreg data= ADTTE;
  class REGION(ref='Western Europe') SITE(ref='Other') N(ref='N0-1')
        T(ref='Other') TRTP (ref='Placebo') /param=ref;
  model AVAL*CNSR(1,2,3,4) = TRTP REGION SITE N T / rl ties = discrete;
  hazardratio 'treatment' TRTP/ diff=ref;
  hazardratio 'region' REGION / diff=ref;
  hazardratio 'tumor site' SITE / diff=ref;
  hazardratio 'lymph node' N / diff=ref;
  hazardratio 'T size' T / diff=ref;
  ods output HazardRatios= HR;
run;

CNSR ≠ 0 if censored, 0 otherwise
AVAL = EFS time in months
Randomization stratification factors:
REGION = Region, SITE = Tumor Site, T = Tumor size, N = Lymph node involvement
TRTP = planned treatment
```

Kaplan-Meier Estimates:

```
proc lifetest data=ADTTE timelist=9,12,18,24,30,36 outsurv=surv reduceout;
  time AVAL * CNSR(1,2,3,4);
  strata TRTPN;
  ods output Quartiles=quar;
  ods output ProductLimitEstimates=est;
run;

CNSR ≠ 0 if censored, 0 otherwise
AVAL = EFS time in months
TRTPN = planned treatment
```

Logistic regression model:

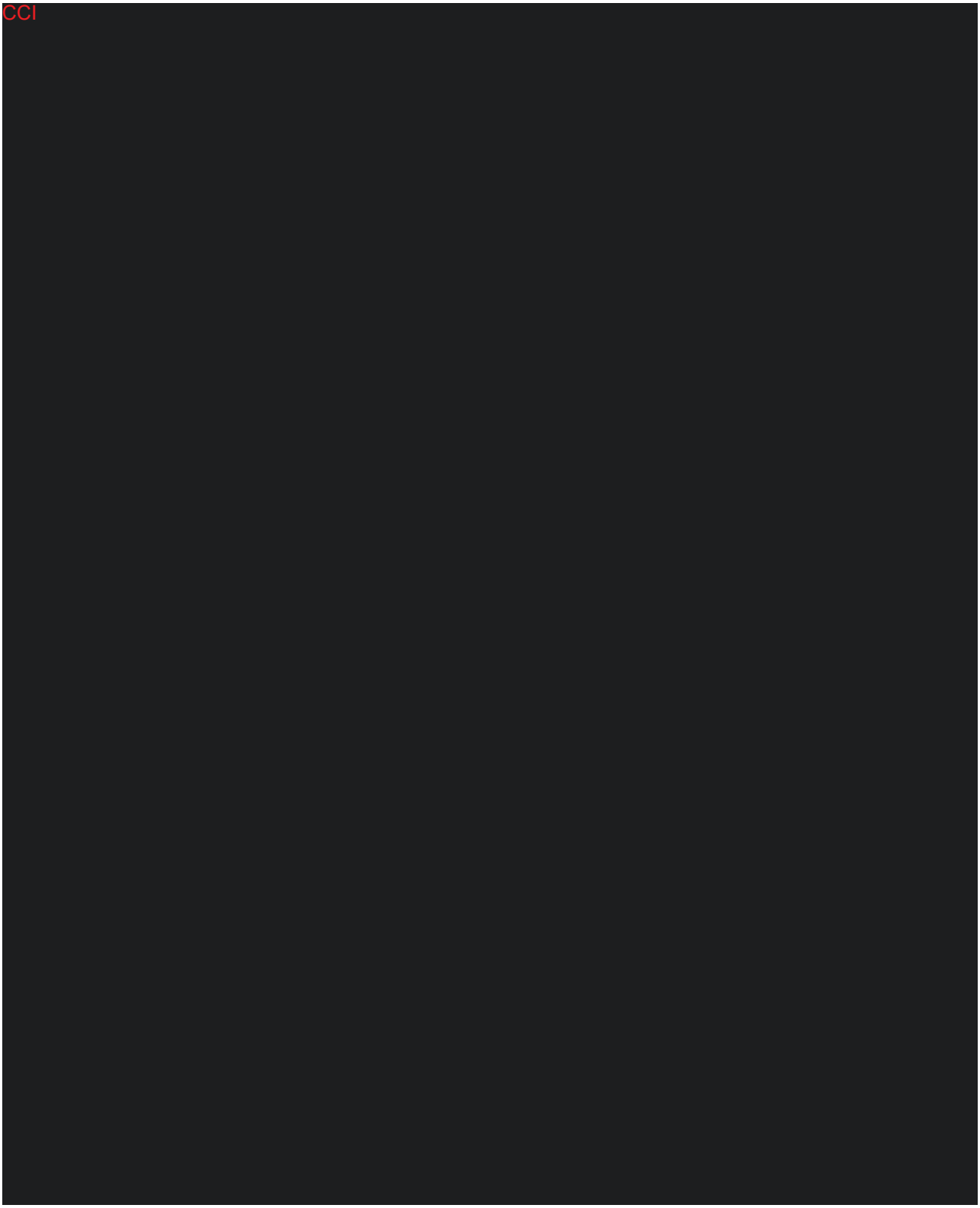
```
proc logistic data= ADTUM;
    class REGION(ref='Western Europe') SITE(ref='Other') N(ref='N0-1')
          T(ref='Other') TRTP (ref='Placebo') /param=ref;
  model AVAL(event='1') = TRTP REGION SITE N T / cl orpvalue ;
  oddsratio TRTP;
run;

AVAL = Response status (0: 'no response'; 1: 'response')
Randomization stratification factors:
REGION = Region, SITE = Tumor Site, T = Tumor size, N = Lymph node involvement
TRTP = planned treatment
```

CCI

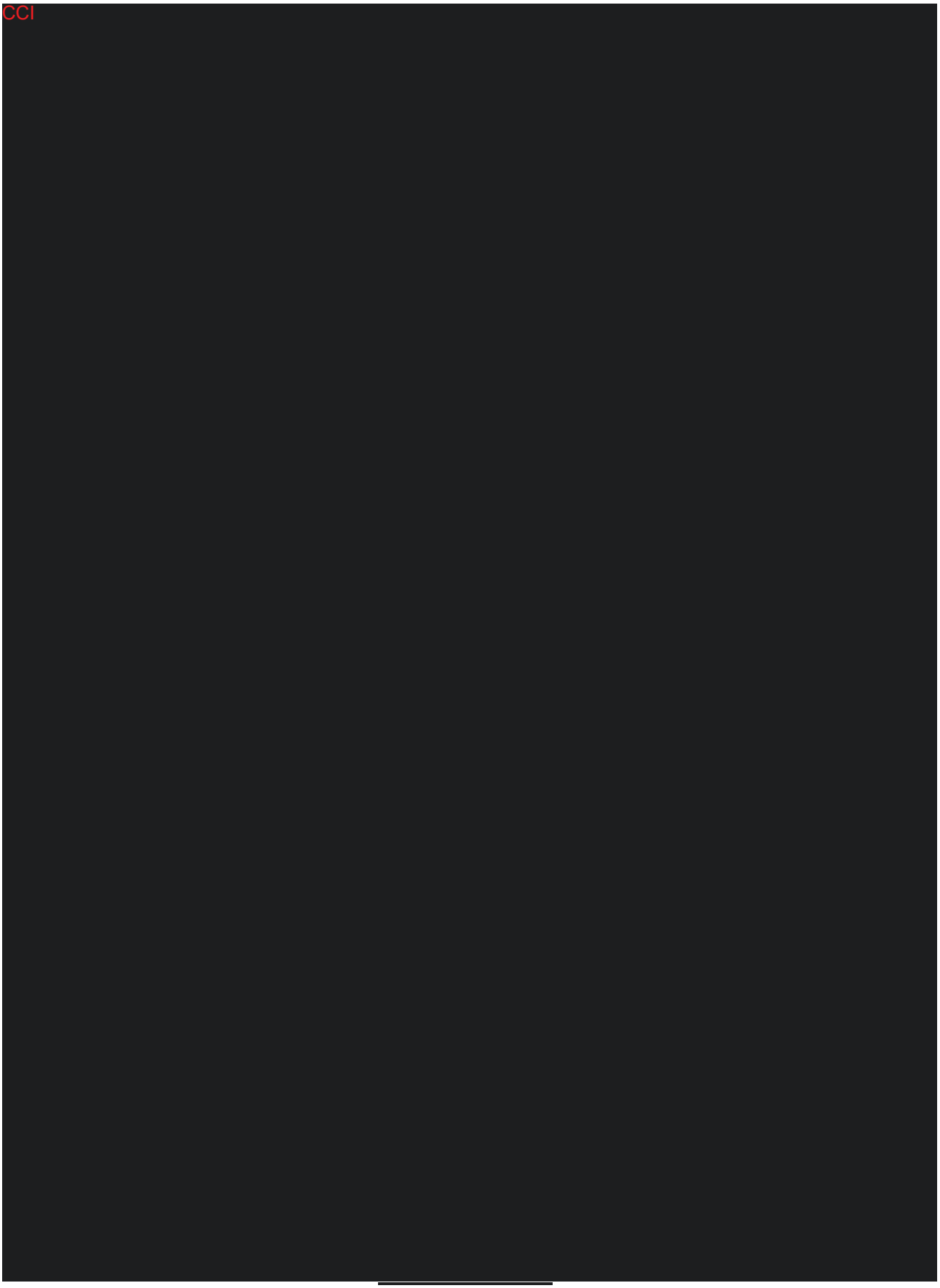
[redacted]

CCI [redacted]

CCI [redacted]



CCI

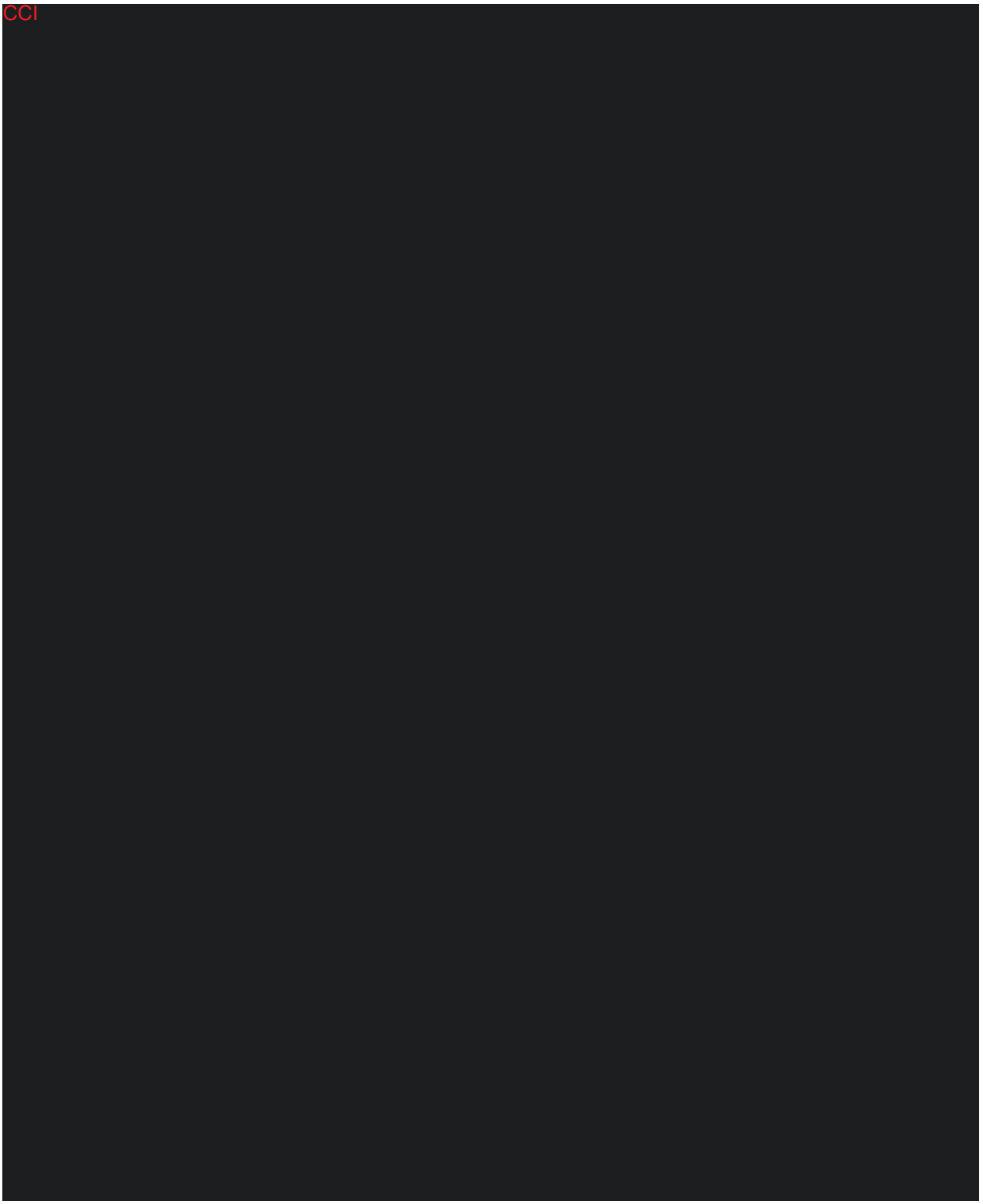


CCI

[redacted]

CCI [redacted]

CCI [redacted]



CCI

CCI

Appendix 8 - Country-specific Amendment for China

In order to meet the requirement of regulatory consultation and submission to the Chinese National Medical Products Administration, the following separate analyses will be considered based on the global study population and subjects included in China specific extension cohort (if any),

- Subjects from East Asia (China, Japan, Korea, Taiwan)
- China Mainland subjects

The overall global and Chinese population, and accordingly the China-specific Safety and Pharmacokinetics analysis sets will be analyzed following the same principles as outlined in Section "Definition of Analysis Sets".

The analyses described in this section will not be part of the main CSR. The analyses outlined in this section will be used to support China NDA filling.

The following summaries will be prepared by treatment group in subjects from East Asian and Chinese mainland population (Chinese vs. non-Chinese):

- Screen failure
- Participant disposition
- Participants randomized vs treated
- Analysis sets
- Important protocol deviations
- Demographic characteristics and baseline characteristics
- Baseline cancer characteristics
- Medical and surgical history
- Prior medication
- Concomitant medication
- Disease history
- Subsequent antineoplastic therapies
- Efficacy (EFS assessed by BIRC and investigator, OS, PFS assessed by BIRC and investigator, LRC assessed by investigator and BIRC, Tumor response assessed by BIRC and investigator, CR assessed by BIRC and investigator, DoR assessed by BIRC and investigator. Some efficacy analyses may not be performed at the time of interim analysis or primary analysis due confidence in evidence of efficacy for expected events in China mainland subjects from global ITT population and the specified extension arm.)
- Treatment exposure
- Treatment exposure by treatment cycle
- Number of subjects with dose changes per treatment cycle

- TEAEs (overall, TEAE, related, grade ≥ 3, related grade ≥ 3, serious, serious treatment related, leading to death, serious leading to death, serious related leading to death, AESI, leading to temporary treatment interruption and leading to permanent treatment discontinuation, pre-specified tier 1/3 treatment related, Xevinapant/matched placebo related)
- Deaths
- Key lab parameters (hematology, blood biochemistry), vital signs, and ECG
- ECOG
- Forest plot for EFS by BIRC
- Kaplan-Meier plot for EFS assessed by BIRC, OS, PFS assessed by BIRC, LRC assessed by investigator and DoR by BIRC

In additional, the following summaries will also be prepared by treatment group in global population and subjects from China specific extension cohort by region/subgroup and treatment group using the layout as below,

	Chinese SAF		East Asian SAF		Global SAF	
	Xevinapant + CRT (N = xxx)	Placebo + CRT (N=xxx)	Xevinapant + CRT (N = xxx)	Placebo + CRT (N=xxx)	Xevinapant + CRT (N = xxx)	Placebo + CRT (N=xxx)

- TEAE by SOC and PT
- TEAE with at least 10% overall incidence by PT (any study treatment in Chinese mainland subjects, subjects from East Asian, or Global safety population)
- TEAE by SOC, PT and worst NCI-CTCAE (only for East Asian and global safety population)
- Treatment related TEAE by SOC and PT
- Treatment related TEAE with at least 10% overall incidence by PT (any study treatment in Chinese mainland subjects, subjects from East Asian, or Global safety population and by study treatment components)
- Treatment related TEAE by SOC, PT and worst NCI-CTCAE (only for East Asian and global safety population)
- Serious TEAE by SOC and PT
- TEAE resulting in permanent treatment discontinuation by SOC and PT
- TEAE resulting in temporary treatment interruption by SOC and PT

The general specifications for data analyses and statistical methods for analyzing respective endpoints described in this Integrated Analysis Plan are applicable for the analysis with the following additions and exceptions:

- Subgroup Definition and Parameterization: the subgroup analysis by region will not be performed.

- Efficacy Analyses: no confirmatory statistical hypothesis will be tested.
- Safety Analyses: in addition to the safety analyses in the overall Chinese population, analyses for the pooled population (global SAF set combined with the China-specific extension cohort) will be performed

The overall Chinese population analysis will be triggered once 65 EFS events across both arms (as assessed by the BIRC) are observed in China Mainland participants.

Additional outputs (summary table and listing) will be generated for a description of the impact by COVID-19 on the study in subjects from East Asian and Chinese mainland population. The number and percentage of participants will be presented for the following findings due to COVID-19:

- Adverse Events (MedDRA SMQ "COVID-19")
- Important protocol deviations
- Missed Visits (including number of missed visits and detailed reason)
- Missed efficacy evaluations (including number of missed efficacy evaluations)
- Drug administration - missed doses
- Drug administration - dose adjustments
- Treatment discontinuation
- Study discontinuation
- Death

A frequency table will be produced for the SAF analysis set to present the number of participants with important protocol deviations related to COVID-19 (categorized by frequency of participants with an important protocol deviation overall as well as by category of protocol deviation and type of protocol deviation).

In addition, important COVID-19 protocol deviations will also be listed.

Outputs related to disposition and exposure will be amended to present reason of treatment/study discontinuations due to COVID-19 and treatment changes due to COVID-19 (if possible).

Signature Page for VV-CLIN-355012 v1.0

Approval Task	PPD

Signature Page for VV-CLIN-355012 v1.0