
STATISTICAL ANALYSIS PLAN

Study Code D9077C00001

Edition Number 6.0

Date 26-Mar-2026

**A Phase II Open-label, Multicentre, Randomised Study of
Neoadjuvant and Adjuvant Treatment in Patients with Resectable,
Early-stage (II to IIIB) Non-small Cell Lung Cancer
(NeoCOAST-2)**

TABLE OF CONTENTS

TITLE PAGE.....	1
TABLE OF CONTENTS	2
LIST OF ABBREVIATIONS	6
AMENDMENT HISTORY	9
1 INTRODUCTION	16
2 CHANGES TO PROTOCOL PLANNED ANALYSES	16
3 DATA ANALYSIS CONSIDERATIONS	16
3.1 Timing of Analyses.....	16
3.2 Analysis Populations	17
3.3 General Considerations.....	20
3.3.1 General Study Level Definitions	20
3.3.2 Visit Window	23
3.3.3 Handling of Unscheduled Visits.....	26
3.3.4 Multiplicity/Multiple Comparisons	26
3.3.5 Handling of Protocol Deviations in Study Analysis.....	26
3.3.6 Missing Dates	27
3.3.7 Sample Size	28
4 STATISTICAL ANALYSIS	29
4.1 Study Population.....	29
4.1.1 Patient Disposition and Completion Status	29
4.1.1.1 Definitions and Derivations	29
4.1.1.2 Presentation.....	29
4.1.2 Analysis Sets.....	31
4.1.2.1 Definitions and Derivations	31
4.1.2.2 Presentation.....	31
4.1.3 Protocol Deviations	31
4.1.3.1 Definitions and Derivations	31
4.1.3.2 Presentation.....	31
4.1.4 Demographics	31
4.1.4.1 Definitions and Derivations	31
4.1.4.2 Presentation.....	31
4.1.5 Baseline Characteristics	32
4.1.5.1 Definitions and Derivations	32
4.1.6 Disease Characteristics	32
4.1.6.1 Definitions and Derivations	32
4.1.6.2 Presentation.....	33
4.1.7 Medical History and Concomitant Disease	33
4.1.7.1 Definitions and Derivations	33
4.1.7.2 Presentation.....	33
4.1.8 Prior and Concomitant Medications	34
4.1.8.1 Definitions and Derivations	34

4.1.8.2	Presentation.....	34
4.1.9	Study Drug Compliance	34
4.1.9.1	Definitions and Derivations	34
4.1.9.2	Presentation.....	34
4.2	Endpoint Analyses	34
4.2.1	Primary Endpoint: pCR	35
4.2.1.1	Definition (pCR).....	35
4.2.1.2	Derivations (pCR).....	35
4.2.1.3	Handling of Dropouts and Missing Data (pCR)	35
4.2.1.4	Primary Analysis of Primary Endpoint (pCR).....	35
4.2.1.5	Sensitivity Analyses of the Primary Endpoint (pCR).....	36
4.2.1.6	Supplementary Analyses of the Primary Endpoint (pCR).....	36
4.2.1.7	Subgroup Analyses (pCR)	36
4.2.2	Secondary Endpoint: EFS.....	36
4.2.2.1	Definition (EFS)	36
4.2.2.2	Derivations (EFS)	37
4.2.2.3	Handling of Dropouts and Missing Data (EFS)	41
4.2.2.4	Primary Analysis of Secondary Endpoint (EFS)	41
4.2.2.5	Sensitivity Analyses of the Secondary Endpoint (EFS)	41
4.2.2.6	Supplementary Analyses of the Secondary Endpoint (EFS)	41
4.2.2.7	Subgroup Analyses (EFS).....	41
4.2.3	Secondary Endpoint: DFS	41
4.2.3.1	Definition (DFS).....	42
4.2.3.2	Derivations (DFS).....	42
4.2.3.3	Handling of Dropouts and Missing Data (DFS)	43
4.2.3.4	Primary Analysis of Secondary Endpoint (DFS).....	43
4.2.3.5	Sensitivity Analyses of the Secondary Endpoint (DFS).....	43
4.2.3.6	Supplementary Analyses of the Secondary Endpoint (DFS).....	43
4.2.3.7	Subgroup Analyses (DFS)	43
4.2.4	Secondary Endpoint: Feasibility to Surgery	43
4.2.4.1	Definition (Feasibility to Surgery).....	44
4.2.4.2	Derivations (Feasibility to Surgery)	44
4.2.4.3	Handling of Dropouts and Missing Data (Feasibility to Surgery).....	44
4.2.4.4	Primary Analysis of Secondary Endpoint (Feasibility to Surgery)	44
4.2.4.5	Sensitivity Analyses of the Secondary Endpoint (Feasibility to Surgery) ..	44
4.2.4.6	Supplementary Analyses of the Secondary Endpoint (Feasibility to Surgery)	44
4.2.4.7	Subgroup Analyses (Feasibility to Surgery).....	44
4.2.5	Secondary Endpoint: mPR.....	44
4.2.5.1	Definition (mPR)	45
4.2.5.2	Derivations (mPR)	45
4.2.5.3	Handling of Dropouts and Missing Data (mPR)	45
4.2.5.4	Primary Analysis of Secondary Endpoint (mPR).....	45
4.2.5.5	Sensitivity Analyses of the Secondary Endpoint (mPR)	45
4.2.5.6	Supplementary Analyses of the Secondary Endpoint (mPR)	45
4.2.5.7	Subgroup Analyses (mPR)	46

4.2.6	Secondary Endpoint: ORR	46
4.2.6.1	Definition (ORR)	46
4.2.6.2	Derivations (ORR)	47
4.2.6.3	Handling of Dropouts and Missing Data (ORR)	47
4.2.6.4	Primary Analysis of Secondary Endpoint (ORR)	47
4.2.6.5	Sensitivity Analyses of the Secondary Endpoint (ORR)	47
4.2.6.6	Supplementary Analyses of the Secondary Endpoint (ORR)	47
4.2.6.7	Subgroup Analyses (ORR)	48
4.2.7	Secondary Endpoint: OS	48
4.2.7.1	Definition (OS)	48
4.2.7.2	Derivations (OS)	48
4.2.7.3	Handling of Dropouts and Missing Data (OS)	49
4.2.7.4	Primary Analysis of Secondary Endpoint (OS)	49
4.2.7.5	Sensitivity Analyses of the Secondary Endpoint (OS)	49
4.2.7.6	Supplementary Analysis of the Secondary Endpoint (OS)	50
4.2.7.7	Subgroup Analyses (OS)	50
4.2.8	Secondary Endpoint: ctDNA	50
4.2.8.1	Definition (ctDNA)	50
4.2.8.2	Derivations (ctDNA)	50
4.2.8.3	Handling of Dropouts and Missing Data (ctDNA)	51
4.2.8.4	Primary Analysis of Secondary Endpoint (ctDNA)	51
4.2.8.5	Sensitivity Analyses of the Secondary Endpoint (ctDNA)	51
4.2.8.6	Supplementary Analysis (ctDNA)	51
4.2.8.7	Subgroup Analyses (ctDNA)	52
4.2.9	Secondary Endpoint: PK	52
4.2.9.1	Definition (PK)	52
4.2.9.2	Derivations (PK)	52
4.2.9.3	Handling of Dropouts and Missing Data (PK)	52
4.2.9.4	Primary Analysis of Secondary Endpoint (PK)	53
4.2.9.5	Additional Analyses of the Secondary Endpoint (Pharmacokinetics)	54
4.2.9.6	Subgroup Analyses (PK)	54
4.2.10	Secondary Endpoint: Immunogenicity	54
4.2.10.1	Definition (Immunogenicity)	54
4.2.10.2	Presentations (Immunogenicity)	55
4.2.11	Safety Analyses	56
4.2.11.1	Exposure	56
4.2.11.2	Adverse Events	61
4.2.11.3	Clinical Laboratory, Blood Sample	68
4.2.11.4	Clinical Laboratory, Urinalysis	69
4.2.11.5	Other Laboratory Evaluations	70
4.2.11.6	Vital Signs	70
4.2.11.7	Electrocardiogram	70
4.2.11.8	Other Safety Assessments	70
4.2.12	Other Endpoint: Biomarkers	71
4.2.12.1	Definition: Biomarkers	71
4.2.12.2	Presentations: Biomarkers	71

4.3	Pharmacodynamic Endpoint(s).....	71
4.3.1	Analysis	71
4.3.2	Definitions and Derivations	71
4.3.3	Presentation.....	71
5	INTERIM ANALYSIS	71
6	REFERENCES	73
7	APPENDICES	74
	Appendix A RECIST	74

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ADA	Anti-drug antibody
AE	Adverse event
AEF	Alive and Event-free
AEPI	Adverse event of possible interest
AESI	Adverse event of special interest
AJCC	American Joint Committee on Cancer
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
BILI	Bilirubin
BIPR	Blinded Independent Pathological Review
BMI	Body mass index
BOR	Best overall response
BP	Blood pressure
BSR	Baseline scaled ratio
CCr	Creatinine clearance rate
CDL	Clinical Data Lock
CI	Confidence interval
COVID-19	Coronavirus Disease 2019
CR	Complete response
CRF	Case report forms
CSP	Clinical Study Protocol
CSR	Clinical Study Report
ctDNA	Circulating tumour DNA
CT	Computerised tomography
CV	Coefficient of variation
Dato-DXd	Datopotamab Deruxtecan
DCO	Data cut off
DFS	Disease-free Survival
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
EFS	Event-free Survival

Abbreviation or special term	Explanation
GFR	Glomerular filtration rate
gCV	Geometric coefficient of variation
gSD	Geometric standard deviation
IASLC	International Association for the Study of Lung Cancer
ICF	Informed Consent Form
imAE	Immune mediated adverse event
IP	Investigational Product
IPD	Important Protocol Deviation
ITT	Intention-to-Treat
IRT	Interactive Response Technology
IV	Intravenous
LLOQ	Lower limit of quantification
LD	Longest Diameter
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention-to-Treat
mPR	Major pathological response
MRI	Magnetic resonance imaging
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	Non-evaluable
NED	No evidence of disease
NQ	Non-quantifiable
NS	No sample
NTL	Non-target lesion
OAE	Other significant adverse events
OR	Objective response
ORR	Objective response rate
OS	Overall survival
pCR	Pathological complete response
PD	Progressive disease
PD-L1	Programmed cell death ligand 1
PFS	Progression-free survival
PK	Pharmacokinetics
PR	Partial response
PT	Preferred term

Abbreviation or special term	Explanation
RDI	Relative dose intensity
RECIST v1.1	Response Evaluation Criteria in Solid Tumours version 1.1
RVT	Residual viable tumour
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SoA	Schedule of assessment
SOC	System organ class
SRC	Safety Review Committee
Std Dev	Standard deviation
TEAE	Treatment-emergent adverse events
TFL	Tables, figures, listings
TL	Target lesion
ULN	Upper limit of normal range
WHO	World Health Organisation

AMENDMENT HISTORY

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
N/A	19 Apr 2022	Initial approved SAP	N/A	N/A
Data presentation	23 May 2023	Section 1 updated to include reference to the latest CSP and CRF and schedule of assessments for all the 5 arms of CSP V4.	Yes	New update
Data presentation	23 May 2023	Analysis population (Section 3.2 Updated. CCI [REDACTED] [REDACTED] Resected set and modified ITT set are added. Wording for safety and immunogenicity set are updated)	Yes	Clarification
Data presentation	23 May 2023	Analyses based on PD-L1 results and definition of periods in Section 3.3.1 are updated	Yes	New update
Data presentation	23 May 2023	Visit windows updated in section 3.3.2 of MEDI5752 and AZD0171 arms for adjuvant period	Yes	Update based on CSP V4.0
Data presentation	23 May 2023	Imputation rule for partial death date is updated	Yes	This is to align the imputation rule of death date with latest early oncology SAP template
Data presentation	23 May 2023	Concomitant medication text removed from section 4.1.8.2	Yes	Concomitant medication is described in section 4.1.8.1 and text in section 4.1.8.2 is redundant.
Subgroup analyses of primary endpoint	23 May 2023	Description of subgroup analysis updated for patients who received surgery by treatment group	Yes	Clarification
Secondary endpoint(s)	23 May 2023	- Definition of event free survival updated in section 4.2.2.1 to include local or distant recurrence as determined by investigator. - Added more wording for censoring rule in section Table 4.	Yes	Clarification. Also to update the two missed visit rule according to the schedule of RECIST assessments of CSP V4.0

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
		- Removed ‘potential’ from duration of follow up. - Removed supplementary analysis from section 4.2.2.6.		
Derivation of secondary endpoint(s)	23 May 2023	Removed ‘No tumour assessments post surgery’ from situation column of Table 5 in section 4.2.3.2	Yes	Clarification
Secondary endpoint(s)	23 May 2023	Updated the primary analysis of DFS using the resected set in section 4.2.3.4	Yes	New update
Derivation of secondary endpoint(s)	23 May 2023	Added neoadjuvant in the definition of feasibility of surgery in section 4.2.4.1	Yes	Clarification
Secondary endpoint(s)	23 May 2023	Updated the primary analysis of feasibility to surgery using the safety set in section 4.2.4.4	Yes	New update
Secondary endpoint(s)	23 May 2023	Added supplementary analyses for feasibility to surgery	Yes	New update
Secondary endpoint(s)	23 May 2023	Updated subgroup analyses for mPR	Yes	Clarification
Data presentation	23 May 2023	New categories added for immunogenicity in section 4.2.10.2	Yes	New update
Data presentation	23 May 2023	Section 4.2.11.1.1 updated for the definitions of all the arms including MEDI5752, Dato-DXd and AZD0171	Yes	CSP V4.0 updated for new arms. Clarification
Data presentation	23 May 2023	Section 4.2.11.1.2 updated to include the arms MEDI5752, Dato-DXd and AZD0171	Yes	CSP V4.0 updated for new arms
Data presentation	23 May 2023	Section 4.2.11.2.1 updated to include arms MEDI5752, Dato-DXd and AZD0171	Yes	CSP V4.0 updated for new arms
Data presentation	23 May 2023	Section 4.2.11.2.2 updated to include arms MEDI5752, Dato-DXd and AZD0171.	Yes	CSP v4.0 updated for new arms

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
		Adverse events of possible interest added		New update
Derivation of secondary endpoint(s)	23 May 2023	Treatment-emergent and off-treatment definitions updated to add surgery and exclude radiotherapy in section 4.2.11.2.1	Yes	Clarification
Data presentation	23 May 2023	Section 5 (interim analysis) updated for futility criteria	Yes	Threshold changed for futility criteria
Data presentation	23 May 2023	- Appendix A Table 7 and Table 11 updated for Not Evaluable visit response definition - Nadir definition and new lesion clarified - Irradiated lesions/lesion intervention section added in Appendix A	Yes	Clarification
N/A	02 Jun 2023	Approved SAP v2.0	Yes, version 4	N/A
Data presentation	20 Sep 2023	PD-L1 values to be presented as continuous values when ranges are collected	Yes	New update
Data presentation	23 Nov 2023	Analysis population (Section 3.2 Updated. CCI [REDACTED], [REDACTED], exploratory mITT added for Final efficacy analyses, and PK and immunogenicity analysis sets definitions updated)	Yes	New update
Data presentation	23 Nov 2023	Section 3.3.7 Sample size and Section 5 Interim Analysis updated to accommodate changes of LRV and TV in the GNG decision framework based on emerging data	Yes	Clarification
Secondary endpoint(s)	23 Nov 2023	Section 4.2.4.6 Secondary Endpoint: ORR Definition updated to add details about BOR and best response of SD	Yes	New update

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Secondary endpoint(s)	23 Nov 2023	Section 4.2.6.1 Secondary Endpoint: the Clavien-Dindo Classification of surgical complications will be presented as supplementary analyses	Yes	New update
Data presentation	23 Nov 2023	Naming for MEDI5752 used in previous SAP versions has been updated to volrustomig	Yes	To align with CSP updates
Data presentation	23 Nov 2023	Database lock (DBL) updated to clinical data lock (CDL) throughout the document	Yes	To align with new AZ SOP
Derivation of secondary endpoint(s)	16 Jan 2024	Section 4.2.11.1.1: Dose intensity section updated to add definitions of actual cumulative dose, intended cumulative dose, and planned dose at each cycle for each study drug	Yes	Clarification
Secondary endpoint(s)	16 Jan 2024	Sections 4.2.1.6 and 4.2.5.6: a Waterfall plot will be presented for pCR and mPR as supplementary analyses	Yes	New update
Data presentation	16 Jan 2024	Section 4.2.11.2.2: Data collected for Interstitial Lung Disease / pneumonitis investigation will be listed as part of Dato-DXd AESI	Yes	New update
Secondary endpoint(s)	16 Jan 2024	Sections 4.2.11.4.1 and 4.2.11.4.2: Urinalysis data presentation has been updated.	Yes	New update
Secondary endpoint(s)	06 Feb 2024	Sections 4.2.2.2 and 4.2.3.2: Removed the word “evaluable” from the definitions of censoring rules for EFS and DFS	Yes	New update
N/A	14 Feb 2024	Approved SAP v3.0	Yes, version 5	N/A
Secondary endpoint(s)	11 Apr 2024	Section 4.2.10.2: New ADA categories and summary tables were added. Volrustomig-specific categories were added.	Yes	New update
Secondary endpoint(s)	17 Jun 2024	Subgroup analyses by stage (II, IIIA, IIIB) were added in sections 4.2.1.7, 4.2.2.7, 4.2.3.7, 4.2.5.7, 4.2.6.7 and 4.2.7.7	Yes	New update
Data presentation	17 Jun 2024	Section 3.2: ctDNA analysis population added	Yes	New update

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Data presentation	28 Jun 2024	Flexible language added for optional analysis	Yes	New update
Secondary endpoint(s)	02 Jul 2024	Supportive analyses using mITT population were added in sections 4.1.4.2 , 4.1.5.1 , 4.1.6.2 , 4.2.1.6 , 4.2.2.6 , 4.2.5.6 , 4.2.6.6 and 4.2.7.6	Yes	New update
Primary and secondary endpoint(s)	13 Sep 2024	CCI [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED]	Yes	New update
Secondary endpoint(s)	13 Sep 2024	Text added to clarify that new equivocal lesions must be converted to unequivocal at a follow-up visit to contribute to a PD	Yes	New update
Stratification factors	13 Sep 2024	Section 4.1.3.2: Stratification factors added for new treatment arms 6 and 7 (PD-L1 1-49% versus $\geq 50\%$)	Yes	New update
Secondary endpoint(s)	13 Sep 2024	Section 4.1.6.2: For treatment arms 6 and 7, local results of PD-L1 should be used. For all other treatment arms, central results of PD-L1 should be used when possible.	Yes	New update
Secondary endpoint(s)	30 Oct 2024	Section 4.2.8.6 : The association of ctDNA with clinical endpoints may be analyzed based ctDNA clearance status alone or in combination with pCR/mPR status.	Yes	New update
Secondary endpoint(s)	19 Sep 2024	Section 4.2.11.2.2 : List of AE summary tables amended to keep only the ones required as per the core list of tables, study-specific, or needed for the SRC. Flexible language used for optional tables.	Yes	New update

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Secondary endpoint(s)	13 Nov 2024	Section 4.2.11.2.2: AEs leading to dose omission of each IP and chemotherapy added to the overall summary table and to the summary by SOC and PT	Yes	New update
N/A	04 Dec 2024	Approved SAP v4.0	Yes, version 8	N/A
Data presentation	02 June 2025	Section 3.2: added definitions of ctDNA baseline evaluable, ctDNA clearance evaluable at C3D1 and pre-surgical timepoints	Yes	New update
Derivation of secondary endpoints(s)	02 June 2025	Section 4.2.2: to the definition EFS event added case of progression for patients who didn't undergo or completed surgery for reasons other than PD	Yes	New update
Secondary endpoint(s)	02 June 2025	Section 4.2.8: added text and figure to clarify the derivation of ctDNA clearance at specific timepoint	Yes	Clarification
Secondary endpoint(s)	02 June 2025	Section 4.2.8.4: in line with the updated derivation of ctDNA clearance, clarified analyses around association of ctDNA clearance and pathological response/EFS	Yes	Clarification
N/A	03 July 2025	Approved SAP v5.0	Yes, version 8	N/A
Data presentation	12 Sep 2025	Section 3.2: added definitions of PK and immunogenicity sets specific to each treatment	Yes	New update
Secondary endpoint(s)	16 Dec 2025	Section 4.2.10.2: updated presentation of immunogenicity data	Yes	New update
Secondary endpoint(s)	17 Dec 2025	Sections 3.2 and 4.2.6.4: Response evaluable set replaced with mITT for ORR and removed from the Analysis Populations.	Yes	New update

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Primary and secondary endpoint(s)	17 Dec 2025	Sections 4.2.1.4, 4.2.2.4, 4.2.5.4 and 4.2.7.4: mITT population used as primary analysis instead of ITT for pCR, EFS, mPR and OS. Section 4.2.6.4: mITT population used as primary analysis instead of Response evaluable set for ORR.	Yes	New update
N/A	26 Mar 2026	Approved SAP v6.0	Yes, version 8	N/A

1 INTRODUCTION

The purpose of this document is to give details for the statistical analysis of study D9077C00001 supporting the clinical study report (CSR). The reader is referred to the latest clinical study protocol (CSP) and the latest case report form (CRF) for details of objectives, study design, study conduct and data collection.

The study design schema is presented in Figure 1 of the CSP, and the Schedule of Activities (SoA) are presented in Table 1, Table 2, Table K18, Table K19, Table L20, Table L21, Table M22, Table M23, Table N26, Table N27, Table O29, Table O30, Table P31, Table P32, Table Q33 and Table Q34 of the CSP.

2 CHANGES TO PROTOCOL PLANNED ANALYSES

No changes from the planned analyses specified in the protocol.

3 DATA ANALYSIS CONSIDERATIONS

3.1 Timing of Analyses

There are interim analysis and final analysis planned for each arm separately.

An interim futility analysis primarily evaluating pCR will be performed once approximately 35 patients in each arm have received neoadjuvant treatment and received surgery or discontinued without receiving surgery. Refer to Section 5 for further details of the planned interim analyses.

The final analysis for each arm and the overall study will take place when all patients in the arm or the overall study complete 1 year of adjuvant therapy plus the follow-up period, or complete the end-of-study visit if withdrawing early from treatment, respectively. In the event that the study or certain treatment arms are terminated early by sponsor decision, the data cut-off may occur before every patient completes all the scheduled assessments. Primary and secondary endpoints including pathological complete response (pCR), objective response rate (ORR), event-free survival (EFS), disease-free survival (DFS), feasibility to surgery, major pathological response (mPR), and overall survival (OS) will be analysed during the final analysis.

Individual CSRs for each arm may be written after the interim clinical data lock (CDL) preceding the final analysis for each arm respectively. A final integrated CSR will be written after the final CDL containing all arms.

An early safety evaluation will be done by a safety review committee (SRC) to review all available data when the first **CC** patients in each treatment arm have completed **CC** cycles of

neoadjuvant treatment. An additional review of safety data by the SRC will occur when the first **CC** patients in each treatment arm have undergone surgery and have had **CC** days of follow-up, in order to assess perioperative mortality and surgery delays. For more details, refer to the SRC charter.

Additionally, throughout the study if more than **CC**% of patients experience a delay in surgery of more than 8 weeks or become unable to receive surgery due to toxicity on any treatment, enrolment in that treatment will be halted and SRC members will be notified to review these cases and provide their recommendation to the Sponsor.

The SRC will meet regularly at approximately 6-month intervals to review the safety and tolerability of treatment as neoadjuvant and adjuvant regimen, until all patients have had the opportunity to undergo surgery and those having surgery had at least 6 months of adjuvant treatment.

3.2 Analysis Populations

All patients who received any amount of study interventions are included in the Safety set. For the safety and PK analyses, patients are classified according to the actual treatment received. For all efficacy analyses, and for baseline and demography, patients are classified according to the treatment they were randomised to (i.e. the planned treatment).

The following populations are defined in [Table 1](#):

Table 1 Populations for Analysis

Population/Analysis set	Description	Endpoint/Output
Enrolled	All patients who sign the ICF	Disposition
ITT	All randomised patients in the study	Baseline and demography
modified ITT (mITT)	All randomised patients with confirmed NSCLC histology who received at least 1 dose of any study intervention	pCR, mPR EFS, EFS12 and EFS24 OS, OS12 and OS24 ORR & Subgroup / Supplementary analyses, unless otherwise specified
Resected	All the patients in the mITT who had surgical resection following the neoadjuvant period, who do not have R2 margins, and whose first scan following surgery shows no evaluable disease (defined as no post-surgery R2 margins and no RECIST evidence of disease based on the investigator data i.e., no target lesions, non-target lesions or new lesions (unless they are	DFS, DFS12 and DFS24 & Subgroup / Supplementary analyses

Population/Analysis set	Description	Endpoint/Output
	equivocal or pathologically confirmed to be a new primary malignancy))	
Exploratory mITT	All patients in the mITT who received surgery	pCR mPR
[REDACTED]	CCI [REDACTED] [REDACTED] [REDACTED] [REDACTED]	[REDACTED] [REDACTED] [REDACTED]
[REDACTED] [REDACTED]	[REDACTED] [REDACTED]	[REDACTED] [REDACTED]
Safety	All patients who received at least 1 dose of any study intervention	Exposure Adverse Events Laboratory evaluations Vital Signs ECG Feasibility to surgery
Pharmacokinetics (PK)	All patients who received at least 1 dose of any study intervention with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Durvalumab Pharmacokinetics (PK)	All patients who received at least 1 dose of durvalumab with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Oleclumab Pharmacokinetics (PK)	All patients who received at least 1 dose of oleclumab with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Monalizumab Pharmacokinetics (PK)	All patients who received at least 1 dose of monalizumab with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Volrustomig Pharmacokinetics (PK)	All patients who received at least 1 dose of volrustomig with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Dato-DXd Pharmacokinetics (PK)	All patients who received at least 1 dose of Dato-DXd with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
AZD0171 Pharmacokinetics (PK)	All patients who received at least 1 dose of AZD0171 with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations

Population/Analysis set	Description	Endpoint/Output
Rilvegostomig Pharmacokinetics (PK)	All patients who received at least 1 dose of rilvegostomig with at least 1 post-dose reportable PK concentration for that study intervention.	PK concentrations
Immunogenicity	All patients who received at least 1 dose of any study intervention with at least one non-missing ADA result at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Durvalumab Immunogenicity	All patients who received at least 1 dose of durvalumab with at least one non-missing ADA result of durvalumab at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Oleclumab Immunogenicity	All patients who received at least 1 dose of oleclumab with at least one non-missing ADA result of oleclumab at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Monalizumab Immunogenicity	All patients who received at least 1 dose of monalizumab with at least one non-missing ADA result of monalizumab at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Volrustomig Immunogenicity	All patients who received at least 1 dose of volrustomig with at least one non-missing ADA result of volrustomig at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Dato-DXd Immunogenicity	All patients who received at least 1 dose of Dato-DXd with at least one non-missing ADA result of Dato-DXd at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
AZD0171 Immunogenicity	All patients who received at least 1 dose of AZD0171 with at least one non-missing ADA result of AZD0171 at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
Rilvegostomig Immunogenicity	All patients who received at least 1 dose of rilvegostomig with at least one non-missing ADA result of rilvegostomig at any time, either baseline or post-baseline, according to the actual treatment received	ADAs
ctDNA	All patients in the mITT who had at least one sample processed at any timepoint.	ctDNA
ctDNA baseline evaluable	All patients in ctDNA set with baseline sample evaluable (result either detected or non-detected)	ctDNA
ctDNA clearance evaluable at C3D1	All patients in the ctDNA evaluable set that are evaluable for clearance (achieved either clearance or non-clearance) at Cycle 3 Day 1	ctDNA

Population/Analysis set	Description	Endpoint/Output
ctDNA clearance evaluable at pre-surgical assessment	All patients in the ctDNA evaluable set that are evaluable for clearance (achieved either clearance or non-clearance) at pre-surgical assessment	ctDNA

ADA = anti-drug antibody ; ctDNA = Circulating tumour DNA; DFS = Disease free survival; DFS12 = Disease free survival at 12 months; DFS24 = Disease free survival at 24 months; ECG = electrocardiogram; ICF = informed consent form; ITT = intention-to-treat; mPR = major pathological response; ORR = objective response rate; OS = overall survival; OS12 = overall survival at 12 months; OS24 = overall survival at 24 months; pCR = pathological complete response, EFS = event free survival; EFS12 = event free survival at 12 months; EFS24 = event free survival at 24 months; PK = pharmacokinetics.

* Individual PK concentration and parameter data for any patients who are excluded from the descriptive summary tables, figures and/or inferential statistical analyses are included in the listings and are flagged with an appropriate footnote.

3.3 General Considerations

3.3.1 General Study Level Definitions

The general principles described below are followed throughout the study:

- Continuous endpoints will be summarised by the number of observations, mean, standard deviation (Std Dev), median, upper and lower quartiles (as applicable), minimum, and maximum. For data that requires log-transformation, geometric mean, coefficient of variation (CV), median, minimum and maximum will be presented. Categorical endpoints will be summarised by frequency counts and percentages for each category.
- If data are available for less than 3 patients, no summary statistics other than minimum, maximum and number of observations will be presented.
- Unless otherwise stated, percentages will be calculated out of the analysis set total (excluding efficacy and exposure) and for treatment group.
- For continuous data, descriptive summary statistics (mean, median, standard deviation, standard error, confidence intervals) will be rounded to 1 additional decimal place compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.
- Derived variables will be rounded to 1 more decimal place compared to the least number of decimal places among the raw data used for calculation, provided the scale of the data is not changing.
- For categorical data, percentages will be rounded to 1 decimal place.
- SAS® version 9.4 (as a minimum) will be used for all analyses.

- Baseline is the last non-missing value obtained prior to the first dose/administration of any study treatment and any information taken after first dose/administration of study treatment is regarded as post-baseline information. If two visits are equally eligible to assess patient status at baseline (eg, screening and baseline assessments both on the same date prior to first dose/administration with no washout or other intervention in the screening period), the average is taken as the baseline value with the exception of ctDNA. For ctDNA the Cycle 1 pre-dose sample will be regarded as baseline if multiple pre-treatment samples are collected and analysed. For non-numeric laboratory tests (i.e. some of the urinalysis parameters) where taking an average is not possible then the best value is taken as baseline as this is the most conservative. In the scenario where there are two assessments on Day 1 prior to first dose, one with time recorded and the other without time recorded, the one with time recorded is selected as baseline. Where safety data are summarised over time, study day is calculated in relation to date of first treatment. For assessments on the day of first dose where time is not captured, a nominal pre-dose indicator, if available, serves as sufficient evidence that the assessment occurred prior to first dose. Assessments on the day of the first dose where neither time nor a nominal pre-dose indicator are captured is considered prior to the first dose if such procedures are required by the protocol to be conducted before the first dose. If no value exists before the first dose/administration, then the baseline value is treated as missing. For the analysis which summarise values by visit will use different baseline values for visit windowing (see section [3.3.2](#))
- In all summaries, change from baseline endpoints will be calculated as the post-treatment value minus the value at baseline. The percentage change from baseline is calculated as $(\text{post-baseline value} - \text{baseline value}) / \text{baseline value} \times 100$. For any endpoint subjected to log transformation, the change from baseline calculated and summarised on the log scale are back-transformed and presented as a 'baseline scaled ratio' (BSR). Percentage change is then calculated as $(\text{BSR} - 1) \times 100$.]
- Unless stated otherwise, two-sided confidence intervals are produced at 95%.
- For percentiles of survival times based on the Kaplan-Meier method (eg, median survival), CI will be calculated using the default method available in the SAS LIFETEST procedure (ie, the Klein and Moeschberger extension of the Brookmeyer-Crowley method).
- For point-estimates of survival based on the Kaplan-Meier method (eg, for EFS and DFS), CI will be calculated using the default method available in the SAS LIFETEST procedure (ie, using Greenwood's estimate of standard error and a log-log transformation).
- For the purposes of summarising safety data assessed at visits, in addition to baseline data, only on treatment data are included in the summary tables. On

treatment data is defined as data after the first dose of investigation product (IP) and with assessment date up to and including the date of last IP (or surgery, whichever is later) + 90 days and prior to start of any subsequent cancer therapy (excluding radiotherapy).

- Missing safety data will generally not be imputed. However, safety assessment values of the form of “<x” (i.e, below the lower limit of quantification) or >x (ie, above the upper limit of quantification) will be imputed as “x” in the calculation of summary statistics but displayed as “<x” or “>x” in the listings.
- Global country situation referred to in this plan (if any) is related to COVID-19 analysis.
- PD-L1 source data will be used for subgroup analyses. For treatment arms 6 and 7, local results should be used. For all other treatment arms, central results will be used for patients who have central results, otherwise local results will be used. In absence of source data, stratification factors as recorded in the IRT will be used. PD-L1 cut offs other than 1% may be explored. If the original value is a range, then values will be rounded to the lowest integers for continuous summary, for example, 20-30 will be converted into 20, ≥ 1 and < 5 will be converted into 1, and > 50 will be converted into 50.

For any safety summaries by period, the following data will be included:

- **Neoadjuvant period:** Date of first dose of neoadjuvant study treatment until the date of surgery for patients who undergo the surgery; For patients who did not undergo surgery, this period is up to min (date of last dose of neoadjuvant treatment + 90 days, date of first dose of subsequent anti-cancer therapy (excluding radiotherapy), date of DCO). Note: For assessments recorded on the day of surgery, time will be used to determine if it's pre or post surgery, if time is not available it will be assumed to occur post surgery.
- **Post-surgery period:** Date of the day of surgery until min (day of first dose of study treatment post-surgery, date of DCO, date of first dose of subsequent anti-cancer therapy (excluding radiotherapy), date of surgery + 90 days).
- **Adjuvant period:** Date of first dose of study treatment post-surgery until min (last study treatment post-surgery + 90 days, date of DCO, date of first dose of subsequent anti-cancer therapy (excluding radiotherapy)).
- **Overall period:** First dose of study treatment until min (last dose of study treatment (or surgery, whichever is later) + 90 days, date of DCO, date of first dose of subsequent anti-cancer therapy (excluding radiotherapy)).

3.3.2 Visit Window

For safety, time windows are defined for any presentations that summarise values by visit. The following conventions apply:

- The time windows are exhaustive so that data recorded at any timepoint has the potential to be summarised. Inclusion within the time window are based on the actual date and not the intended date of the visit.
- All unscheduled visit data have the potential to be included in the summaries.
- The window for the visits following baseline are constructed in such a way that the upper limit of the interval falls half way between the two visits (the lower limit of the first post-baseline visit is Day 2). If an even number of days exists between two consecutive visits, then the upper limit is taken as the midpoint value minus 1 day.
- Time windowing will be done separately for each assessment based on the schedule of events specific to that assessment.
- Should Study Day be missing (due to partial or missing dates), then visit will be assigned to the nominal visit at which that assessment was recorded, and no windowing will be performed.
- Visit windowing will be conducted up to and including the end of treatment visit. That is, the end of treatment visit will be reassigned to a scheduled visit based on the study day the end of treatment visit occurred at.

Neoadjuvant period (21 day cycle), calculated using the day of the first dose of study interventions as Day 1. For assessments up to and including last neoadjuvant dose, the following windows are used:

- Week 0 (Neoadjuvant baseline), visit window (up to and including the date of first dose) Note: only assessments before the first dose of neoadjuvant treatment are considered valid for baseline, e.g., Neoadjuvant cycle 1 day 1 pre-dose or screening)
- Week 3; nominal day 22, visit window (latest of Day 2 or date of first dose+1) – 32
- Week 6; nominal day 43, visit window 33 – 53
- Week 9; nominal day 64, visit window 54 – 74

In addition, there will be a summary of patients who complete or discontinue treatment in the neoadjuvant phase, but do not enter the adjuvant phase. For this summary these patients only will have their post-treatment discontinuation follow-up assessments windowed to the 'Follow-up schedule' based on the last day of study medication. For compliance calculations, all data collected after a subject permanently discontinues study treatment will use these windows and neoadjuvant treatment discontinuation visits and adjuvant treatment discontinuation visits will be presented separately.

- Follow up Day 30, visit window Days 2 - 46
- Follow up Month 2, visit window Days 47 - 76
- Follow up Month 3, visit window Days 77 - 136

Adjuvant period (28 day cycle), calculated using the day of the first dose of durvalumab/oleclumab/monalizumab following surgery as Day 1, or the date of the first assessment post-surgery for subjects who have not started adjuvant treatment. Use windows regardless of treatment discontinuation:

- Week 0 (Adjuvant baseline), visit window: Date of surgery + 1 to day 1, Note: only assessments before the first dose of adjuvant treatment, or the first assessment post-surgery for subjects who have not started adjuvant treatment are considered valid for baseline e.g., Adjuvant cycle 1 day 1 pre-dose
- Week 4; nominal day 29, visit window 2 – 43
- Week 8; nominal day 57, visit window 44 – 71
- Week 12; nominal day 85, visit window 72 – 99
- Week 16; nominal day 113, visit window 100 – 127
- Week 20; nominal day 141, visit window 128 – 155
- Week 24; nominal day 169, visit window 156 – 183
- Week 28; nominal day 197, visit window 184 – 211
- Week 32; nominal day 225, visit window 212 – 239
- Week 36; nominal day 253, visit window 240 – 267
- Week 40; nominal day 281, visit window 268 – 295
- Week 44; nominal day 309, visit window 296 – 323

CCI
[Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]

- [Redacted]
[Redacted]
[Redacted]
[Redacted]
[Redacted]
- [Redacted]
- [Redacted]
- [Redacted]

- CCI [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

...

CCI [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

- [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

...

- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment are used (regardless of where it falls in an interval).
- Listings display all values contributing to a time point for a patient.
- For visit based summaries, if there is more than one value per patient within a time window then the closest value to the scheduled visit date are summarised, or the earlier, in the event the values are equidistant from the nominal visit date. The listings highlight the value for the patient that contributed to the summary table, wherever feasible. Note: in summaries of extreme values, all post baseline values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.

- For summaries at a patient level, all values are included, regardless of whether they appear in a corresponding visit-based summary, when deriving a patient level statistic such as a maximum.

3.3.3 Handling of Unscheduled Visits

See Section 3.3.2 for safety.

3.3.4 Multiplicity/Multiple Comparisons

Not Applicable.

3.3.5 Handling of Protocol Deviations in Study Analysis

Important protocol deviations (IPDs) may include, but are not limited to the following:

- Patients randomised but who did not receive study treatment (Deviation 1)
- Patients who deviate from any entry criteria per the CSP (Deviation 2)
- Baseline RECIST CT/MRI scan >28 days before start date of randomisation and/or baseline PET scan >42 days before start date of randomisation (Deviation 3)
- No baseline RECIST 1.1 assessment on or before date of randomisation (Deviation 4)
- Patients who received prohibited concomitant medications during study period. (Deviation 5). See CSP Appendix H1
- Patients who met study intervention discontinuation criteria but continued study intervention, and potentially had major impact to safety of patients according to clinical judgement (Deviation 6)
- RECIST scans performed outside of the scheduled window on more than 1 occasion (Deviation 7)
- Patients randomised who received their randomised study treatment at an incorrect dose or received an alternate study treatment to that which they were randomised (Deviation 8)
- Missed visits, assessments, or treatments that, in the opinion of the investigator were due to the COVID-19 global pandemic and where there was a significant effect on either completeness, accuracy and/or reliability of the patient data, or the patient's rights, safety or wellbeing (Deviation 9)

More details will be provided in protocol deviation assessment plan. None of the deviations lead to patients being excluded from any analysis populations described in the SAP, unless otherwise specified. If a deviation is serious enough to have a potential impact on the

primary analysis, sensitivity analyses may be performed. A list of all protocol deviations is reviewed and decisions regarding how to handle these deviations are documented by the study team physician, clinical pharmacology scientist and statistician prior to database lock.

3.3.6 Missing Dates

Generally, the imputation of dates is used to decide if an observation is treatment emergent for adverse events or concomitant medications. The imputed dates should not be used to calculate durations, where the results would be less accurate.

The following are the guidelines used when partial dates are detected in the study:

- For missing diagnostic dates (e.g., disease diagnosis), if day and/or month are missing use 01 and/or Jan. If year is missing, put the complete date to missing.
- For missing AE and concomitant medication start dates, the following is applied:
 - a. Missing day - impute the 1st of the month unless month is the same as month of the first dose of study drug then impute first dose date.
 - b. Missing day and month - impute 1st January unless year is the same as first dose date then impute first dose date.
 - c. Completely missing - impute first dose date unless the end date suggests it could have started prior to this in which case impute the 1st January of the same year as the end date.
 - d. Imputed start date should be no later than the end date.
- For missing AE and concomitant medication end dates, the following is applied:
 - a. Missing day - impute the last day of the month unless both the month and the year are the same as the last dose date or the primary analysis data cut-off date then impute the last dose date or the primary analysis data cut-off date.
 - b. Missing day and month - impute 31st December unless the year is the same as the last dose date or the primary analysis data cut-off date then impute the last dose date or the primary analysis data cut-off date. Flags are retained in the database indicating where any programmatic imputation has been applied, and in such cases, any durations would not be calculated.
 - c. Completely Missing – need to look at whether the AE/medication is still ongoing before imputing a date and also when it started in relation to study drug. If the ongoing flag is missing, then assume that AE is still present / medication is still being taken (i.e. do not impute a date). If the AE/medication has stopped and start date is prior to first dose date, then impute first dose date.

Or if it started on or after first dose date then impute a date that is after the last dose of study drug date.

- If a patient is known to have died where only a partial death date is available, then the date of death will be imputed according to the rules for imputing AE start dates unless this date is before the last date the patient is known to be alive then the date of death will be imputed as the date the patient was last known to be alive +1.

If death has been recorded but the date is entirely missing, then date of death will be imputed as the date the patient was last known to be alive +1.

3.3.7 Sample Size

The sample size of up to 70 patients per treatment arm is to:

- Ensure the accuracy of making a Go or No-Go decision at the interim analysis and final analysis.
- Obtain a preliminary assessment of anti-tumour activity with a certain degree of precision.

The sample size is not based on type I error and power considerations. Instead, it is based on the model-based drug development approach (Lalonde et al., 2007; Frewer et al., 2016). In this decision framework, the maximum false-go risk is set to be CCI while maximum false-stop risk is CCI%. The sample size of up to 70 is calculated such that there is at least an CCI% chance of making a go decision at the final analysis if the true pCR rate is at target value for the respective treatment arm. Single-arm analyses are conducted. Meanwhile, the ultimate Go No-Go decision will be based on the totality of the data including, but not limited to, efficacy, safety, and biomarker data. The estimated pCR rate and its two-sided CCI% exact CIs can be found in Table 2.

CCI				

CC				

CC plan to enroll CC patients per arm to evaluate tolerability of alternate dose regimens, with a potential to increase to CC patients per arm to further understand safety and efficacy. Arm 4 may enroll CC patients if it is sufficient to understand safety and efficacy. Arm 6 may enroll CC patients if it is sufficient to understand safety and preliminary efficacy.

4 STATISTICAL ANALYSIS

This section provides information on definitions, derivation and analysis/data presentation per domain.

4.1 Study Population

The domain study population covers patient disposition, analysis sets, protocol deviations, demographics, baseline characteristics medical history, prior and concomitant medication and study drug compliance.

4.1.1 Patient Disposition and Completion Status

4.1.1.1 Definitions and Derivations

Study participation (i.e., a patient is “enrolled”) and randomisation is defined in CSP Section 5. Completion of study is defined in CSP Section 4.4.

4.1.1.2 Presentation

Patient disposition including screen failures and reason for screen failure is summarised and listed based on all patients enrolled (i.e., informed consent signed) by treatment and for all patients combined as defined by the current relevant tables, figures, and listings (TFL) standards. The number and percentage of patients for the following will be summarised:

- Patients enrolled;
- Patients who screen failed;
- Reason for screen failures;
- Patients randomised;
- Patients who received treatment;
- Patients who were randomised but did not receive treatment;
- Patients who discontinued treatment in the neoadjuvant period, i.e. discontinued any part of the study intervention;

- Patients who discontinued treatment in the adjuvant period, i.e. discontinued any part of the study intervention;
- Reasons for treatment discontinuation for both neoadjuvant and adjuvant period (including the reason due to global/country situation) – presented separately for each study intervention;
- Patients who completed or discontinued the neoadjuvant period but did not undergo surgery;
- Reasons for surgery not done;
- Patients who completed or discontinued the neoadjuvant period and received surgery;
- Patients who did not enter the adjuvant period;
- Patients completing surgery but not continuing to adjuvant period;
- Reasons for adjuvant treatment not given;
- Patients who had delay in receiving adjuvant treatment;
- Reasons for delay in adjuvant treatment;
- Patients ongoing study treatment at data cut-off (DCO) date;
- Patients ongoing study at DCO date;
- Patients who terminated study;
- Reasons for study termination.

The number of patients by region, country and centre and disruptions due to global/country situation (COVID-19) may also be summarised. The number of patients with confirmed/suspected COVID-19, and number of patients with confirmed/suspected COVID-19 who died may also be presented for the ITT analysis set.

Listings will be presented for disposition details for discontinued patients and patients ongoing in the study.

Listings may be presented for:

- Disposition due to global/country situation
- Patients with reported issues in the Clinical Trial Management System due to the global/country situation

4.1.2 Analysis Sets

4.1.2.1 Definitions and Derivations

For the definitions of each analysis set, refer to Section [3.2](#).

4.1.2.2 Presentation

The analysis sets are summarised by treatment group and for all patients combined. Any exclusions from analysis sets are listed.

4.1.3 Protocol Deviations

4.1.3.1 Definitions and Derivations

The list of the categories is presented in Section [3.3.5](#).

4.1.3.2 Presentation

The incidence of IPDs is summarised for the ITT analysis set by deviation categories. The number and percentage of patients with at least 1 important protocol deviation are summarised. A listing is provided with the important protocol deviation details.

Errors in stratifications (based upon stratification information recorded in the IRT and source data) may also be summarised separately to the important protocol deviations in order to show any discrepancies. Patients will be stratified by baseline PD-L1 expression status (1-49% versus $\geq 50\%$ for treatment arms 6 and 7, and $< 1\%$ versus $\geq 1\%$ for all other treatment arms).

4.1.4 Demographics

4.1.4.1 Definitions and Derivations

Age (years) will be grouped accordingly: $\geq 18 - < 50$, $\geq 50 - < 65$, $\geq 65 - < 75$ and ≥ 75 . Each race category counts patients who selected only that category.

4.1.4.2 Presentation

Demographics will be summarised and listed based on the ITT analysis set by treatment group and for all patients combined as defined by the current relevant TFL standards. The following will be summarised:

- age (years)
- age group
- sex

- race
- ethnicity
- country.

A supportive analysis of demographics may be based on the mITT.

Demographic characteristics in patients with confirmed/suspected COVID-19 infection may also be summarised and listed similarly.

4.1.5 Baseline Characteristics

4.1.5.1 Definitions and Derivations

Baseline characteristics are listed and summarised for the ITT analysis set population by treatment group and for all patients combined as defined by the current relevant TFL standards. Body mass index (BMI) (kg/m^2) will be calculated as: $\text{weight}/(\text{height}^2)$. Weight (kg) will be grouped accordingly: <50 , $\geq 50 - < 70$, $\geq 70 - < 90$ and ≥ 90 , and BMI will be grouped accordingly: <18.5 , $\geq 18.5 - < 25.0$, $\geq 25.0 - < 30.0$ and ≥ 30.0 . The latest assessment prior to randomization should be used as baseline ECOG score for patients who did not receive treatment.

The following will be summarised by treatment group using the ITT analysis set:

- weight (kg)
- weight group
- BMI (kg/m^2)
- BMI group
- Smoking status (current, former, never)
- Number of pack-years
- Eastern Cooperative Oncology Group (ECOG) performance status.

A supportive analysis of baseline characteristics may be based on the mITT.

4.1.6 Disease Characteristics

4.1.6.1 Definitions and Derivations

Not applicable

4.1.6.2 Presentation

Disease characteristics at baseline are listed and summarised for the ITT analysis set population by treatment group and for all patients combined as defined by the current relevant TFL standards.

Summaries are produced that present the number and percentage of patients on their:

- Histology type
- American Joint Committee on Cancer (AJCC) stage at diagnosis
- Regional lymph nodes
- Stage at study entry
- Baseline PD-L1 (1-49% and $\geq 50\%$ for treatment arms 6 and 7, and $<1\%$, $\geq 1\%$, 1-49% and $\geq 50\%$ for all other treatment arms)
- Baseline ctDNA (detected, undetected and unevaluable).

Both categorical and continuous data of baseline PD-L1 and ctDNA will be summarised. For treatment arms 6 and 7, local results of PD-L1 should be used. At the end of the study, if retrospective PD-L1 central results are available, a corresponding table based on central results will be presented. For all other treatment arms, central results of PD-L1 should be used for patients who have central results, otherwise local results will be used. In the absence of source data, stratification factors as recorded in the IRT will be used for categorical summaries.

Summary statistics will also be presented for patients' time from original diagnosis to randomisation and sum of target lesions (longest diameter [or short axis for lymph nodes], mm).

A supportive analysis of disease characteristics may be based on the mITT.

4.1.7 Medical History and Concomitant Disease

4.1.7.1 Definitions and Derivations

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

4.1.7.2 Presentation

Medical history and concomitant disease are summarised and may be listed for the ITT analysis set by treatment group, system organ class (SOC) and preferred term (PT). Medical

history for patients with confirmed/suspected COVID-19 infection may be summarised and listed similarly.

4.1.8 Prior and Concomitant Medications

4.1.8.1 Definitions and Derivations

For the purpose of inclusion in prior and/or concomitant medication or therapy summaries, incomplete medication or radiotherapy start and stop dates are imputed as detailed in Section 3.3.6.

Prior medications, concomitant and post-treatment medications are defined based on imputed start and stop dates as follows:

- Prior medications are those taken prior to study treatment with a stop date prior to the first dose of study treatment.
- Concomitant medications are those with a stop date on or after the first dose date of study treatment and must have started prior to or during treatment so there is at least one day in common with the study treatment.
- Post-treatment medications are those with a start date after the last dose date or last study treatment (including surgery).

4.1.8.2 Presentation

The number and percentage of patients who took prior and concomitant medications will be summarised and may be listed by Anatomical therapeutic chemical (ATC) classification (level 4) and generic term coded by World Health Organisation (WHO) Drug Dictionary.

4.1.9 Study Drug Compliance

4.1.9.1 Definitions and Derivations

Not applicable.

4.1.9.2 Presentation

Not applicable.

4.2 Endpoint Analyses

This section covers details related to the endpoint analyses such as primary, secondary, other endpoints including sensitivity and supportive analyses.

Efficacy assessments such as pCR and mPR will be performed and determined by a central blinded independent pathological review (BIPR), while other efficacy assessments (EFS, DFS, ORR, and OS) are conducted by Investigator. The EFS event of PD discovered during surgery will be determined by the investigator.

Efficacy analyses of EFS, DFS and ORR are based on programmatic application of Response Evaluation Criteria in Solid Tumours (RECIST) 1.1 ([Eisenhauer et al., 2009](#)) to investigator assessed tumour measurements. Programmatic derivation guidance used for the application of RECIST 1.1 are provided in [Appendix A](#), which is used to determine disease response.

RECIST data, including visit-level overall response is listed. Additional listing(s) are produced for pCR and mPR data.

All efficacy analyses are presented by planned treatment group.

4.2.1 Primary Endpoint: pCR

The primary efficacy endpoint is pCR.

4.2.1.1 Definition (pCR)

Pathological complete response is defined as the proportion of patients who have 0% residual viable tumour cells within all resected tissue (including primary lung lesion and lymph nodes) following neoadjuvant treatment as assessed by central BIPR laboratory. Central pathology assessment of pCR will be performed according to the recommended methods and definitions described by IASLC 2020 ([Travis et al., 2020](#)). Samples may also be evaluated by a different method ([Cottrell et al., 2018](#)) if deemed appropriate by the Sponsor.

4.2.1.2 Derivations (pCR)

The measure of interest is the proportion of patients with 0% residual viable tumour cells within all resected tissue as assessed by the central blinded pathologist.

4.2.1.3 Handling of Dropouts and Missing Data (pCR)

Patients who are not evaluable per central pathology assessment (this includes patients with R2 margins) or who do not have a surgical specimen will be considered as non-pCR (eg, pathology assessments captured as “non-evaluable” or “missing”, as appropriate).

4.2.1.4 Primary Analysis of Primary Endpoint (pCR)

Pathological complete response will be presented by the number and percentage of patients with pCR including 95% and 80% Clopper-Pearson (exact) ([Clopper-Pearson et al., 1934](#)) CIs for each arm using the mITT analysis set. CCI

[REDACTED]

4.2.1.5 Sensitivity Analyses of the Primary Endpoint (pCR)

Sensitivity analysis may be performed using the method of Cottrell mentioned in section [4.2.1.1](#).

4.2.1.6 Supplementary Analyses of the Primary Endpoint (pCR)

A waterfall plot of pathological regression in primary tumour based on independent pathological review may be presented.

4.2.1.7 Subgroup Analyses (pCR)

The pCR data may be analysed by treatment group using the exploratory mITT analysis set.

The pCR data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours [including adenocarcinoma and other histology]), and stage (II, IIIA and IIIB). Baseline PD-L1 expression <1% will not be applicable to arms 6 and 7.

4.2.2 Secondary Endpoint: EFS

EFS is a secondary efficacy endpoint.

4.2.2.1 Definition (EFS)

Event-free survival is defined as the time from randomisation to the first of the following: a) documented local or distant recurrence after completed surgery date, as determined by investigator using RECIST 1.1 assessment; b) death due to any cause (event date is the date of death); c) PD that precludes surgery as determined by Investigator (event date is date of this determination) or PD discovered and reported by the Investigator upon attempting surgery that prevents completion of surgery (event date is the date of the first attempt at surgery); d) date of the documented RECIST 1.1 progression that occurred after the surgery eligibility decision date (for participants who didn't undergo surgery for reason other than PD) or after the attempted surgery date (for participants who attempted, but didn't complete the surgery for reasons other than PD).

Patients with R1 and/or R2 margins are not considered to represent PD and will remain in EFS follow-up, although, in the case of R2, they will not be eligible for adjuvant therapy and should be treated in accordance with local standard of care. Pathological confirmation from biopsied lesions, if performed according to Investigator's judgment and local practice, will also be taken into consideration (as applicable). All new lesions should be biopsied where possible. A new primary malignancy confirmed by pathology is not considered an EFS event, but if not confirmed by pathology, a new equivocal lesion will need to be confirmed at a following visit to be converted to a new

unequivocal lesion and thus be considered as an EFS event. The date of EFS event will then be attributed to the first visit with new equivocal lesion (later converted to unequivocal).

4.2.2.2 Derivations (EFS)

$$\text{EFS (months)} = (\text{date of EFS event (progression/death) or censoring} - \text{date of randomisation} + 1) / (365.25/12)$$

Patients who have not experienced an event or died at the time of analysis are censored at the time of the latest date of assessment from their last disease assessment. For the purpose of EFS, the date of determination of whether or not to proceed to surgery as well as a decision to continue surgery are also considered disease assessments. However, if the patient has an event or dies immediately after two or more consecutive missed visits, the patient is censored at the time of the latest disease assessment prior to the two missed visits. Note: a NE visit is not considered as a missed visit.

If a patient has no disease assessments post-baseline, or does not have any baseline data, they will be censored at baseline, unless they die or progress (either RECIST 1.1 progression or a progression which precludes surgery, or with progression as the reason for surgery not being completed) in the period prior to surgery (Expected date of surgery is 40 days after the last neoadjuvant dose, or day 104 for subjects who did not receive neoadjuvant treatment), in which case they will have an EFS event on the date of death/progression.

Definition of two missed RECIST visits after surgery (calculated from the date of surgery) are defined as follows:

For subjects without a surgery date, an expected surgery date will be imputed as 40 days after the last neoadjuvant dose or day 104 for subjects who did not receive neoadjuvant treatment. For subjects with surgery performed outside the study, the surgery date will be used where available, otherwise the expected surgery date will be imputed as detailed above.

Given the scheduled visit assessment scheme (i.e. 12-weekly for the first 48 weeks after surgery/pre-planned surgery for no surgery performed, 24-weekly until week 192 and then 48-weekly thereafter) the definition of 2 missed visits changes.

For example, if the previous RECIST assessment is from day 50 after surgery to day 245 (i.e. week 7 to week 35) then two missing visits equates to 26 weeks since the previous RECIST assessment, allowing for early and late visits (i.e. $2 \times 12 \text{ weeks} + 1 \text{ week for an early assessment} + 1 \text{ week for a late assessment} = 26 \text{ weeks}$).

If the two missed visits occur over the period when the scheduled frequency of RECIST assessments changes from 12-weekly to 24-weekly this equates to 39 weeks (i.e. take the average of 12 and 24 weeks which gives 18 weeks and then apply same rationale, hence 2 x 18 weeks + 1 week for an early assessment + 2 week for a late assessment = 39 weeks). The time period for the previous RECIST assessment is from days 245 to 329 after surgery (i.e. week 35 to week 47).

Table 3 Two missed visits rule for EFS for schedule changes

Scheduled Assessment	Previous RECIST assessment	Two missed RECIST visits window
No disease assessment after surgery, or assessment prior to Day 21	(<Day 21)	13 weeks (91 days) after surgery*
Post-surgery assessment	Week 5±2 weeks (Day 21 – 49)	22 Weeks (154 days)
Q12W±1 week up to Week 48	Week 7 – Week 35 after surgery (>Day 49 – Day 245)	2 x 12 weeks + 2 weeks = 26 weeks (182 days)
	Week 35 – Week 47 (>Day 245 – Day 329) (change period from Q12W to Q24W)	2 x [(12 weeks+24 weeks)/2] + 1 week for an early assessment +2 week for a late assessment= 39 weeks (273 days)
Q24W±2 weeks up to week 192	Week 47 – Week 70 (Day 330 – 490)	2 x 24 weeks + 1 week for an early assessment +2 week for a late assessment = 51 weeks (357 days)
	Week 70 – Week 166 (Day 491 – 1162)	2 x 24 weeks + 4 weeks = 52 weeks (364 days)
	Week 166 – Week 190 (Day 1163 – 1330) (change period from Q24W to Q48W)	2 x [(24 weeks+48 weeks)/2] + 2 weeks for an early assessment +2 weeks for a late assessment= 76 weeks (532 days)
Q48W±2 weeks thereafter	>Day 1330 onwards	2 x 48 weeks + 4 weeks=100 weeks (700 days)

* Window, in this case only, is measured from date (or planned date) of surgery, as this is the first scan post-surgery.

A summary of censoring rules and the date of PD/death or censoring after surgical resection are given in [Table 4](#). Note that censoring overrides event in certain specified cases.

Table 4 **Summary of Censoring Rules for EFS after Surgical Resection/Pre-planned surgery**

Situation	Date of PD/Death or Censoring	EFS Outcome
Progressive disease (PD) or death in the absence of progression	Date of earliest documentation of PD or date of death in the absence of progression	Event
No tumour assessment after surgery/pre-planned surgery AND death prior to second scheduled post-surgery disease assessment	Date of death	Event
No tumour assessment after surgery/pre-planned surgery AND no death prior to second scheduled post-surgery disease assessment	Date of surgery/last disease assessment for no surgery.	Censored
PD or death (in the absence of progression) immediately after ≥ 2 consecutive missed disease assessments as per the protocol specified assessment schedule	Last event-free disease assessment prior to missed assessments	Censored
On-going with neither PD nor death at the time of analysis or lost to follow-up or withdrawn consent	Date of last disease assessment	Censored

PD = progressive disease; EFS = Event-free survival

The EFS time for RECIST scans is always derived based on scan/assessment dates, not visit dates.

RECIST assessments/scans contributing towards a particular visit may be performed on different dates. The following rules are applied:

- The date of progression is determined based on the earliest of the dates of the component that triggered the progression.
- When censoring a patient for EFS the patient is censored at the latest of the dates contributing to a particular overall visit assessment.

Note: for TLs only the latest scan date is recorded out of all scans performed at that assessment for the TLs and similarly for NTLs only the latest scan date is recorded out of all scans performed at that assessment for the NTLs.

The proportion of patients alive and event free at 12, 18 and 24 months (AEF-12/18/24 [Alive and Event Free] or EFS-12/18/24) is defined as the Kaplan-Meier estimate of EFS (per RECIST 1.1 as assessed by the site investigator) at 12/18/24 months.

The duration of follow-up is defined from randomisation to date of censoring (date last known to have not progressed) in censored (not progressed) patients only.

4.2.2.3 Handling of Dropouts and Missing Data (EFS)

Please see section [4.2.2.2](#).

4.2.2.4 Primary Analysis of Secondary Endpoint (EFS)

The main analysis of EFS is based on the mITT. The number and percentage of patients experiencing a EFS event broken down by type of event/censoring and Kaplan-Meier plots of EFS are presented by treatment group. The median EFS and its two-sided 95% CI are estimated using the Kaplan-Meier method.

The treatment status at EFS event of patients at the time of analysis are summarised. This includes the number (%) of patients who were on treatment at the time of event, the number (%) of patients who discontinued study treatment prior to event, the number (%) of patients who have no event and were on treatment or discontinued treatment.

A summary of the duration of follow-up for EFS is included using median (range). This is presented for censored patients (including all types of EFS censoring).

The proportion of patients alive and event free at 12, 18 and 24 months (landmark analysis) and associated two-sided 95% CI are estimated using the Kaplan-Meier method.

4.2.2.5 Sensitivity Analyses of the Secondary Endpoint (EFS)

Sensitivity analysis may be performed by censoring any patient who had an EFS event due to death where: a) primary/secondary cause of death was due to Confirmed/Suspected COVID-19 Infection, or b) a Confirmed/Suspected COVID-19 infection reported as a fatal AE, at their last assessment prior to their Confirmed/Suspected COVID infection death date.

4.2.2.6 Supplementary Analyses of the Secondary Endpoint (EFS)

Supplementary analysis of EFS data may be performed by summarising the results of R0, R1 and R2 resection separately.

4.2.2.7 Subgroup Analyses (EFS)

The EFS data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours), and stage (II, IIIA and IIIB).

4.2.3 Secondary Endpoint: DFS

DFS is a secondary endpoint.

4.2.3.1 Definition (DFS)

Disease-free survival will only be evaluated for patients who had surgical resection following the neoadjuvant period, who have R0/R1 margins, and whose first post-surgical RECIST 1.1 scan shows no evaluable disease (defined as no post-surgery R2 margins and no RECIST evidence of disease). Disease-free survival is defined as the time from the date of surgery until the first date of disease recurrence as determined by Investigator using RECIST 1.1 assessments (local or distant), or date of death due to any cause, whichever occurs first. Pathological confirmation from biopsied lesions, will also be taken into consideration (as applicable). A new primary malignancy, confirmed by pathology, is not considered a DFS event.

Patients who did not have an event at the time of analysis will be censored at the latest date of disease assessment after the surgical resection date.

However, if any of these events occur after 2 or more consecutive missed visits after the first post-surgery RECIST 1.1 scan then the patient will be censored at the time of latest disease assessment prior to the 2 missed visits. If a patient dies between the first post-surgical RECIST scan and the next scheduled RECIST scan (12 weeks relative to surgery date), this will still be considered as an event.

4.2.3.2 Derivations (DFS)

DFS (months) = (date of DFS event (recurrence/death) or censoring – date of surgery + 1) / (365.25/12)

For two missed visit rules please see section 4.2.2.2 after post-surgery. The censoring rules for DFS is defined in Table 5.

Table 5 Censoring rules used in the calculation of DFS

Situation	Date of PD/Death or Censoring	DFS outcome
Documented recurrence (local or metastatic) during the study	Date of the 1 st assessment of the series of tests that determined recurrence	Event ^a
Death during the study before recurrence	Date of death	Event
Death or recurrence immediately after two or more missed scheduled assessments	Date of last tumour assessment prior to missed visits as defined in section 4.2.2.2 (post-surgery)	Censored
Patients without recurrence as of data cut-off	Date of last tumour assessment	Censored

^a Earliest date among the dates used in calculating the disease-free survival.

4.2.3.3 Handling of Dropouts and Missing Data (DFS)

Please see section [4.2.3.2](#)

4.2.3.4 Primary Analysis of Secondary Endpoint (DFS)

The main analysis of DFS is based on the resected analysis set. The number and percentage of patients experiencing a DFS event broken down by type of event/censoring and Kaplan-Meier plots of DFS are presented by treatment group. The median DFS and its two-sided 95% CI are estimated using the Kaplan-Meier method.

The treatment status at DFS event of patients at the time of analysis are summarised. This includes the number (%) of patients who were on treatment at the time of event, the number (%) of patients who discontinued study treatment prior to event, the number (%) of patients who have no event and were on treatment or discontinued treatment.

A summary of the duration of follow-up for DFS is included using median (range). This is presented for censored patients (including all types of DFS censoring).

The proportion of patients alive and disease free at 12 and 24 months (landmark analysis) and associated two-sided 95% CI are estimated using the Kaplan-Meier method.

4.2.3.5 Sensitivity Analyses of the Secondary Endpoint (DFS)

Sensitivity analysis may be performed by censoring any patient who had a DFS event due to death where: a) primary/secondary cause of death was due to Confirmed/Suspected COVID-19 Infection, or b) a Confirmed/Suspected COVID-19 infection reported as a fatal AE, at their last assessment prior to their Confirmed/Suspected COVID infection death date.

4.2.3.6 Supplementary Analyses of the Secondary Endpoint (DFS)

Supplementary analysis of DFS data may be performed by summarising the results of R0 and R1 resection separately.

4.2.3.7 Subgroup Analyses (DFS)

The DFS data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours), and stage (II, IIIA and IIIB).

4.2.4 Secondary Endpoint: Feasibility to Surgery

Feasibility to surgery is a secondary endpoint.

4.2.4.1 Definition (Feasibility to Surgery)

Feasibility to surgery is defined as having the planned surgical resection within 40 days from the last dose of neoadjuvant study interventions administration. The measure of interest is the proportion of patients that have intended surgery within 40 days from the last neoadjuvant study interventions administered. Patients who do not intend to have surgery within 40 days of last dose of neoadjuvant study interventions will be considered as not feasible for surgery.

4.2.4.2 Derivations (Feasibility to Surgery)

See Section [4.2.4.1](#)

4.2.4.3 Handling of Dropouts and Missing Data (Feasibility to Surgery)

Please see section [4.2.4.1](#)

4.2.4.4 Primary Analysis of Secondary Endpoint (Feasibility to Surgery)

Summaries of number and percentage of patients having surgical resection within 40 days from the last neoadjuvant study interventions administered will be presented by treatment group using the safety set.

4.2.4.5 Sensitivity Analyses of the Secondary Endpoint (Feasibility to Surgery)

No sensitivity analysis is planned for feasibility to surgery.

4.2.4.6 Supplementary Analyses of the Secondary Endpoint (Feasibility to Surgery)

A summary of surgical details will be produced for the safety set and will include:

- Number of subjects who underwent surgery
- Number of subjects who completed surgery
- Frequency of surgery type/procedure performed
- Summary statistics for days from last IP administration (derived using date of last dose of IP and date of surgery)
- Frequency of R0, R1, R2
- Time from surgery to first dose of adjuvant treatment (days).

A summary of surgical complications will be presented for the safety set. Clavien-Dindo assessment will be utilised for grading post-operative complications (See CSP Appendix I for more details).

4.2.4.7 Subgroup Analyses (Feasibility to Surgery)

No subgroup analysis is planned for feasibility to surgery.

4.2.5 Secondary Endpoint: mPR

mPR is a secondary endpoint.

4.2.5.1 Definition (mPR)

Major pathological response rate is defined as the proportion of patients with 10% or less residual viable tumour tissue in lung primary tumour after neoadjuvant treatment at the time of resection as assessed per central BIPR laboratory. Central pathology assessment of mPR will be performed according to the recommended methods and definitions described by IASLC 2020 (Travis et al., 2020). Samples may also be evaluated by a different method (Cottrell et al., 2018) if deemed appropriate by the Sponsor. Estimated % residual viable tumour cells will be recorded. This will be used to programmatically assign a response of Yes if the value is 10% or less, or No if the value is above 10%. Patients who are not evaluable per central pathology assessment (including patients with R2 margins) or who do not have a surgical specimen will be considered as having non-mPR (eg, response captured as “non-evaluable” or “missing” as appropriate).

4.2.5.2 Derivations (mPR)

Please see section [4.2.5.1](#)

4.2.5.3 Handling of Dropouts and Missing Data (mPR)

Please see section [4.2.5.1](#)

4.2.5.4 Primary Analysis of Secondary Endpoint (mPR)

Major pathological response will be presented by the number and percentage of patients with mPR including 95% and 80% Clopper-Pearson (exact) CIs for each treatment using the mITT analysis set. CCI

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

4.2.5.5 Sensitivity Analyses of the Secondary Endpoint (mPR)

Sensitivity analysis may be performed using the method of Cottrell mentioned in section [4.2.1.1](#)

4.2.5.6 Supplementary Analyses of the Secondary Endpoint (mPR)

Supplementary analysis of mPR data may be performed by summarising the results of R0, R1 and R2 resection separately.

A waterfall plot of pathological regression in primary tumour based on independent pathological review may be presented.

4.2.5.7 Subgroup Analyses (mPR)

The mPR data may be analysed by treatment group using the exploratory mITT analysis set.

The mPR data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours), and stage (II, IIIA and IIIB).

4.2.6 Secondary Endpoint: ORR

ORR is a secondary endpoint.

4.2.6.1 Definition (ORR)

Objective response rate is defined as the proportion of patients achieving either CR or PR, prior to surgery, as assessed by the Investigator.

Objective Response (OR) is defined as a programmatically derived, best overall response of CR or PR that occurs prior to the initiation of subsequent anticancer treatment and prior to surgery. Objective Response Rate (ORR) is defined as the percentage of patients with objective response.

Best Overall Response (BOR) is calculated based on the overall visit responses from each RECIST assessment, described in [Appendix A](#). It is the best response a subject has had following randomisation up until surgery, or the last evaluable assessment up to and including RECIST progression in the absence of surgery, but prior to starting any subsequent cancer therapy (excluding radiotherapy).

Data obtained up until surgery, or last evaluable assessment in the absence of surgery will be included in the assessment of ORR, regardless of whether the patient withdraws from therapy. Patients who discontinue treatment without a response or surgery, receive subsequent anti-cancer therapy prior to surgery and then respond will not be included as responders in the ORR.

A confirmed response of CR/PR means that a response of CR/PR is recorded at 1 visit and confirmed by repeat imaging not less than 4 weeks after the visit when the response was first observed with no evidence of progression between the initial and CR/PR confirmation visit. For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment is used. SD should be recorded at least 6 weeks from randomisation.

In the case where a patient has two non-consecutive visit responses of PR, then, as long as the time between the 2 visits of PR is greater than 4 weeks and there is no PD between the PR visits, the patient is defined as a responder. Similarly, if a patient has visit responses of CR, NE, CR, then, as long as the time between the 2 visits of CR is greater than 4 weeks, then a best response of CR is assigned.

Both unconfirmed and confirmed ORR will be reported, considering there are only 2 post-baseline scheduled tumour assessments prior to surgery.

4.2.6.2 Derivations (ORR)

Please see section [4.2.6.1](#)

4.2.6.3 Handling of Dropouts and Missing Data (ORR)

Please see section [4.2.6.1](#)

4.2.6.4 Primary Analysis of Secondary Endpoint (ORR)

Summaries are produced that present the number and percentage of patients with a CR or PR as best overall tumour response. Confirmed and unconfirmed ORR will be calculated and presented as well. The ORR is presented with a two-sided 95% CI using the Clopper-Pearson (exact probability) method. CCI

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Patients that have missing overall response assessments at all visits are considered as non-responders, and are therefore counted in the denominator of ORR. The main analysis of ORR is based on the mITT.

4.2.6.5 Sensitivity Analyses of the Secondary Endpoint (ORR)

No sensitivity analysis is planned for ORR.

4.2.6.6 Supplementary Analyses of the Secondary Endpoint (ORR)

Supplementary analysis may be conducted by calculating the percentage (and its 95% CI) of patients who achieve either objective response (CR/PR) or pathological response (pCR/mPR), based on the mITT set.

Supplementary analysis of ORR data may be performed by summarising the results of R0, R1 and R2 resection separately.

4.2.6.7 Subgroup Analyses (ORR)

The ORR data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours), and stage (II, IIIA and IIIB).

4.2.7 Secondary Endpoint: OS

OS is a secondary endpoint.

4.2.7.1 Definition (OS)

Overall survival is defined as the time from the date of randomisation until death due to any cause regardless of whether the patient withdraws from study therapy or receives another anti-cancer therapy.

4.2.7.2 Derivations (OS)

OS (months) = (date of death or censoring – date of randomisation + 1) / (365.25/12)

Any patient not known to have died at the time of analysis is censored based on the last recorded date on which the patient was known to be alive.

Note: Survival calls are made in the week following the date of data cut-off (DCO) for the analysis, and if patients are confirmed to be alive or if the death date is post the DCO date these patients are censored at the date of DCO. The status of ongoing, withdrawn (from the study) and “lost to follow-up” patients at the time of the final OS analysis should be obtained by the site personnel by checking the patient’s notes, hospital records, contacting the patient’s general practitioner and checking publicly-available death registries. In the event that the patient has actively withdrawn consent to the processing of their personal data, the vital status of the patient can be obtained by site personnel from publicly available resources where it is possible to do so under applicable local laws.

Note: For any OS analysis performed (prior to the final OS analysis), in the absence of survival calls being made, it may be necessary to use all relevant CRF fields to determine the last recorded date on which the patient was known to be alive for those patients still on treatment (since the *SURVIVE* module is only completed for patients off treatment if a survival sweep is not performed). The last date for each individual patient is defined as the latest among the following dates recorded on the CRFs:

- AE start, stop and change in severity dates
- Admission and discharge dates of hospitalisation
- Study treatment date

- End of treatment date
- Concomitant medication start and stop dates
- Laboratory test dates
- Date of vital signs
- Disease assessment dates on RECIST CRF
- Start and stop dates of alternative anticancer treatment
- Date last known alive on survival status CRF
- End of study date.

The proportion of patients alive at 12/24 months (OS-12/24) is defined as the Kaplan-Meier estimate of OS at 12/24 months.

Duration of follow-up for OS of all patients is defined as follows:

Duration of follow-up for OS (months) = (date of death or censoring (date last known to be alive) – date of randomisation + 1) / (365.25/12).

4.2.7.3 Handling of Dropouts and Missing Data (OS)

See section [3.3.6](#).

4.2.7.4 Primary Analysis of Secondary Endpoint (OS)

The analysis of OS is based on the mITT analysis set. The number and percentage of patients experiencing an OS event and Kaplan-Meier plots of OS will be presented. The median OS and two-sided 95% CI are estimated using the Kaplan-Meier method.

Summaries on the number and percentage of patients who have died, those still in survival follow-up and those lost to follow-up and those who have withdrawn consent will be provided along with the median OS for each treatment.

The proportion of patients alive at 12 and 24 months (landmark analysis) and associated two-sided 95% CI are estimated using the Kaplan-Meier method.

A summary of the duration of follow-up for OS is included using median (range). This is presented separately for censored and non-censored patients.

4.2.7.5 Sensitivity Analyses of the Secondary Endpoint (OS)

Sensitivity analysis may be performed by censoring any patient who had a death where: a) primary/secondary cause of death was due to Confirmed/Suspected COVID-19 Infection, or b) a Confirmed/Suspected COVID-19 infection reported as a fatal AE, at their last assessment prior to their Confirmed/Suspected COVID infection death date.

4.2.7.6 Supplementary Analysis of the Secondary Endpoint (OS)

Supplementary analysis of OS data may be performed by summarising the results of R0, R1 and R2 resection separately.

4.2.7.7 Subgroup Analyses (OS)

The OS data may be analysed by treatment group and baseline ctDNA (detected, undetected and unevaluable), ctDNA clearance status, baseline PD-L1 expression (<1%, ≥1%, 1-49% and ≥50%), histology (squamous cells tumours, non-squamous cells tumours), and stage (II, IIIA and IIIB).

4.2.8 Secondary Endpoint: ctDNA

ctDNA clearance on-treatment prior to surgery is a secondary endpoint for this study.

4.2.8.1 Definition (ctDNA)

The outcome of a ctDNA assessment at any timepoint can be one of the following: detected (positive), undetected (negative), or unevaluable/missing result.

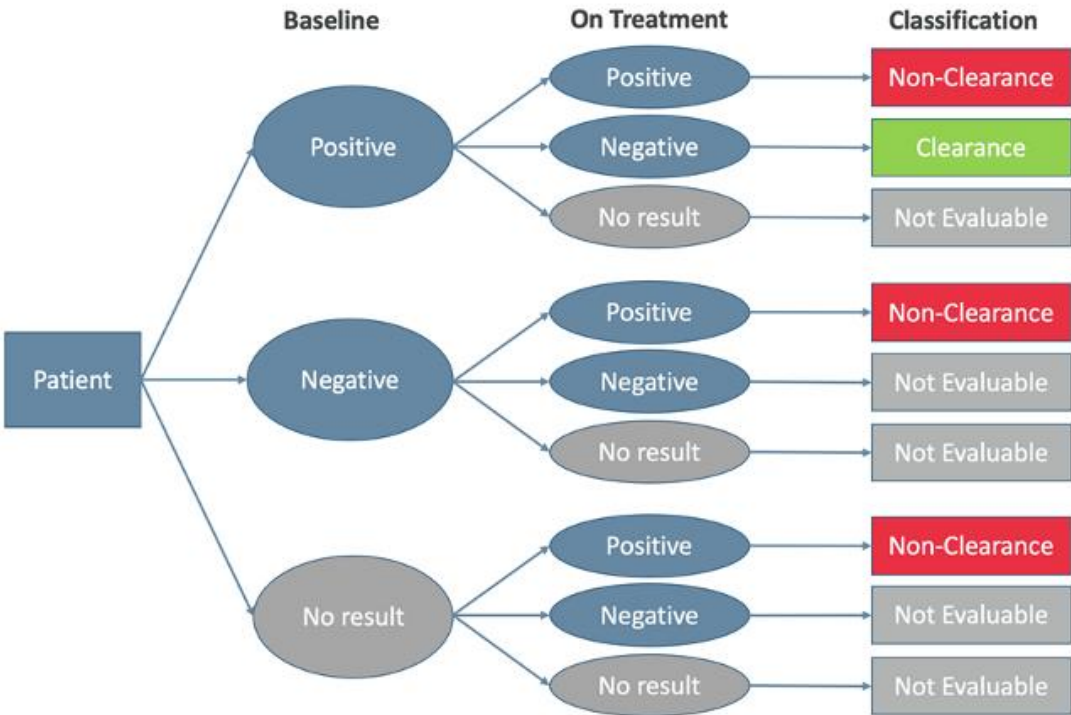
Complete molecular response or "ctDNA clearance" is defined as the change of ctDNA status from detected at baseline to undetected at post-baseline prior to surgery.

The ctDNA clearance rate at specified on-treatment timepoint (C3 or pre-surgical assessment) is defined as proportion of patients who achieved clearance among all patients evaluable for ctDNA clearance at that specific timepoint (i.e. who have as their result either clearance or non-clearance).

4.2.8.2 Derivations (ctDNA)

ctDNA clearance at specified on-treatment timepoint (C3 or pre-surgical assessment) is defined as positive ctDNA result at baseline followed by negative ctDNA result at that timepoint. ctDNA non-clearance is defined as positive ctDNA result at post-baseline assessment, regardless of result at baseline. Patients with negative ctDNA or missing result at baseline followed by negative result at post-baseline, as well as patients with missing result of ctDNA clearance for post-baseline timepoint are considered unevaluable for ctDNA clearance at that specific post-baseline timepoint. Please see [Figure 1](#) below for graphical representation of derivation rule:

Figure 1 Classification rules for ctDNA status



4.2.8.3 Handling of Dropouts and Missing Data (ctDNA)

No imputation will be performed for ctDNA.

4.2.8.4 Primary Analysis of Secondary Endpoint (ctDNA)

The number and percentages of patients with evaluable baseline assessment (detected or undetected) will be summarised across each treatment within the ctDNA baseline evaluable population, while number and percentage of patients who achieve ctDNA clearance at C3 or at pre-surgical assessment will be summarised across each treatment within respective ctDNA clearance evaluable populations. CCI

4.2.8.5 Sensitivity Analyses of the Secondary Endpoint (ctDNA)

No sensitivity analysis is planned for ctDNA.

4.2.8.6 Supplementary Analysis (ctDNA)

The association between ctDNA status at baseline/ctDNA clearance at post-baseline timepoints and clinical endpoints (pCR, mPR, EFS and DFS etc.) will also be assessed by cross-tabulations and Kaplan-Meier analysis for EFS and DFS if numbers allow. The

association with EFS may be analyzed based on ctDNA clearance status alone or in combination with pCR/mPR status (e.g. pCR+/ctDNA+ vs pCR+/ctDNA- vs pCR-/ctDNA+ vs pCR-/ctDNA-, and mPR+/ctDNA+ vs mPR-/ctDNA+ vs mPR-/ctDNA- vs mPR+/ctDNA-) provided big enough sample sizes.

The association between ctDNA clearance and histology, tumor size, smoking history and stage of disease may also be explored.

Supplementary analysis of ctDNA data may be performed by summarising the results of R0, R1 and R2 resection separately.

4.2.8.7 Subgroup Analyses (ctDNA)

No subgroup analysis is planned for ctDNA.

4.2.9 Secondary Endpoint: PK

4.2.9.1 Definition (PK)

This section covers details related to pharmacokinetics endpoints and analyses.

4.2.9.2 Derivations (PK)

The gmean is calculated as $\exp(\mu)$, where μ is the mean of the data on the natural log scale.

The gSD is calculated as $\exp(\sigma)$, where σ is the standard deviation of the data on the natural log scale.

The gCV is calculated as $100 \times \sqrt{\exp(s^2)-1}$, where s is the Std Dev of the data on the natural log scale.

The $\text{gmean} \pm \text{gSD}$ ($\text{gmean}-\text{gSD}$ and $\text{gmean}+\text{gSD}$) are calculated as $\exp[\mu \pm s]$.

4.2.9.3 Handling of Dropouts and Missing Data (PK)

Handling of Non-Quantifiable Concentrations

Individual concentrations below the LLOQ of the bioanalytical assay are reported as NQ in the listings with the LLOQ defined in the footnotes of the relevant TFLs. Individual plasma concentrations that are Not Reportable are reported as NR and those that are missing are reported as NS (No Sample) in the listings. Plasma concentrations that are NQ, NR or NS are handled as follows for the provision of descriptive statistics:

- Any values reported as NR or NS are excluded from the summary tables and corresponding figures.

- At a time point where less than or equal to 50% of the concentration values are NQ, all NQ values are set to the LLOQ, and all descriptive statistics are calculated accordingly.
- At a time point where more than 50% (but not all) of the values are NQ, the gmean, $\text{gmean} \pm \text{gSD}$ and gCV% are set to NC. The maximum value is reported from the individual data, and the minimum and median are set to NQ.
- If all concentrations are NQ at a time point, no descriptive statistics are calculated for that time point. The gmean, minimum, median and maximum are reported as NQ and the gCV% and $\text{gmean} \pm \text{gSD}$ as NC.
- The number of values below LLOQ ($n < \text{LLOQ}$) are reported for each time point together with the total number of collected values (n).

Three observations $> \text{LLOQ}$ are required as a minimum for a plasma concentration or PK parameter (e.g. C_{max}, C_{min}, C_{last}) to be summarised. Two observations $> \text{LLOQ}$ are presented as minimum and maximum with the other summary statistics as NC.

4.2.9.4 Primary Analysis of Secondary Endpoint (PK)

Precision and Rounding Rules for Pharmacokinetic Data

PK concentration data

PK concentration data listings present to the same number of significant figures as the data received from the bioanalytical laboratory (usually but not always to 3 significant figures) and against the same units as received.

PK concentration descriptive statistics present 4 significant figures with the exception of the min and max which present 3 significant figures and n and $n < \text{LLOQ}$ which present as integers.

For each analyte, serum concentrations for each scheduled time-point are summarised by analyte, cohort, Visit and scheduled timepoint using appropriate descriptive statistics.

The following descriptive statistics are presented for serum concentrations:

- n
- n below LLOQ
- geometric mean (gmean)
- geometric standard deviation (gSD)
- $\text{gmean} \pm \text{gSD}$
- geometric coefficient of variance (%) (gCV)
- arithmetic mean of non log-transformed data (mean)
- standard deviation of non log-transformed data (Std Dev)

- coefficient of variance (%) (CV)
- median
- minimum (min)
- maximum (max).

Protocol scheduled times are used to present the PK concentration summary tables and corresponding gmean concentration-time figures.

Graphical presentation of PK data

All mean (arithmetic mean and/or gmean) plots or combined plots showing all patients by treatment are based on the PK analysis set.

For consistency, the serum concentration values used in the mean (arithmetic mean and/or gmean) data graphs are those given in the descriptive statistics summary table for each time point.

For gmean concentration-time plots, NQ values are handled as described for the descriptive statistics; if the geometric mean is NQ, the value plotted is zero for linear plots and missing for semi-logarithmic plots. Any $\text{gmean} \pm \text{gSD}$ error bar values that are negative are truncated at zero on linear concentration-time plots and omitted from semi-logarithmic plots.

For individual plots, serum concentrations which are NQ prior to the first quantifiable concentration are set to a value of zero (linear plots only). After the first quantifiable concentration, any NQ plasma concentrations are regarded as missing.

4.2.9.5 Additional Analyses of the Secondary Endpoint (Pharmacokinetics)

All the analysis of PK are mentioned in Section [4.2.9.4](#).

4.2.9.6 Subgroup Analyses (PK)

No subgroup analysis is planned for PK data.

4.2.10 Secondary Endpoint: Immunogenicity

4.2.10.1 Definition (Immunogenicity)

Blood samples for determination of ADA for IP(s) in serum will be taken according to the SoA in the CSP and assayed by bioanalytical test sites operated by or on behalf of AstraZeneca, using an appropriately validated bioanalytical method. Full details of the methods used will be described in a separate report.

ADA samples may also be further tested for characterisation of the ADA response.

4.2.10.2 Presentations (Immunogenicity)

ADA table and listings will be produced according to all AZ corporate and/or Oncology TA standard TFL shells. Immunogenicity results will be listed for each patient and summarised for the immunogenicity analysis population. ADA results from each sample will be reported as either positive or negative. If the sample is positive, the ADA titer will be reported as well. Number and percentage of patients in the following categories will be provided.

- ADA positive post-baseline and positive at baseline.
- ADA not detected post-baseline and positive at baseline.
- Anti-drug antibody (ADA) positive at baseline and/or post-baseline visits. The percentage of these subjects in a population is known as ADA prevalence.
- Treatment-induced ADA positive (positive post-baseline and not detected at baseline).
- Treatment-boosted, defined as baseline positive ADA titre that was boosted to a 4-fold or higher level following drug administration.
- Treatment-emergent ADA positive (either treatment-induced ADA positive or treatment-boosted ADA positive). The percentage of these subjects in a population is known as ADA incidence.
- Non-treatment-emergent ADA positive, defined as being ADA positive but not fulfilling the conditions for treatment-emergent ADA positive.
- Treatment-emergent persistently positive ADA, defined as being TE ADA+ and having at least 2 post-baseline ADA-positive measurements with at least 16 weeks (112 days) between the first and last positive measurements.
- Treatment-emergent transiently positive ADA, defined as being TE ADA+ and having at least 1 post-baseline ADA-positive measurement and not fulfilling the conditions for persistently positive.
- Treatment-emergent ADA positive with maximum post-baseline titre $\geq$ median of maximum post-baseline titres (include for Volrustomig and Rilvegostomig only).

A summary will be provided of the number and percentage of patients who develop detectable anti-drug antibodies by ADA categories using the immunogenicity set. Descriptive statistics (minimum, Q1, median, Q3, and maximum) for the maximum post-baseline ADA titers of patients in each ADA category will also be included.

CCI

CCI

The impact of ADA on PK and safety may be evaluated, if the data allow. Spaghetti plots of individual concentration data over time profiles of subjects TE-ADA+, non-TE-ADA+, and ADA negative will be presented.

All valid assay results from patients who received any study drug are included in immunogenicity summaries. Blood samples collected outside of the protocolled window are summarised at the closest nominal time point that does not already have a value.

Immunogenicity results will be listed for all patients in safety set regardless of ADA-evaluable status. Anti-drug antibody titer data will be listed for samples confirmed positive for the presence of anti-drug antibodies. AEs in ADA positive patients by ADA positive category will be listed.

4.2.11 Safety Analyses

The domain safety covers exposure, adverse events, clinical laboratory, vital signs, and electrocardiogram (ECG).

Tables are provided for the safety set, listings are provided for all patients or the safety set depending on the availability of data.

4.2.11.1 Exposure

4.2.11.1.1 Definitions and Derivations

Total Exposure during Neoadjuvant period

Duration of exposure (weeks) will be defined separately for each study intervention and for neoadjuvant and adjuvant period as follows:

- Duration of exposure (weeks) of durvalumab in neoadjuvant period= (earliest of (last date of actual dosing of durvalumab where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of oleclumab in neoadjuvant period= (earliest of (last date of actual dosing of oleclumab where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of monalizumab in neoadjuvant period= (earliest of (last date of actual dosing of monalizumab where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7

- Duration of exposure (weeks) of chemotherapy in neoadjuvant period= (earliest of (last date of actual dosing of chemotherapy where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of volrustomig in neoadjuvant period= (earliest of (last date of actual dosing of volrustomig where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of Dato-DXd in neoadjuvant period= (earliest of (last date of actual dosing of Dato-DXd where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of AZD0171 in neoadjuvant period= (earliest of (last date of actual dosing of AZD0171 where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of rilvegostomig in neoadjuvant period= (earliest of (last date of actual dosing of rilvegostomig where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (cycles) is defined as the number of cycles in which at least one portion of investigational agent was administered (i.e. dose > 0 mg). If a cycle is prolonged due to toxicity, this should still be counted as one cycle.

Total Exposure during Adjuvant period

- Duration of exposure (weeks) of durvalumab in adjuvant period= (earliest of (last date of actual dosing of durvalumab where dose>0 mg in the last cycle+ 27, death, DCO)– first dose date+1)/7.
- Duration of exposure (weeks) of oleclumab in adjuvant period= (earliest of (last date of actual dosing of oleclumab where dose>0 mg in the last cycle+ 27, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of monalizumab in adjuvant period= (earliest of (last date of actual dosing of monalizumab where dose>0 mg in the last cycle+ 27, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of volrustomig in adjuvant period= (earliest of (last date of actual dosing of volrustomig where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of AZD0171 in adjuvant period= (earliest of (last date of actual dosing of AZD0171 where dose>0 mg in the last cycle+ 13, death, DCO)– first dose date+1)/7
- Duration of exposure (weeks) of rilvegostomig in adjuvant period= (earliest of (last date of actual dosing of rilvegostomig where dose>0 mg in the last cycle+ 20, death, DCO)– first dose date+1)/7
- Duration of exposure (cycles) is defined as the number of cycles in which at least one portion of investigational agent was administered (i.e. dose > 0 mg). If a cycle is prolonged due to toxicity, this should still be counted as one cycle.

Actual Exposure

Duration of actual exposure (which takes into account dose delay) of durvalumab/oleclumab/monalizumab/chemotherapy/volrustomig/Dato-DXd/AZD0171/rilvegostomig:

- Actual exposure = total exposure – total duration of dose delays, where total exposure will be calculated as above, and a dose delay is defined as any length of time (weeks) where the patient has not taken any of the planned dose according to the last dose date.

Neoadjuvant period

Calculation of duration of dose delays (for actual exposure) in the neoadjuvant period:

Since patients will receive durvalumab/oleclumab/monalizumab/volrustomig/Dato-DXd/AZD0171/rilvegostomig via IV infusion q3w, the duration of dose delays will be calculated as follows:

- For all dosing dates which are delayed per protocol: Total duration of dose delays = Sum of (Date of the dose – Date of previous dose – 21 days).

Adjuvant period

For patients receiving durvalumab/oleclumab/monalizumab via IV infusion q4w, the duration of dose delays will be calculated as follows:

- For all dosing dates which are delayed per protocol: Total duration of dose delays = Sum of (Date of the dose – Date of previous dose – 28 days).

CCI

[REDACTED]

[REDACTED]

■ [REDACTED]
[REDACTED]

[REDACTED]
[REDACTED]

■ [REDACTED]
[REDACTED]

Number of treatment cycles received

Exposure will also be measured by the number of cycles received. A cycle corresponds to a period of 21 days in the neoadjuvant period and 28 days in the adjuvant period except volrustomig and rilvegostomig for which a cycle corresponds to **cc** days. If a cycle is prolonged due to toxicity, this should still be counted as one cycle. A cycle will be counted if treatment is started even if the full dose is not delivered.

Dose Intensity

Dose intensity of IP(s) is addressed by considering relative dose intensity (RDI), where RDI is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation. More specifically, RDI is defined as follows:

- $RDI = 100\% \times d/D$, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the actual last day of dosing. D is the total dose that would be delivered, if there were no modification to dose or schedule.

Actual cumulative dose = sum over all study drug administration recorded on the study drug exposure form up to min (date of last dose date where dose > 0, date of death, date of DCO).

The percentage of volume left after the infusion will be used to calculate how much of the study drug they received, i.e.

- Volume left = Volume after infusion/Volume before infusion
- Actual dose received = (1 - Volume left) x Total dose

For example, if a patient has 100ml left after the infusion from 250ml before the infusion, then:

Volume left = $100/250=40\%$

Actual mg of study drug received= $(1 - 0.4) \times 1500 \text{ mg} = 900 \text{ mg}$

Intended cumulative dose is calculated by summing the planned individual doses that should have been received up to and including the last day of treatment according to the planned dose and schedule.

The planned dose at each cycle for each study drug will be calculated as:

- Durvalumab: 1500 mg IV Q3W (pre-surgery) or Q4W (post-surgery)
- Oleclumab: 3000 mg IV Q3W (pre-surgery) or Q4W (post-surgery)
- Monalizumab: 1500 mg IV Q3W (pre-surgery) or Q4W (post-surgery)

- CCI
 - CCI
 -
 -
 -
- Dato-DXd: 6.0 mg/kg x baseline weight
Weight at day 1 of each cycle will be used if the percentage change from baseline weight is $\geq \pm 10\%$. Otherwise, baseline weight will be used. After the recalculation, the updated patient's weight will be used as the new baseline weight. If the weight at day 1 of a cycle is not available, the last weight measurement prior to the day 1 of that cycle will be used.
- CCI
-

RDI will be calculated for each for the following:

- Neoadjuvant period - durvalumab/oleclumab/monalizumab/volrustomig/Dato-DXd/AZD0171/rilvegostomig
- Adjuvant period - durvalumab/oleclumab/monalizumab/volrustomig/AZD0171/rilvegostomig
- Overall - durvalumab/oleclumab/monalizumab/volrustomig /AZD0171/rilvegostomig, calculated as the average between the neoadjuvant and adjuvant RDI. Overall RDI for Dato-DXd is the same as the RDI for the neoadjuvant period because Dato-DXd is administered during the neoadjuvant period only.

Duration of safety follow-up

Duration of safety follow-up for all subjects included in the study:

- Total safety follow-up (months) = $[\min(\text{date of safety follow-up assessment, last dose date (or surgery date, whichever is later)} + 90, \text{date of subsequent cancer therapy (excluding radiotherapy), date of study withdrawal, date of death, DCO date}) - \text{first dose date} + 1] / (365.25/12)$.

4.2.11.1.2 Presentation

The summaries related to study treatment will be produced for the safety set by actual treatment group and will be repeated for each treatment period (neoadjuvant, adjuvant).

Duration of exposure to IP(s) in weeks (and cycles) is summarised by descriptive statistics and by frequency using the safety set for each treatment group. RDI is summarised by

descriptive statistics using the safety set for each treatment period (neoadjuvant, adjuvant and overall). Exposure to IP(s) i.e. total amount of study drug received is listed for all patients. Exposure swimmer plot(s) are produced, with a line presented for each patient to display relevant exposure and disposition details.

Dosing deviations for IP(s) will be summarised with reasons for deviations for the following categories: dose delays, omissions, interruptions and reductions (if applicable). Dosing delays are derived based on the scheduled dosing dates using the previous dose given as reference. The number of patients with dosing delays and total dose delays are summarised. In addition, the number of patients who had delay to the first adjuvant cycle will be summarised with reasons for delay.

4.2.11.2 Adverse Events

4.2.11.2.1 Definitions and Derivations

The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) is used to code the AEs. AEs are graded according to the National Cancer Institute Common Terminology Criteria for AEs (using the CTCAE version referenced in the CSP).

Treatment emergent adverse events (TEAEs) are all AEs which onset or worsen in severity following the first administration of IP within the duration of the treatment period, up to and including 90 days after the last dose of study treatment (or surgery, whichever is later) as per the study safety follow-up period but prior to subsequent cancer therapy (excluding radiotherapy). Worsening in severity is determined by comparison with the pre-treatment CTCAE grade of the AE recorded closest to the start of dosing.

For rules on missing or partial dates, see Section 3.3.6. AEs with a missing start time (or where time is not collected) which occur on the same day as first IP administration are reported as treatment emergent.

When assigning AEs to the relevant phase of the study the following rules apply and any deviations must be agreed by the study team:

- Pre-treatment phase: All AEs with a start date after signing the informed consent form (ICF), prior to the first administration of IP that do not subsequently go on to worsen during the treatment emergent phase.
- Treatment emergent phase: All AEs (starting or worsening) on or after the first dose of study intervention and within 90 days after the last dose of study intervention (or surgery, whichever is later) or up to the day prior to start of subsequent therapy (excluding radiotherapy), whichever comes first. The

treatment emergent phase is separated by neoadjuvant, post-surgery, adjuvant period and overall period.

- Off-treatment phase: All AEs starting more than 90 days after last dose of study intervention (or surgery, whichever is later) or once subsequent cancer therapy (excluding radiotherapy) is started, whichever is earlier.

SAEs

A serious adverse event (SAE) is any AE that:

- Results in death
- Is immediately life-threatening
- Requires inpatient hospitalisation or prolongation of existing hospitalisation
- Results in persistent or significant disability/incapacity
- Is a congenital anomaly/birth defect in offspring of the patient
- Is an important medical event that may jeopardise the patient.

AEs of special interest/possible interest

An adverse event of special interest (AESI) is one of scientific and medical interest specific to understanding of the IP and may require close monitoring and rapid communication by the investigator to the sponsor. An AESI may be serious or non-serious. The rapid reporting of AESIs allows ongoing surveillance of these events in order to characterise and understand them in association with the use of study intervention.

AESIs for durvalumab include but are not limited to events with a potential inflammatory or immune-mediated mechanism and which may require more frequent monitoring and/or interventions such as steroids and other immunosuppressants, and/or hormone replacement therapy. These AESIs are being closely monitored in clinical studies with durvalumab monotherapy and combination therapy.

An adverse event of possible interest (AEPI) is defined as an AE that could have a potential inflammatory or immune-mediated pathophysiological basis resulting from the mechanism of action of durvalumab but are more likely to have occurred due to other pathophysiological mechanisms, thus, the likelihood of the event being inflammatory or immune-mediated in nature is not high and/or is most often or usually explained by the other causes. These AEs not routinely arising from an inflammatory or immune-mediated mechanism of action – typically quite general clinical terms that usually present from a multitude of other causes – are classified as AEPIs.

An immune mediated Adverse Event (imAE) is defined as an AE that is associated with drug exposure and is consistent with an immune-mediated mechanism of action and where

there is no clear alternate aetiology. The list of AESIs/imAEs for durvalumab is presented in Section 8.3.10.1 of the CSP. There are no specific AESIs for monalizumab. AESIs for oleclumab, volrustomig, Dato-DXd, AZD0171 and rilvegostomig are specified in CSP Appendix K7, M7, N7, O7 and P7 respectively.

More information regarding AESIs can be found in the protocol. Other categories may be added, or existing terms may be modified as necessary. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which higher-level terms and which preferred terms contribute to each AESI. Further reviews may take place prior to CDL to ensure any further terms not already included are captured within the categories. Preferred terms used to identify AESI is listed before CDL.

Other significant adverse events (OAE)

During the evaluation of the AE data, an AstraZeneca medically qualified expert reviews the list of AEs that were not reported as SAEs and AEs leading to discontinuation of study treatment. Based on the expert's judgement, adverse events of particular clinical importance may, after consultation with the Global Safety Physician, be considered other significant adverse events (OAEs) and reported as such in the Clinical Study Report. A similar review of laboratory values, vital signs, ECGs and other safety assessments are performed for identification of other significant adverse events. This review takes place prior to database lock, and any AEs identified are fully documented in meeting minutes. Further review following database lock may result in ad-hoc OAEs being identified, in this case, the OAEs and resulting summaries are fully documented in the CSR.

Examples of these are marked haematological and other laboratory abnormalities, and certain events that lead to intervention (other than those already classified as serious), dose reduction or significant additional treatment.

4.2.11.2.2 Presentation

All AEs are presented based on safety set.

All treatment emergent AEs (TEAEs) are summarised and listed.

TEAEs are counted once for each patient for calculating percentages of patients experiencing TEAE. In addition, if the same TEAE occurs multiple times within a particular patient, the highest severity and level of relationship observed are reported. For tables by MedDRA system organ class (SOC) and MedDRA preferred term (PT), patients with multiple TEAEs are counted once for each SOC/PT.

An overall summary table of the number of patients experiencing each category of adverse event is produced, including but not limited to:

- Any AE
- AEs with CTCAE grade 3 or higher
- AEs leading to discontinuation of each IP and chemotherapy
- AEs leading to interruption of each IP and chemotherapy
- AEs leading to dose omission of each IP and chemotherapy
- AEs leading to dose reduction of Dato-DXd
- SAEs
- SAEs with outcome of death
- AEs causally related to each IP and chemotherapy (as determined by the reporting investigator)
- Grade 3 or higher AEs causally related to any IP and chemotherapy (as determined by the reporting investigator)
- AEs leading to surgery not done
- AEs leading to a delay in surgery (i.e. surgery more than 40 days after the last dose of study treatment in the neoadjuvant period)
- AEs leading to a delay in adjuvant treatment
- AEs leading to adjuvant treatment not given.

The number of patients experiencing TEAEs by MedDRA SOC and PT are summarised. Summary information (the number and percent of patients) by SOC and PT will be tabulated for:

- All AEs
- AEs by maximum CTCAE grade
- AEs with CTCAE grade 3 or higher
- AEs leading to discontinuation of each IP and chemotherapy
- AEs leading to interruption of each IP and chemotherapy
- AEs leading to dose omission of each IP and chemotherapy

- AEs leading to dose reduction of Dato-DXd
- SAEs
- SAEs with outcome of death.

Summary information (the number and percent of patients) by SOC and PT may be tabulated for:

- All AEs causally related to each IP and chemotherapy (as determined by the reporting investigator)
- All AEs causally related to any IP or chemotherapy (as determined by the reporting investigator)
- All AEs causally related to surgery (as determined by the reporting investigator)
- AEs by maximum reported CTCAE grade on a subset of key group term and preferred term level (as determined by the reporting investigator)
- Immune mediated AEs (as derived programmatically or determined by the reporting investigator, if applicable).

An overall summary table of the number and percentage of patients experiencing each category of adverse event will be produced, as well as an overall summary of the number of events in each category. In addition, a truncated AE table of most common AEs and another table showing most common AEs with CTCAE grade 3 or higher, showing all events that occur in at least 5% of patients overall will be summarised by preferred term, by decreasing frequency. This cut-off may be modified after review of the data. When applying a cut-off (i.e., x %), the raw percentage should be compared to the cut-off, no rounding should be applied first (i.e., an AE with frequency of 4.9% will not appear if a cut-off is 5%).

AEs will be assigned CTCAE grades and summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, system organ class, preferred term and actual treatment group.

Details of any deaths will be summarised and listed for all patients. AEs leading to death will also be summarised.

The overall AE summary table, and all AEs by system organ class and preferred term will be repeated firstly by period as defined in section 3.3.1 and then repeated for the overall period including AEs observed up until 90 days after last dose of study drugs (or surgery, whichever is later) or up to the day prior to start of subsequent anti-cancer therapy (excluding radiotherapy), whichever comes first.

Additionally, summary information (number and percent of patients) may be presented for AEs related to COVID-19. A listing may be provided for patients with AEs with confirmed/suspected COVID-19. If there are no patients with COVID-19, these tables and listings will not be required.

SAEs

SAEs are summarised as described above for the TEAEs.

Any SAE occurring before study treatment will be included in the data listings but will not be included in the summary tables of AEs.

Any SAE occurring within 90 days of discontinuation of IP (i.e. the last dose of study treatment) will be included in the relevant SAE summaries. Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of study treatment) will be flagged in the data listings.

AEs of special interest or possible interest

Grouped summary tables of certain MedDRA preferred terms are produced and may also show the individual preferred terms which constitute each AESI grouping. Groupings are based on preferred terms provided by the medical team prior to CDL, and a listing of the preferred terms in each grouping are provided. There is no AESI for monalizumab.

An overall summary of AESI or AEPI may include number and percentage (%) of patients who have:

- Any AESI or AEPI (per preferred terms disclosed as of clinical data lock)
- Any AESI or AEPI leading to a delay in surgery (i.e., surgery more than 40 days after the last dose of study treatment in the neoadjuvant period)
- At least one durvalumab AESI or AEPI causally related to durvalumab
- At least one oleclumab AESI causally related to oleclumab
- At least one volrustomig AESI causally related to volrustomig
- At least one Dato-DXd AESI causally related to Dato-DXd
- At least one AZD0171 AESI or AEPI causally related to AZD0171
- At least one rilvegostomig AESI causally related to rilvegostomig
- At least one durvalumab AESI or AEPI leading to discontinuation of durvalumab

- At least one oleclumab AESI leading to discontinuation of oleclumab
- At least one volrustomig AESI leading to discontinuation of volrustomig
- At least one Dato-DXd AESI leading to discontinuation of Dato-DXd
- At least one AZD0171 AESI or AEPI leading to discontinuation of AZD0171
- At least one rilvegostomig AESI leading to discontinuation of rilvegostomig.

Data collected for Interstitial Lung Disease / pneumonitis investigation for patients exposed to Dato-DXd will be listed.

Listings

Key subject information listings will be produced as follows:

- A key subject information listing of SAEs
- A key subject information listing of SAEs with outcome of death
- A key subject information listing of deaths
- A key subject information listing of AEs leading to discontinuation of IP(s)
- A key subject information listing of AESIs.

By patient listings of AEs will be produced.

Deaths

A summary of deaths will be provided with number and percentage of patients, categorised as:

- Related to disease under investigation,
- AE outcome=death,
- Both related to disease under investigation and with AE outcome=death,
- AE with outcome=death ≥ 90 days after last treatment dose,
- Deaths ≥ 90 days after last treatment dose, unrelated to AE or disease under investigation, and
- Patients with unknown reason for death.

The following summary tables of AEs with an outcome of death may be prepared and presented by actual treatment group.

- By SOC and PT
- Causally related to IP(s) by SOC and PT.

4.2.11.3 Clinical Laboratory, Blood Sample

4.2.11.3.1 Definitions and Derivations

Laboratory tests are grouped according to chemistry and haematology. Laboratory parameters are assessed at baseline as well as throughout the study. The parameters from Table 7, 8 and 9 of CSP will be presented.

For chemistry and haematology parameters, laboratory abnormalities with toxicity grades according to the NCI CTCAE version 5.0 are derived.

Change from baseline in haematology and clinical chemistry parameters are calculated for each post-dose visit. CTCAE grade is calculated at each visit. Maximum post-baseline CTCAE grades are also calculated. Absolute values are compared to the local laboratory reference range and classified as low (below range), normal (within range or on limits of range) and high (above range). All values classified as high or low are flagged on the listings.

Liver Function Parameters

Patients with elevated post-baseline alanine aminotransferase (ALT), aspartate aminotransferase (AST) or Total Bilirubin that fall into these categories are identified.

Table 6 Liver Function Parameters

Liver Function Parameters	Category
ALT	• $\geq 3 \times - \leq 5 \times \text{ULN}$ • $> 5 \times - \leq 8 \times \text{ULN}$ • $> 8 \times - \leq 10 \times \text{ULN}$ • $> 10 \times - \leq 20 \times \text{ULN}$ • $> 20 \times \text{ULN}$
AST	• $\geq 3 \times - \leq 5 \times \text{ULN}$ • $> 5 \times - \leq 8 \times \text{ULN}$ • $> 8 \times - \leq 10 \times \text{ULN}$, • $> 10 \times - \leq 20 \times \text{ULN}$ • $> 20 \times \text{ULN}$
Total bilirubin	• $\geq 2 \times - \leq 3 \times \text{ULN}$ • $> 3 \times - \leq 5 \times \text{ULN}$ • $> 5 \times \text{ULN}$
ALT or AST	• $\geq 3 \times - \leq 5 \times \text{ULN}$ • $> 5 \times - \leq 8 \times \text{ULN}$ • $> 8 \times - \leq 10 \times \text{ULN}$, • $> 10 \times - \leq 20 \times \text{ULN}$ • $> 20 \times \text{ULN}$

Liver Function Parameters	Category
Potential Hy's law	(AST $\geq 3 \times$ ULN or ALT $\geq 3 \times$ ULN) and (Total Bilirubin $\geq 2 \times$ ULN)^a

ULN: upper limit of normal range.

^a It includes all patients who have ALT or AST $\geq 3 \times$ ULN and total bilirubin (BILI) $\geq 2 \times$ ULN, and in which the elevation in transaminases precede or coincide with (that is, on the same day as) the elevation in BILI.

Assessment of Nephrotoxicity

Creatinine clearance rate (CCr) is calculated using serum creatinine and the Cockcroft-Gault formula to estimate glomerular filtration rate (GFR). Baseline and “worst-case” on treatment C_{Cr} value are categorised for the following categories:

- Normal: ≥ 90 mL/min
- Mild Impairment: $\geq 60 - < 90$ mL/min
- Moderate Impairment: $\geq 30 - < 60$ mL/min
- Severe Impairment: $\geq 15 - < 30$ mL/min
- Kidney Failure: < 15 mL/min

4.2.11.3.2 Presentations

Summaries of results over time and change from baseline will be presented for each laboratory parameter. Shift from baseline to the worst on-treatment value is presented.

Listings may be provided for laboratory results.

Liver Function Parameters

Individual patient data where elevated ALT or AST plus total bilirubin fall into the “Potential Hy's law” are summarised and/or listed. Maximum on-treatment ALT and AST versus maximum on-treatment total bilirubin will be presented.

Assessment of Nephrotoxicity

Not applicable.

4.2.11.4 Clinical Laboratory, Urinalysis

4.2.11.4.1 Definitions and Derivations

Laboratory parameters are assessed at baseline as well as throughout the study.

Change from baseline in urinalysis parameters are calculated for each post-dose visit. CTCAE grade is calculated at each visit. Maximum post-baseline CTCAE grade is also

calculated. Absolute values are compared to the local laboratory reference range and classified as low (below range), normal (within range or on limits of range) and high (above range). All values classified as high or low are flagged in the listings.

4.2.11.4.2 Presentations

Not applicable.

4.2.11.5 Other Laboratory Evaluations

4.2.11.5.1 Definitions and Derivations

Not applicable.

4.2.11.5.2 Presentations

Not applicable.

4.2.11.6 Vital Signs

4.2.11.6.1 Definitions and Derivations

Vital signs (BP, pulse, temperature, and respiration rate) will be evaluated according to the SoAs in CSP. Body weight is also recorded at each visit along with vital signs.

Vital signs are assessed at baseline and throughout the study.

4.2.11.6.2 Presentations

Vital signs will be summarised by study visit which may include descriptive statistics for the value of the parameters and the changes from baseline.

4.2.11.7 Electrocardiogram

4.2.11.7.1 Definitions and Derivations

Electrocardiogram (ECG) parameters are assessed at baseline as well as throughout the study. ECG parameters include: QT and QTcF. The QTcF is considered as the primary correction method to assess patient cardiac safety.

The notable ECG interval values while on treatment are:

- Maximum QTcF intervals > 450 milliseconds, $\geq$ 470 milliseconds, and > 500 milliseconds.
- Maximum changes from baseline in QTcF > 30, >60, and > 90 milliseconds.

4.2.11.7.2 Presentations

Not applicable.

4.2.11.8 Other Safety Assessments

No other safety assessments are planned.

4.2.11.8.1 Definitions and Derivations

Not applicable.

4.2.11.8.2 Presentations

Not applicable.

4.2.12 Other Endpoint: Biomarkers

CCI

4.2.12.1 Definition: Biomarkers

This will be discussed in a separate analysis plan.

4.2.12.2 Presentations: Biomarkers

The presentations will be discussed in a separate analysis plan.

4.3 Pharmacodynamic Endpoint(s)

Exploratory analyses pertaining to pharmacodynamic endpoint(s) may be conducted. More details will be specified upon data availability.

4.3.1 Analysis

More details about exploratory analyses of pharmacodynamic endpoint(s) will be specified upon data availability.

4.3.2 Definitions and Derivations

More details about exploratory analyses of pharmacodynamic endpoint(s) will be specified upon data availability.

4.3.3 Presentation

More details about exploratory analyses of pharmacodynamic endpoint(s) will be specified upon data availability.

5 INTERIM ANALYSIS

Continuous Monitoring

CCI

CCI Each treatment arm will be evaluated separately and independently. CCI Enrolment is not paused during the interim analysis before the No-Go decision is made.

CCI For any interim analyses and the final analysis, a pCR response is defined as lack of any viable tumour cells after complete evaluation in the resected lung cancer specimen and all sampled regional lymph nodes as determined by central BIPR and described by IASLC 2020 (Travis et al., 2020). In addition to pCR, mPR, Feasibility to surgery, ORR, PK and ADA may also be evaluated at the interim using the appropriate analysis population as specified in Section 3.2.

The interim conducted at CCI patients will be evaluated to determine futility. Based on the Go/No-Go decision framework (Frewer et al., 2016), a conclusion of lack of evidence of the targeted anti-tumour activity is reached if 6 or fewer pCRs are observed out of CCI. The TV and LRV values may be revised based on emerging data. The futility criterion at interim is calculated such that interim results would lead to a predictive probability of a stop decision at final analysis being at least CCI. Moreover, the ultimate Go No-Go decision will be based on totality of the data including, but not limited to, efficacy, safety, and biomarker data.

6 REFERENCES

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika* 1934; 26:404–416.

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

Cottrell, T. R, Thompson ED, Forde P M, Stein J E, Duffield A S, Anagnostou V, Rekhtman N, Anders R A, Cuda J D, Illei P B, Gabrielson E, Askin F B, Niknafs N, Smith K N, Velez M J, Sauter J L, Isbell J M, Jones D R, Battafarano R J, Yang S C, Danilova L, Wolchock J D, Topalian S L, Velculescu V E, Pardoll D M, Brahmer J R, Hellmann M D, Chaft J E, Cimino-Mathews A, Taube J M. Pathologic features of response to neoadjuvant anti-PD-1 in resected non-small-cell lung carcinoma: a proposal for quantitative immune-related pathologic response criteria (irPRC). *Annals of Oncology* 2018 ; 29.8: 1853-1860

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guidelines (version 1.1). *European Journal of Cancer* 2009; 45:228-47.

Frewer P, Mitchell P, Watkins C, Matcham J. Decision-making in early clinical drug development. *Pharma Statist.* 2016; 15:255-263.

Lalonde R L, Kowalski K G, Hutmacher M M, Ewy W, Nichols D J, Milligan P A, Corrigan B W, Lockwood P A, Marshall S A, Benincosa L J, Tensfeldt T G, Parivar K, Amantea M, Glue P, Koide H, Miller R. Model-based drug development. *Clin Pharmacol Ther.* 2007 Jul; 82(1):21-32

Travis et al, 2020 IASLC Multidisciplinary Recommendations for Pathologic Assessment of Lung. Cancer Resection Specimens After Neoadjuvant Therapy. *J Thorac Oncol.* 2020 May;15(5):709-740. doi:10.1016/j.jtho.2020.01.005.

7 APPENDICES

Appendix A RECIST

Derivation of RECIST Visit Responses

For all patients, the RECIST tumour response data is used to determine each patient's visit response according to RECIST version 1.1. It is also used to determine if and when a patient has progressed in accordance with RECIST and their best objective response to study treatment.

RECIST assessments will be performed according to the following schedule:

- 1 Baseline scan – performed within 28 days prior to and as close as possible to randomisation.
- 2 Neoadjuvant follow-up scan – at Cycle 2 and acquired upon completion of neoadjuvant period, i.e. end of neoadjuvant treatment if applicable, and prior to surgery.
- 3 First post-surgical scan – acquired 5 weeks (+/- 2 weeks) after surgery and prior to, but as close as possible to, the start of adjuvant study interventions. For subjects taking optional post-surgery radiotherapy, this scan should be performed prior to starting radiotherapy.
- 4 Post-surgical follow-up scans –
 - Follow-up scans are acquired Q12W $\pm$ 1 week (relative to the date of surgery) until week 48;
 - Then Q24W $\pm$ 2 weeks (relative to the date of surgery) until week 192 (approximately 4 years);
 - And then Q48W $\pm$ 2 weeks (relative to the date of surgery) thereafter, until RECIST 1.1-defined radiological PD, consent withdrawal, death, or study completion as determined by Sponsor.

For patients who do not have surgery and therefore will not have a first post-surgical scan or adjuvant treatment, follow-up scans are acquired Q12 $\pm$ 1 week (relative to the date of the pre-planned surgery) until week 48 and then Q24 $\pm$ 2 weeks (relative to the date of pre-planned surgery) until week 192 (approximately 4 years) and then Q48 $\pm$ 2 weeks (relative to the date of pre-planned surgery). These follow-up scans will use the original neoadjuvant screening as a baseline scan for RECIST 1.1 assessments.

If a pre-planned surgery date is not available for planning RECIST 1.1 follow-up scans on patients who did not undergo surgery (for reasons other than PD), the follow-up schedule will use 40 days after the last study drugs administration as the date of pre-planned surgery.

If an unscheduled assessment is performed, and the patient has not progressed, every attempt is made to perform the subsequent assessments at their scheduled visits. This schedule is followed in order to minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

From the investigator's review of the imaging scans, the RECIST tumour response data is used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients are programmatically assigned a RECIST 1.1 visit response of CR, PR, SD or PD, using the information from target lesions (TLs), non-target lesions (NTLs) and new lesions and depending on the status of their disease compared with baseline and previous assessments. If a patient has had a tumour assessment that cannot be evaluated then the patient is assigned a visit response of not evaluable (NE) (unless there is evidence of progression in which case the response is assigned as PD). For subjects following surgery, given the surgery is considered an intervention, the lesion intervention = Yes will be marked on all subsequent RECIST eCRFs and hence only NE or PD are potential outcomes unless scaling results in SD/PR/CR.

Please refer to [Table 7](#) for the definitions of CR, PR, SD and PD.

RECIST outcomes (i.e. EFS, ORR etc.) are calculated programmatically for the site investigator data (see last section of this Appendix) from the overall visit responses.

During the adjuvant period, the imaging schedule must be followed regardless of any delays in dosing. Patients who do not have surgery, will not have adjuvant treatment and scans for these patients will be conducted according to the post-surgical follow-up schedule relative to the date of the pre-planned surgery. Follow-up scans will use the original neoadjuvant screening scan as a baseline scan for RECIST 1.1 assessments.

Target lesions (TLs) – site investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD), (except lymph nodes which must have short axis ≥ 15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements. A patient can have a maximum of five measurable lesions recorded at baseline with a maximum of two lesions per organ (representative of all lesions involved and suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to randomisation/first dose is used to define the baseline sum of TLs. It is the case that, on

occasion, the largest lesion does not lend itself to reproducible measurement. In which circumstance the next largest lesion, which can be measured reproducibly, is selected.

All other lesions (or sites of disease) not recorded as TL are identified as non-target lesions (NTLs) at baseline. Measurements are not required for these lesions, but their status are followed at subsequent visits.

Note: For patients who do not have measurable disease at entry (i.e. no TLs) but have non-measurable disease, evaluation of overall visit responses are based on the overall NTL assessment and the absence/presence of new unequivocal lesions (see section below for further details). If a patient does not have measurable disease at baseline then the TL visit response is Not Applicable (NA).

Table 7 TL Visit Responses (RECIST 1.1)

Visit Responses	Description
Complete response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A $\geq 20\%$ increase in the sum of diameters of TLs and an absolute increase of $\geq 5\text{mm}$, taking as reference the smallest sum of diameters pre-surgery since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.
Not evaluable (NE)	Only relevant in certain situations (i.e. if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response. If no PD, subjects with surgery are considered having an intervention and hence a response of NE unless scaling results in SD/PR/CR.
Not Applicable (NA)	No TLs are recorded at baseline.

Rounding of TL data

For calculation of PD and PR for TLs percentage changes from baseline and previous minimum are rounded to one d.p. before assigning a TL response. For example 19.95% is rounded to 20.0% but 19.94% is rounded to 19.9%

Missing TL data

For a visit to be evaluable then all TL measurements are recorded. However, a visit response of PD is still assigned if any of the following occurred

- A new unequivocal lesion, or a new equivocal lesion, converted to unequivocal during a follow-up visit, is recorded,
- A NTL visit response of PD is recorded,
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of $\geq 5\text{mm}$, from nadir even assuming the non-recorded TLs have disappeared.

Note: the nadir can only be taken from assessments where all the TLs had a LD recorded. If a particular TL measurement is missing, progression can still be assessed by assuming the TL LD = 0mm. Nadir is classed as the prior smallest sum of diameters pre-surgery since treatment started including the baseline sum of diameters.

If the TL visit response is not recorded as PD, then the TL visit response is NE.

If there is at least one TL measurement missing and a visit response of PD cannot be assigned, the visit response is NE.

If all TL measurements are missing then the TL visit response is NE. Overall visit response is also NE, unless there is a progression of non-TLs or new unequivocal lesions, in which case the response is PD.

Lymph nodes

For lymph nodes, if the size reduces to $< 10\text{mm}$ then these are considered non-pathological. However, a size is still given and this size is still used to determine the TL visit response as normal. In the special case where all lymph nodes are $< 10\text{mm}$ and all other TLs are 0mm then although the sum may be $> 0\text{mm}$ the calculation of TL response should be over-written as a CR. Note that deviations to this approach, such as counting such diameters as 0mm are not recommended.

TL visit responses subsequent to CR

Only CR, PD or NE can follow a CR. If a CR has occurred then the following rules at the subsequent visits must be applied:

- If all lesions meet the CR criteria (i.e. 0mm or $< 10\text{mm}$ for lymph nodes) then response is set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node short axis increases by 20% but remains $< 10\text{mm}$.

- If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm for lymph nodes) then response is set to NE irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node short axis increases by 20% but remains < 10mm.
- If not all lesions are missing, and those that are non-missing do not meet the CR criteria (i.e. a pathological lymph node selected as TL has short axis ≥ 10 mm or the reappearance of previously disappeared lesion), then response is set to PD.
- If all lesions are missing the response is set to NE.

TL too big to measure

If a TL becomes too big to measure this is indicated in the database and a size ('x') above which it cannot be accurately measured is recorded. If using a value of x in the calculation of TL response does not give an overall visit response of PD, then this is flagged and reviewed by the study team blinded to treatment assignment. It is expected that a visit response of PD remains in the vast majority of cases.

TL too small to measure

If a TL becomes too small to measure then this is indicated as such on the case report form and a value of 5mm is entered into the database and used in TL calculations. However a smaller value is used if the radiologist has not indicated 'too small to measure' on the case report form and has entered a smaller value that can be reliably measured. If a TL response of PD results (at a subsequent visit) then this is reviewed by the study team blinded to treatment assignment.

Lesions that split in two

If a TL splits in two, then the LDs of the split lesions should be summed and reported as the LD for the lesion that split.

Lesions that merge

If two TLs merge, then the LD of the merged lesion should be recorded for one of the TL sizes and the other TL size should be recorded as 0 mm.

Change in method of assessment of TLs

CT, MRI and clinical examination are the only methods of assessment that can be used within a trial, with CT and MRI being the preferred methods and clinical examination only used in special cases. If a change in method of assessment occurs, between CT and MRI this will be considered acceptable and no adjustment within the programming is needed.

If a change in method involves clinical examination (e.g., CT changes to clinical examination or vice versa), any affected lesions should be treated as missing.

Irradiated lesions/lesion intervention

The surgical resection in this study will be considered a lesion intervention and should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumors.

After surgery the investigator must record lesion measurements as usual, but also record a TL Intervention for all visits after the surgery. In the case of a complete surgical resection, a lesion measurement of 0 mm will be recorded with the TL intervention flag. If an unequivocal lesion, or a new equivocal lesion, converted to unequivocal during a follow-up visit, appears in the same position as one that “disappears” and is confirmed during a follow-up visit, it is considered a New Unequivocal Lesion and PD.

Table 8 Example: RECIST visit responses following complete resection

	Baseline	Neoadjuvant follow-up scan	Complete Surgical Resection	1 st post-surgery scan	2 nd post-surgery scan	3 rd post-surgery scan
Sum of diameters	80 mm	50 mm		0 mm	0 mm	0 mm (lesion that appears in same position as lesion that has disappeared is recorded as new lesion)
TL Intervention Flag	No	No		Yes	Yes	Yes

VISIT RESPONSE	-	PR		NE	NE	PD (due to new unequivocal lesion otherwise it would be NE)
-------------------	---	----	--	----	----	---

In the scenario where a subject has a partial resection, it is marked as having intervention and assessed as NE at all subsequent visits until there's clear progression by increase in sum of diameters of $\geq 20\%$ and 5 mm over nadir or new unequivocal lesion. Nadir can only be taken from assessments where all the TLs had a LD recorded or scaling applied. Nadir is classed as the smallest sum of diameters pre-surgery since treatment started, including the baseline sum of diameters.

Table 9 Example: RECIST visit responses following partial resection

	Baseline	Neoadjuvant follow-up scan	Partial Surgical Resection	1st post-surg scan	2nd post-surg scan	3rd post-surg scan
Sum of diameters	80 mm	50 mm		4 mm	4 mm	66 mm
TL Intervention Flag	No	No		Yes	Yes	Yes
VISIT RESPONSE	-	PR		NE	NE	PD

Table 10 Example: RECIST visit responses following partial resection

	Baseline	Neoadjuvant follow-up scan	Partial Surgical Resection	1st post-surg scan	2nd post-surg scan	3rd post-surg scan
Sum of diameters	80 mm	50 mm		4 mm	4 mm	51 mm
TL Intervention Flag	No	No		Yes	Yes	Yes
VISIT RESPONSE	-	PR		NE	NE	NE

Any TL (including lymph nodes), which has had intervention on the study should be handled in the following way in programming:

- Step 1: the diameters of the TLs (including the lesions that have had intervention) will be summed and the calculation will be performed in the usual manner. If the visit response is PD, this will remain as a valid response category.
- Step 2: If there was no evidence of progression after step 1, treat the lesion diameter (for those lesions with intervention) as missing and if $\leq 1/3$ of the TLs have missing measurements then scale up as described in the ‘Scaling’ section below. If the scaling results in a visit response of PD then the subject would be assigned a TL response of PD.

- Step 3: If, after both steps, PD has not been assigned then, if appropriate (i.e. if $\leq 1/3$ of the TLs have missing measurements), the scaled sum of diameters calculated in step 2 should be used, and PR or SD then assigned if applicable as the visit response. Subjects with intervention are evaluable for CR as long as all non-intervened lesions are 0 (or $<10\text{mm}$ for lymph nodes) and the lesions that have been subject to intervention have a value of 0 (or $<10\text{mm}$ for lymph nodes) recorded. If scaling up is not appropriate due to too few non-missing measurements then the visit response will be set as NE.

At subsequent visits, the above steps will be repeated to determine the TL and overall visit response.

Scaling (applicable only for irradiated lesions/lesion intervention)

If $> 1/3$ of TL measurements are missing (because of intervention) then the TL response will be NE, unless the sum of diameters of non-missing TL would result in PD (i.e. if using a value of 0 for missing lesions, the sum of diameters has still increased by 20% or more compared to nadir and the sum of TLs has increased by $\geq 5\text{mm}$ from nadir).

If $\leq 1/3$ of the TL measurements are missing (because of intervention) then the results will be scaled up (based on the sizes at the nadir visit to give an estimated sum of diameters) and this will be used in calculations; this is equivalent to comparing the visit sum of diameters of the non-missing lesions to the nadir sum of diameters excluding the lesions with missing measurements.

Example of scaling

Lesion 5 is missing at the follow-up visit; the nadir TL sum including lesions 1-5 was 74 mm.

The sum of lesions 1-4 at the follow-up is 68 mm. The sum of the corresponding lesions at the nadir visit is 62 mm.

Scale up as follows to give an estimated TL sum of 81 mm:
 $68 \times 74 / 62 = 81 \text{ mm}.$

CR will not be allowed as a TL response for visits where there is missing data. Only PR, SD or PD (or NE) could be assigned as the TL visit response in these cases. However, for visits with $\leq 1/3$ lesion assessments not recorded, the scaled up sum of TLs diameters will be included when defining the nadir value prior to surgery for the assessment of progression.

Non-target lesions (NTLs) and new lesions – site investigator data.

At each visit, the investigator is to record an overall assessment of the NTL response. This section provides the definitions of the criteria used to determine and record overall response for NTL at the investigational site at each visit.

NTL response is derived based on the investigator's overall assessment of NTLs as follows:

Table 11 NTL Visit Responses

Visit Responses	Description
Complete response (CR)	Disappearance of all NTLs present at baseline with all lymph nodes non-pathological in size (<10 mm short axis).
Progressive disease (PD)	Unequivocal progression of existing NTLs. Unequivocal progression may be due to an important progression in one lesion only or in several lesions. In all cases, the progression MUST be clinically significant for the physician to consider changing (or stopping) therapy.
Non-CR/Non-PD	Persistence of one or more NTLs with no evidence of progression.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: For patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met. Note: subjects with intervention and no PD will be classed as NE.
Not Applicable (NA)	Only relevant if there are no NTLs at baseline.

To achieve 'unequivocal progression' on the basis of NTLs, there must be an overall level of substantial worsening in non-target disease such that, even in the presence of SD or PR in TLs, the overall tumour burden has increased sufficiently to merit a determination of disease progression. A modest 'increase' in the size of one or more NTLs is usually not sufficient to qualify for unequivocal progression status.

Details of any new lesions are also recorded with the date of assessment. The presence of one or more new lesions is assessed as progression.

A lesion identified at a follow up assessment in an anatomical location that was not scanned at baseline is considered a new lesion and indicates disease progression.

New lesions will be identified at each visit and recorded on the CRF. Biopsy data will be collected for new lesions to determine if the new lesion is a new primary malignancy.

The finding of a new lesion should be unequivocal: i.e. not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumour.

New lesions are identified via a Yes/No tick box. The absence and presence of new lesions at each visit are listed alongside the TL and NTL visit responses.

A new unequivocal lesion indicates progression so the overall visit response is PD irrespective of the TL and NTL response.

For the EFS and DFS endpoints, pathological confirmation from biopsied lesions, if performed according to investigators judgement and local practice, will also be taken into consideration (as applicable). A new malignancy, that is not NSCLC, as confirmed by pathology, is not considered an EFS/DFS event. In the absence of pathology confirmation, a new unequivocal lesion will be considered an EFS/DFS event.

If the question ‘Any new lesions since baseline’ has not been answered with Yes or No and the new lesion details are blank, this is not evidence that no new lesions are present, but should not overtly affect the derivation. This scenario (i.e. whereby new lesion response is NE), should only occur in exceptional cases, as missing data for the new lesion field should always be queried.

Symptomatic progression is not a descriptor for progression of NTLs: it is a reason for stopping study therapy and is not included in any assessment of NTLs.

Patients with ‘symptomatic progression’ requiring discontinuation of treatment without objective evidence of disease progression at that time should continue to undergo tumour assessments where possible until objective disease progression is observed.

Overall visit response – site investigator data

[Table 12](#) defines how the previously defined TL and NTL visit responses are combined with new lesion information to give an overall visit response.

Table 12 Overall visit responses

TARGET	NON-TARGET	NEW LESIONS*	OVERALL VISIT RESPONSE
CR	CR or NA	No (or NE)	CR
CR	Non-CR/Non-PD or NE	No (or NE)	PR
PR	Non-PD or NE or NA	No (or NE)	PR
SD	Non-PD or NE or NA	No (or NE)	SD
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD
NE	Non-PD or NE or NA	No (or NE)	NE
NA	CR	No (or NE)	CR
NA	Non-CR/Non-PD	No (or NE)	SD
NA	NE	No (or NE)	NE
NA	NA	No (or NE)	NED

*One or more new unequivocal lesions of the primary disease (NSCLC). If all lesions are pathologically confirmed to be new primary malignancies (not NSCLC) then PD by new lesions is not met.

Signature Page for VV-RIM-11517573 System v1.0
D9077C00001 SAP v6.0

Approve: Document Level Task Verdict: Approved	Kevin Gris Content Approval 26-Mar-2026 09:13:19 GMT+0000
---	---

Approve: Document Level Task Verdict: Approved	Yutao Liu Content Approval 26-Mar-2026 16:19:02 GMT+0000
---	--

Signature Page for VV-RIM-11517573 System v1.0
D9077C00001 SAP v6.0