An Open-Label, Multinational, Multicenter, Phase 3b/4 Study of Trastuzumab Deruxtecan in Patients With or Without Baseline Brain Metastasis With Previously-Treated Advanced/Metastatic HER2-Positive Breast Cancer (DESTINY-Breast12)

ClinicalTrials.gov Identifier: **NCT04739761**

Statistical Analysis Plan: version 4.0, dated 07 Dec 2023

STATISTICAL ANALYSIS PLAN	
Study Code	D9673C00007
Edition Number	4.0
Date	7-Dec-2023

An Open-Label, Multinational, Multicenter, Phase 3b/4 Study of Trastuzumab Deruxtecan in Patients With or Without Baseline Brain Metastasis With Previously-Treated Advanced/Metastatic HER2-Positive Breast Cancer (DESTINY-Breast12)

Sponsor Name:

AstraZeneca AB, 151 85 Södertälje, Sweden

Daiichi Sankyo Co., Ltd., 3-5-1, Nihonbashi-honcho, Chuo-ku, Tokyo, Japan
(Local sponsor for Japan)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation or Specialized Term	Definition
AE	adverse events
AESI	adverse event of special interest
AJCC	American Joint Committee on Cancer
ALT	alanine aminotransferase
aPTT	activated partial thromboplastin time
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BM	brain metastasis
CCI	CCI
BMI	body mass index
BoR	best objective response
CI	confidence interval
CNS	central nervous system
COVID-19	coronavirus disease 2019
CR	complete response
CRF	case report form
CSP	clinical study protocol
CSR	clinical study report
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCO	data cut-off
DI	dose intensity
DLCO	diffusion capacity of the lungs for carbon monoxide
DoR	duration of response
DoT	duration of treatment
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
ER	estrogen receptor
FAS	Full Analysis Set
FAS-CNS	Full Analysis Set for isolated CNS progression analyses
FAS-ORRc	Full Analysis Set for ORR (non-final) ICR analyses
FAS-ORRi	Full Analysis Set for ORR (non-final) investigator analyses
FAS-ORR-CNSc	Full Analysis Set for CNS ORR (non-final) ICR analyses
FAS-ORR-CNSi	Full Analysis Set for CNS ORR (non-final) investigator analyses
FAS-ST	Full Analysis Set for subsequent treatments
FEV1	forced expiratory volume in 1 second
FVC	forced vital capacity
HER2	human epidermal growth factor receptor 2
CCI	CCI

HRQoL	health-related quality of life
ICR	independent central review
IDMC	Independent Data Monitoring Committee
ILD	interstitial lung disease
INR	international normalized ratio
IV	intravenous
KM	Kaplan-Meier
LD	longest diameter
LVEF	left ventricular ejection fraction
MDASI	MD Anderson Symptom Inventory
MedDRA	Medical Dictionary for Regulator Activities
MRI	magnetic resonance imaging
NA	not applicable
NANO	Neurologic Assessment in Neuro-oncology
NE	not evaluable
NTL	non-target lesion
OAE	other significant adverse event
ORR	objective response rate
OS	overall survival
PAL	Paired Associates Learning
PALFAMS	PAL First Attempt Memory Score
PALTEA	PAL Total Errors Adjusted
PD	progression of disease
PDI	planned dose intensity
PFS	progression-free survival
PFS2	time from first dose of study intervention to second progression or death
PgR	progesterone receptor
PR	partial response
PRO	patient reported outcome
PT	preferred term
PTT	partial thromboplastin time
QLQ-C30	30 item core quality of life questionnaire
QoL	quality of life
QTc	corrected QT interval
QTcF	QT interval corrected by Fridericias's formula
CCI	CCI
RDI	relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
RS	raw score
RTI	Reaction Time Task
RTIFMDMT	RTI Median Five-Choice Movement Time
RTIFMDRT	RTI Median Five-Choice Reaction Time
SAE	serious adverse event

SAF	Safety Analysis Set
SAF-ILD	Safety Analysis Set for ILD/pneumonitis patients
SD	stable disease
SGRQ-I	St. George's Respiratory Questionnaire - idiopathic pulmonary fibrosis-specific version
SOC	system organ class
SpO2	peripheral capillary oxygen saturation
SRS	stereotactic radiosurgery
SWM	Spatial Working Memory
SWMBE	SWM Between Errors
SWMBE4	SWM Between Errors 4 Boxes
SWMBE6	SWM Between Errors 6 Boxes
SWMBE8	SWM Between Errors 8 Boxes
SWMS	SWM Strategy
SWMTE	SWM Total Errors
T-DXd	trastuzumab deruxtecan
TEAE	treatment-emergent adverse event
CCI	CCI
TKI	tyrosine kinase inhibitor
TL	target lesion
TNM	tumor, node and metastasis
ULN	upper limit of normal
WBRT	whole brain radiation therapy
WHODD	World Health Organization - Drug Dictionary

AMENDMENT HISTORY

CATEGORY **Change refers to:**	**Date**	**Description of change**	**In line with CSP?**	**Rationale**
N/A	22-Mar-2021	Initial approved SAP	N/A	N/A
Primary, secondary, safety and exploratory endpoints	28-May-2021	Several updates to align with new version of protocol, including modifications in study objectives.	Yes (PA 1.0)	Protocol Amendment 1.0
Subgroup analyses	28-May-2021	New subgroups added in sections 5.2.1.4, 5.2.16.3, 5.4.2.3 and 5.4.10	Yes (PA 1.0)	Protocol Amendment 1.0
Protocol deviations	28-May-2021	Section 4.3.6 updated with a subset of key entry criteria	Yes (PA 1.0)	Protocol Amendment 1.0
Other minor updates	28-May-2021	Other minor updates and correction of typos	Yes (PA 1.0)	Protocol Amendment 1.0
Primary, secondary, safety and exploratory endpoints	07-Sep-2022	New endpoints and subsets added. CCI [redacted] CCI [redacted] It's been added and other endpoints descriptions have been readjusted accordingly	Yes (PA 2.0)	Protocol Amendment 2.0, Correction
Subgroup analyses	07-Sep-2022	Subgroups modified	Yes (PA 2.0)	Protocol Amendment 2.0
Protocol deviations	07-Sep-2022	Section 4.3.6 updated	Yes (PA 2.0)	Correction
Other minor updates	07-Sep-2022	Other minor updates and correction of typos	Yes (PA 2.0)	Protocol Amendment 2.0, Correction
Primary, secondary, safety and exploratory endpoints	07-Dec-2023	Added analyses for ILD patients in section 5.3.9. Added additional analysis for duration of treatment in section 5.2.7	Yes (PA 2.0)	Correction

Subgroup analyses	07-Dec-2023	Added two subgroup analysis for TEAEs in section 5.3.2.3. Added subgroup to demographics in section 5.1.4.1. Clarification on applicable cohorts for subgroups in section 5.2.1.4	Yes (PA 2.0)	Correction
Other minor updates	07-Dec-2023	Minor word changes and clarifications in primary/secondary endpoint section	Yes (PA 2.0)	Correction

1 INTRODUCTION

The purpose of this document is to give details for the statistical analysis of study D9673C00007 supporting the clinical study report (CSR). The reader is referred to the clinical study protocol (CSP) amendment 2 (dated 07 Jun 2022) and the case report form (CRF) for details of study conduct and data collection.

2 STUDY DETAILS

2.1 Study Objectives

Table 1 Objectives and Endpoints

Objectives	Endpoints
Primary	
To describe the overall treatment effect of T-DXd in HER2-positive MBC patients with or without baseline BM	Patients without BM at baseline (Cohort 1) • ORR by RECIST 1.1 per ICR Patients with BM at baseline (Cohort 2): • PFS by RECIST 1.1 per ICR
Secondary	
To describe the treatment effect on the development and progression of BM in patients with or without baseline BM using additional efficacy measurements	Patients in both cohorts: • OS • DoR by RECIST per ICR • Time to progression by RECIST per ICR • DoT on subsequent lines of therapy • PFS2 Patients without BM at baseline (Cohort 1): • Incidence of new symptomatic CNS metastasis during treatment In patients who develop isolated CNS progression, receive local therapy, and continue on protocol therapy: • Time to next progression (CNS or extracranial) or death • Site (CNS vs extracranial vs both) of next progression

Table 1 Objectives and Endpoints

Objectives	Endpoints
To describe efficacy in patients with stable or untreated BM	Patients with BM at baseline (Cohort 2): • ORR by RECIST 1.1 per ICR • CNS PFS by CNS RECIST 1.1 per ICR • Time to new CNS lesions • CNS ORR by CNS RECIST 1.1 per ICR • CNS DoR by CNS RECIST 1.1 per ICR
To describe the effect of T-DXd on symptoms, functioning and HRQoL in HER2-positive MBC patients with or without baseline BM	Changes in symptoms, functioning, and HRQoL as measured by: • All patients: EORTC QLQ-C30, NANO scale, cognitive tests • BM patients (Cohort 2): MDASI brain tumor-specific items • ILD/pneumonitis patients: SGRQ-I
Safety	
To describe the safety profile of T-DXd	Safety and tolerability will be evaluated in terms of AEs, vital signs, clinical laboratory results, and ECGs. Assessments related to AEs will also include: • Rate of investigator-assessed ILD/pneumonitis - PTs will be matched with most commonly-reported terms within ILD cluster terms - Rate of ILD clinical symptoms resolution among ILD patients who have been treated with high-dose steroid (total daily dose > 2 mg dexamethasone or equivalent) • Rate of AEs among patients with baseline BM who are treated with concurrent high-dose steroid (total

Table 1 Objectives and Endpoints

Objectives	Endpoints
	daily dose > 2 mg dexamethasone or equivalent)
Exploratory	
CCI	CCI CCI CCI CCI
CCI	CCI CCI CCI CCI CCI
CCI	CCI
CCI	CCI
CCI	CCI
CCI	CCI

AE = adverse event; BM = brain metastasis; CCI CNS = central nervous system; CCI DoR = duration of response; DoT = duration of treatment; ECG = electrocardiogram; EORTC = European Organization for the Research and Treatment of Cancer; HER2 = human epidermal growth factor receptor 2; CCI HRQoL = health-related quality of life; ICR = independent central review; ILD = interstitial lung disease; MBC = metastatic breast cancer; MDASI = MD Anderson Symptom Inventory; NANO = Neurologic Assessment in Neuro-oncology; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; PFS2 = time to second progression or death; PT = preferred term; QLQ-30 = 30 item core quality of life questionnaire; CCI RECIST 1.1 = Response Evaluation Criteria in Solid

Tumors Version 1.1; SGRQ-I = St. George's Respiratory Questionnaire – idiopathic pulmonary fibrosis version; T-DXd = trastuzumab deruxtecan; CCI [redacted]

2.2 Study Design

This is a Phase 3b/4, open-label, multicenter, international study assessing the efficacy and safety of trastuzumab deruxtecan (T-DXd) in patients with or without brain metastasis (BM), with previously-treated advanced/metastatic HER2-positive breast cancer whose disease has progressed on prior anti-HER2 based regimens and who received no more than 2 lines/regimens of therapy in the metastatic setting (excluding tucatinib). Patients will be enrolled into 1 of 2 cohorts according to the presence or absence of BMs at baseline:

- Cohort 1: Patients without BM at baseline
- Cohort 2: Patients with BM at baseline

Approximately 500 eligible patients will be treated in this study with 5.4 mg/kg T-DXd every three weeks (q3w) until RECIST 1.1-defined radiological progression outside CNS, unless there is unacceptable toxicity, withdrawal of consent, or another criterion for discontinuation is met. To ensure adequate representation of patients with BM at baseline, the maximum number of patients without BM at baseline will be limited to 250. Cohort 1 (participants without BM at baseline) will additionally be limited to include no more than 25% third-line participants. The study will evaluate ORR by RECIST 1.1, PFS by RECIST 1.1, the incidence of new symptomatic CNS metastasis, and other measures of efficacy, and further characterize the safety and tolerability profile of T-DXd.

2.3 Number of Patients

Approximately 500 eligible patients will be treated in the study. Of these, approximately 250 eligible patients without BM at baseline (Cohort 1) and approximately 250 eligible patients with BM at baseline (Cohort 2) will be treated.

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI

Table 2 CCI

CCI

CCI

CCI

CCI

CCI

CCI

Table 3 CCI

CCI

CCI

3 CHANGES TO PROTOCOL PLANNED ANALYSES

The following change to planned analysis from protocol will be made:

- CCI

4 DATA ANALYSIS CONSIDERATIONS

4.1 Timing of Analyses

The primary analysis for each cohort will be performed independently. If the enrollment completion dates for both cohorts are at least 3 months apart, to avoid delay of primary analysis of the earlier completed cohort, separate database locks will be performed. The final analysis for each cohort will occur when the total duration is 2.5 years and the last patient for the respective cohort is dosed and followed for at least 6 months.

An interim analysis for Cohort 2 will be performed once approximately 50% (125 patients) have been enrolled into this cohort.

4.2 Analysis Populations

Full Analysis Set (FAS)

The FAS is defined as all enrolled patients that received at least 1 dose of study intervention. The FAS will be used for all efficacy, patient reported outcomes (PROs), and study population outputs, unless otherwise specified.

Full Analysis Set for ORR (non-final) ICR analyses (FAS-ORRc)

The FAS-ORRc is a subset of patients in the FAS having at least one ICR post-baseline RECIST tumor assessment at the time of DCO. This population will be considered for all TLFs deliveries up to but not including the final analysis in ORR related analyses so only patients who have the possibility to achieve an un-confirmed RECIST response are assessed.

Full Analysis Set for ORR (non-final) investigator analyses (FAS-ORRi)

The FAS-ORR-CNSi is a subset of patients in the FAS having at least one investigator post-baseline RECIST tumor assessment at the time of DCO. This population will be considered for all TLFs deliveries up to but not including the final analysis in ORR related analyses so only patients who have the possibility to achieve an un-confirmed RECIST response are assessed.

Full Analysis Set for CNS ORR (non-final) ICR analyses (FAS-ORR-CNSc)

The FAS-ORR-CNSc is a subset of patients in the FAS having at least one ICR post-baseline CNS RECIST tumor assessment at the time of DCO. This population will be considered for all TLFs deliveries up to but not including the final analysis in ORR related analyses so only patients who have the possibility to achieve an un-confirmed CNS RECIST response are assessed.

Full Analysis Set for CNS ORR (non-final) investigator analyses (FAS-ORR-CNSi)

The FAS-ORRi is a subset of patients in the FAS having at least one investigator post-baseline CNS RECIST tumor assessment at the time of DCO. This population will be considered for all TLFs deliveries up to but not including the final analysis in ORR related analyses so only patients who have the possibility to achieve an un-confirmed CNS RECIST response are assessed.

Full Analysis Set for subsequent treatment (FAS-ST)

The FAS-ST is a subset of patients in the FAS who had RECIST progression, as assessed by investigator, during the study and received subsequent anti-cancer therapy. This population will be used in analysis of subsequent treatments only.

Full Analysis Set for isolated CNS progression analyses (FAS-CNS)

The FAS-CNS is a subset of patients in the FAS who develop isolated CNS progression as assessed by investigator, receive local therapy, and continue on protocol therapy. This population will be used in analyses related to the time to next progression and site of next progression.

Safety Analysis Set (SAF)

The SAF will consist of all enrolled patients who received at least 1 dose of study intervention. Safety data will not be formally analyzed but summarized using the SAF.

Safety Analysis Set for ILD/pneumonitis patients (SAF-ILD)

The SAF-ILD is a subset of patients in the SAF who developed ILD/pneumonitis. This population will be used in analyses related to ILD/pneumonitis.

A summary on which analysis set will be used for each outcome variable is provided in Table 4.

Table 4 Summary of Outcome Variables and Analysis Sets

Outcome Variable	Analysis Set
Efficacy Data	
RECIST 1.1 related outcome variables (PFS [in patients with BM at baseline], DoR, time to progression, PFS2, tumor size)	FAS
ORR for non-final ICR analysis	FAS-ORRc
ORR for non-final investigator analysis	FAS-ORRi
ORR for final analysis	FAS
OS	FAS

Table 4 Summary of Outcome Variables and Analysis Sets

Outcome Variable	Analysis Set
CCI	FAS
DoT on subsequent lines of therapy	FAS-ST
Incidence of new symptomatic CNS metastasis during study intervention for patients without BM at baseline, CCI (Cohort 1)	FAS
In patients who develop isolated CNS progression, receive local therapy, and continue on protocol therapy: - Time to next progression (CNS or extracranial) or death - Site (CNS vs extracranial vs both) of next progression	FAS-CNS
In patients with BM at baseline (Cohort 2), CNS RECIST 1.1 related outcome variables (CNS PFS, time to new CNS lesions, and CNS DoR)	FAS
CNS ORR for non-final ICR analysis	FAS-ORR-CNSc
CNS ORR for non-final investigator analysis	FAS-ORR-CNSi
CNS ORR for final analysis	FAS
CCI	FAS
Health Related Quality of Life Data	
EORTC QLQ-C30, NANO scale, cognitive tests, MDASI brain tumor-specific items (Cohort 2), compliance to health-related quality of life data	FAS
CCI SGRQ-I	SAF-ILD
Study Population/Demography Data	
Disposition of patients	All subjects
Demography	FAS
Baseline and disease characteristics	FAS
Important protocol deviations	FAS
Medical/surgical history	FAS
Previous anti-cancer therapy	FAS
Concomitant medications/procedures	FAS
Subsequent anti-cancer therapy	FAS
CCI	
CCI	FAS
Safety Data	
Exposure	SAF
AEs	SAF

Table 4 Summary of Outcome Variables and Analysis Sets

Outcome Variable	Analysis Set
Rate of AEs among patients with baseline BM who are treated with concurrent high-dose steroid	SAF
Deaths	SAF
Rate of ILD/pneumonitis	SAF
Laboratory measurements	SAF
Vital signs	SAF
ECGs	SAF
Echocardiogram/MUGA	SAF
SpO2 and pulmonary assessments	SAF
ECOG performance status	SAF
ILD Data	
ILD/pneumonitis investigation	SAF-ILD
CCI	SAF-ILD

AE = adverse event; BM = brain metastasis; CCI CNS = central nervous system; DoR = duration of response; DoT = duration of treatment; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organization for the Research and Treatment of Cancer; FAS = Full Analysis Set; ILD = interstitial lung disease; MDASI = MD Anderson Symptom Inventory; MUGA = multigated acquisition scan; NANO = Neurologic Assessment in Neuro-oncology; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; PFS2 = time to second progression or death; QLQ-C30 = 30 item core quality of life questionnaire; CCI CCI RECIST 1.1 = Response Evaluation Criteria in Solid Tumors Version 1.1; SAF = Safety Analysis Set; SGRQ-I = St. George's Respiratory Questionnaire – idiopathic pulmonary fibrosis version; SpO2 = peripheral capillary oxygen saturation; T-DXd = trastuzumab deruxtecan; CCI CCI

4.3 General Considerations

No formal hypothesis testing is planned for this study.

All analyses will be presented separately for each cohort.

The below mentioned general principles will be followed throughout the study:

- Descriptive statistics will be used for all variables, as appropriate: Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum and maximum. Categorical variables will be summarized by frequency counts and percentages for each category.

- Unless otherwise stated, percentages will be calculated out of the population total for the corresponding analysis set.

- For continuous data, the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2 additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.

- For categorical data, percentages will be presented to 1 decimal place.

- Confidence intervals (CIs) calculated from Kaplan-Meier (KM) plots and CIs on rates will be presented to 1 decimal place.

- All CIs presented will be at the 95% level.

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

- For subgroup analyses, if the subgroup/category contains fewer than 10 patients, point estimates (for example median, mean, percentages) and CIs will not be presented.

- For survival endpoints, median will not be calculated if 50% of patients have not had an event. Quartiles will not be calculated if 25/75% of patients have not had an event.

- SAS® version 9.4 or higher will be used for all analyses.

4.3.1 General Study Level Definitions

In general, the last observation before the first dose of study intervention will be considered the baseline unless otherwise specified. Assessments on the day of first dose will be considered prior to first dose if such procedures are required by the CSP to be conducted before the first dose.

In all summaries change from baseline variables will be calculated as the post-intervention value minus the value at baseline. The percentage change from baseline will be calculated as (post-baseline value – baseline value) / baseline value x 100.

Study Day 1 is defined as the date of first dose of study intervention (Cycle 1 Day 1). For visits (or events) that occur on or after first dose of study intervention, Study Day is derived as (date of visit [event] – date of first dose of study intervention + 1). For visits (or events) that occur prior to first dose, study day is defined as (date of visit [event] – date of first dose of study intervention). There is no Study Day 0 defined for this study.

For listings (such as for adverse events [AEs]) that include the derivation of "days since last dose", this is defined as (event date – date of last dose) where "date of last dose" is defined

as the date of dosing immediately preceding the event occurrence. Events that occur on the same day as the last dose of study intervention will therefore be described as occurring 0 days from last dose of study intervention.

Listings will only include actual data i.e. imputed data will not be included.

Study periods will be defined as follows:

- Pre-treatment: Any result prior to first dose of study intervention.
- On-treatment: Any result from first dose of study intervention up to 47 days after the day of last study intervention.
- Survival follow-up: Any result 48 days or more after the day of last study intervention.

4.3.2 Visit Windows

Visit windows will be defined for any outputs that summarize values by visit. The following conventions will apply:

- Visit windows will be exhaustive so that data recorded at any timepoint has the potential to be summarized. Inclusion within the visit window will be based on the actual date and not the intended date of the visit.
- The window for the visits following baseline will be constructed in such a way that the upper limit of the interval falls halfway between the 2 visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between 2 consecutive visits, then the upper limit will be taken as the midpoint value minus 1 day. For example, the visit windows for vital signs data up to Cycle 5 Day 1 are:
  - Cycle 1 Day 8 (Study Day 8): 2 - 11
  - Cycle 1 Day 15 (Study Day 15): 12 - 18
  - Cycle 2 Day 1 (Study Day 22): 19 – 32
  - Cycle 3 Day 1 (Study Day 43): 33 – 53
  - Cycle 4 Day 1 (Study Day 64): 54 – 74
  - Cycle 5 Day 1 (Study Day 85): 75 – 95
- Visit 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 a scheduled visit based on the study day the end of treatment visit occurred at.

- For summaries showing the maximum or minimum values, the maximum/minimum values recorded while on study intervention will be used (regardless of where it falls in an interval).

- Listings should display all values contributing to a timepoint for a patient.

- If there are more than 1 value per patient within a visit window then the closest value to the scheduled visit date will be summarized, or the earlier in the event the values are equidistant from the nominal visit date. If there is still a tie, preference to the nominal visit at which the results were reported at will be given. The listings will highlight the value for the patient that contributed to the summary table wherever feasible.

- If multiple measurements are taken on the same study day then the reason for multiple measurements will be queried with the site. Only the first measurement will be considered for the scheduled visit, unless there is evidence that the later measurements were conducted due to an issue with the first assessment, where the later measurement will be used instead.

- Note that values which are not summarized will still be listed.

4.3.3 Handling of Unscheduled Visits

All unscheduled visit data have the potential to be included in the summaries, based on the assignment of visit using visit windowing as described in Section 4.3.2. Data from all visits will be considered for summaries showing the maximum or minimum values recorded while on study intervention.

4.3.4 Multiplicity/Multiple Comparisons

Not applicable.

4.3.5 Handling of Missing Data

Missing safety data will generally not be imputed. However, safety assessments of the form "<x" (i.e., below the lower limit of quantification) or ">x" (i.e., 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.

AEs that have missing causality (after data querying) will be assumed to be related to study intervention. Missing Common Terminology Criteria for AEs (CTCAE) grades will not be imputed.

Furthermore:

- For partial or missing AE or medication start dates the following will be applied:
  - Missing day: Impute the 1st day of the month unless month is the same as month of first dose of study intervention then impute first dose date. If this results in a start date after the end date, then the day will be imputed with the 1st day of the month.
  - Missing day and month: Impute 1st January unless year is the same as first dose date then impute first dose date. If this results in a start date after the end date, then the day and month will be imputed with 1st January.
  - Completely missing: Impute first dose date unless the end date indicates it started prior to the first dose date, in which case do not impute.
- For partial or missing AE or medication end dates, the following will be applied:
  - Missing day: Impute with the last day of the month.
  - Missing day and month: Impute 31st December.
  - Completely missing: Do not impute.
- Flags will be retained in the datasets indicating where any programmatic imputation has been applied, and in such cases, any durations would not be calculated.

4.3.6 Handling of Protocol Deviations in Study Analysis

Important protocol deviations (IPDs) are defined as a subset of protocol deviations that may significantly affect the completeness, accuracy, and/or reliability of the study data or that may significantly affect a patient's rights, safety, or well-being.

A list of all protocol deviations, including those reported by monitors, will be reviewed and decisions regarding how to handle these deviations will be documented by the study team physician, clinical pharmacology scientist and statistician prior to database lock. The final classification will be made prior to database lock. Important protocol deviations and any action to be taken regarding the exclusion of participants or affected data from specific analyses are defined in the project-specific Protocol Deviation Specification.

The important protocol deviations will be listed and summarised.

A per-protocol analysis excluding patients with specific important protocol deviations is not planned, however a "deviation bias" sensitivity analysis may be performed on the primary endpoints excluding patients with deviations that may affect the efficacy of the study intervention if more than 10% of patients had any significant deviation deemed to affect the primary endpoint. The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock and will be documented prior to the analysis being conducted.

4.3.7 Derivation of RECIST Visit Responses

For all patients, the RECIST tumor response data will be used to determine each patient's visit response according to RECIST Version 1.1. It will also be used to determine if and when a patient has progressed in accordance with RECIST 1.1 and their best objective response (BoR) to study intervention.

Baseline radiological tumor assessments will be performed no more than 28 days before the start of study intervention and ideally as close as possible to the start of study intervention. Tumor assessments are then performed every 6 weeks (± 1 week) from Cycle 1 Day 1 until Week 48 and every 9 weeks (± 1 week) thereafter starting at Week 57 until RECIST 1.1-defined radiological progression of disease (PD). All patients in this study will receive an intravenous (IV) contrast-enhanced magnetic resonance imaging (MRI) (preferred, unless contraindicated) or IV contrast-enhanced computed tomography (CT) of the brain at screening/baseline. Mandatory follow-up scans are required thereafter at the same time as tumor assessments in the rest of the body, or as needed based on clinical neurological symptoms for all patients with baseline stable BM, while patients without BM do not need additional brain scans for subsequent tumor assessments unless clinically indicated.

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

Assessments per RECIST version 1.1 by independent central review (ICR) will be used for the primary analysis of this study. Results from the investigator's review of imaging scans will be presented as a sensitivity analysis.

From the investigator's review of the imaging scans, the RECIST tumor response data will be used to determine each patient's visit response according to RECIST version 1.1. At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of complete response (CR), partial response (PR), stable disease (SD), or progression of disease (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 tumor assessment that cannot be evaluated then the patient will be assigned a visit response of not evaluable (NE), unless there is evidence of progression in which case the response will be assigned as PD.

Refer to Table 5 for the definitions of CR, PR, SD and PD.

RECIST outcomes (i.e., PFS, ORR, etc.) will be calculated programmatically for the site investigator data from the overall visit responses.

4.3.7.1 Independent Review

A planned ICR of all radiological imaging data will be carried out using RECIST version 1.1. All radiological scans for all patients (including those at unscheduled visits, or outside visit windows) will be collected on an ongoing basis and sent to an AstraZeneca appointed Contract Research Organization for central analysis. The imaging scans will be reviewed by 2 independent radiologists using RECIST 1.1 and will be adjudicated, if required (i.e., 2 reviewers review the scan and adjudication is performed by a separate reviewer in case of disagreement). For each patient, the ICR will define the overall visit response (i.e., the response obtained overall at each visit by assessing TLs, NTLs and new lesions) data and no programmatic derivation of visit response is necessary. (For patients with TLs at baseline: CR, PR, SD, PD, NE; for patients with NTLs only: CR, SD, PD, NE). If a patient has had a tumor assessment that cannot be evaluated, then the patient will be assigned a visit response of NE (unless there is evidence of progression in which case the response will be assigned as PD). RECIST assessments/scans contributing towards a particular visit may be performed on different dates and for the central review the date of progression for each reviewer will be provided based on the earliest of the scan dates of the component that triggered the progression.

If adjudication is performed, the reviewer that the adjudicator agreed with will be selected as a single reviewer (note in the case of more than 1 review period, the latest adjudicator decision will be used). In the absence of adjudication, the records for all visits for a single reviewer will be used. The reviewer selected in the absence of adjudication will be the reviewer who read the baseline scan first. The records from the single selected reviewer will be reported for all ICR RECIST information including dates of progression, visit response, censoring and changes in target lesion dimensions. Endpoints (of ORR, PFS and duration of response [DoR]) will be derived programmatically from this information.

Results from this independent review will not be communicated to the investigators and the management of patients will be based solely upon the results of the RECIST 1.1 assessment conducted by the investigator.

An ICR of all patients will be performed for the final database lock, which will cover all the scans up to the DCO.

Further details of the ICR will be documented in the ICR Charter.

4.3.7.2 Target Lesions (TLs) – Site Investigator Data

Measurable disease is defined as having at least 1 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 CT or MRI and which is suitable for accurate repeated measurements. A patient can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (including lymph nodes collectively considered as a single organ), representative of all lesions involved and these are referred to as TLs. TLs should be selected on the basis of their size (longest diameter for non-nodal lesions or short axis diameter for nodal lesions), but in addition should be those that lend themselves to reproducible repeated measurements. It may be 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, should be selected. If more than 1 baseline scan is recorded, then the measurements from the scan that is closest and prior to first administration of study intervention will be used to define the baseline sum of TLs.

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

For patients who do not have measurable disease at entry (i.e., no TLs), but have non-measurable disease (for this study defined as non-measurable central nervous system [CNS] disease, or non-measurable, bone-only disease that can be assessed by CT, MRI or X-Ray [lytic or mixed lytic bone lesions that can be assessed by CT, MRI or X-Ray are acceptable; sclerotic/osteoblastic bone lesions are not eligible]), evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions. If a patient does not have measurable disease at baseline, then the TL visit response will be not applicable (NA).

Table 5 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 < 10 mm.
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.
Progression of disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5 mm, taking as reference the smallest

Table 5 TL Visit Responses (RECIST 1.1)

Visit Responses	Description
	sum of diameters since study intervention 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 and scaling up could not be performed for lesions with interventions). Note: if the sum of diameters meets the PD criteria, PD overrides not evaluable as TL response.
Not applicable (NA)	No TLs are recorded at baseline.

CR = complete response, NA = not applicable, NE = not evaluable, PD = progression of disease, PR = partial response, SD = stable disease, TL = target lesion.

Rounding of TL data

For calculation of PD and PR for TLs percentage change from baseline and previous minimum should be rounded to 1 decimal place before assigning a TL response. For example: 19.95% should be rounded to 20.0% but 19.94% should be rounded to 19.9%.

Missing TL data

For patients with measurable disease at entry, a visit can only be considered evaluable if all TL measurements have been recorded. However, a visit response of PD should still be assigned if any of the following occurred:

- A new lesion is recorded
- A NTL visit response of PD is recorded
- The sum of TLs sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5 mm from nadir, even assuming the non-recorded TLs have disappeared. The only exception here is if the previous visit response was CR (refer to below to below section “TL visit responses subsequent to CR”)

The nadir (i.e., smallest measurement) can only be taken from assessments where all the TLs had a LD recorded.

Lymph nodes

For lymph nodes, if the size reduces to < 10 mm then these are considered non-pathological. However, a size will still be given, and this size should still be used to determine the TL visit response as normal. In the special case where all lymph nodes are

< 10 mm and all other TLs are 0 mm then although the sum may be > 0 mm the calculation of TL response should be over-written as a CR.

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:

- Step 1: If all lesions meet the CR criteria (i.e., 0 mm or < 10 mm for lymph nodes) then response will be set to CR irrespective to whether the criteria for PD or TL is also met, i.e., if a lymph node LD increased by 20% but remains < 10 mm.
- Step 2: If some lesion measurements are missing but all other non-missing lesions meet the CR criteria (i.e., 0 mm or < 10 mm for lymph nodes) then response will be set to NE irrespective of whether, when referencing the sum of TL diameters, the criteria for PD are also met.
- Step 3: If not all lesions 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) or a new lesion appears, then response will be set to PD.
- Step 4: If after steps 1 – 3 a response can still not be determined the response will be set to remain as CR.

TL too big to measure

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

TL too small to measure

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

Irradiated lesions/lesion intervention

Previously irradiated lesions (i.e., lesion irradiated prior to entry into the study) should be recorded as NTLs and should not form part of the TL assessment.

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation/ palliative surgery/ embolisation), should be handled in the following way. Once a lesion has had intervention then it should be treated as having intervention for the remainder of the study noting that an intervention will most likely shrink the size of tumors:

- 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 ≤ 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 patient will be assigned a TL response of PD.
- Step 3: if, after both steps, PD has not been assigned, then, if appropriate (i.e., if ≤ 1/3 of the TLs have missing measurements), the scaled sum of diameters in step 2 should be used, and PR or SD then assigned as the visit response. Patients with intervention are evaluable for CR as long as all non-intervened lesions are 0 mm (or < 10 mm for lymph nodes) and the lesions that have been subject to intervention have a value of 0 mm (or < 10 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 to NE.

At subsequent visits, the above steps will be repeated to determine the TL and overall visit response. When calculating the previous minimum, lesions with intervention should be treated as missing and scaled up (as per step 2 above).

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 still increased by 20% or more compared to nadir and the sum of TLs has increased by ≥ 5 mm from nadir).

If ≤ 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 estimate TL sum of 81 mm:

68 x 74 / 62 = 81 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 ≤ 1/3 lesion assessments not recorded, the scaled-up sum of TLs diameters will be included when defining the nadir value for the assessment of progression.

Lesions that split in two

If a TL splits in 2, then the lesion diameters of the split lesions should be summed and reported as the lesion diameter for the lesion that split.

Lesions that merge

If 2 TLs merge, then the lesion diameter of the merged lesion should be recorded for 1 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.

4.3.7.3 Non-Target Lesions (NTLs) and New Lesions – Site Investigator Data

At each visit, the investigator should 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 investigation site at each visit.

NTL response will be derived based on the investigator's overall assessment of NTLs as described in Table 6.

Table 6 NTL Visit Responses (RECIST 1.1)

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).
Progression of disease (PD)	Unequivocal progression of existing NTLs. Unequivocal progression may be due to an important progression in 1 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 1 or more NTLs with no evidence of progression.
Not evaluable (NE)	Only relevant when 1 or more 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.
Not applicable (NA)	Only relevant if there are no NTLs at baseline.

CR = complete response, NA = not applicable, NE – not evaluable, NTL = non-target lesion, PD = progression of disease, TL = target lesion.

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 tumor burden has increased sufficiently to merit a determination of disease progression. A modest 'increase' in the size of 1 or more NTLs is usually not sufficient to qualify for unequivocal progression status.

Details of any new lesions will also be recorded with the date of assessment. The presence of 1 or more new lesions is assessed as progression.

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

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 tumor.

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

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

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.

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

Patients with 'symptomatic progression' requiring discontinuation of study intervention without objective evidence of disease progression at that time should continue to undergo tumor assessments where possible until objective disease progression is observed.

4.3.7.4 Overall Visit Response – Site Investigator Data

Table 7 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 7 Overall Visit Responses (RECIST 1.1)

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 (non-CR/non-PD)**
NA	NE	No (or NE)	**NE**

Non-CR/Non-PD for overall response if only NTL (no TLs) are present at baseline.
CR = complete response; NA = not applicable (only relevant if there were no TLs at baseline or NTLs at baseline); NE = not evaluable; NTL = non-target lesion; PD = progression of disease; PR = partial response; SD = stable disease; TL = target lesion.

The following overall visit responses are possible depending on the extent of tumor disease at baseline:

- For patients with TLs (at baseline): CR, PR, SD, PD, or NE
- For patients with NTLs only (at baseline): CR, Non-CR/Non-PD, PD, or NE

4.3.8 Derivation of CNS RECIST Visit Responses

The assessments of CNS RECIST will mimic those of RECIST from Section 4.3.7 but only including up to 5 TLs within the CNS, selected on the basis of their size and the suitability for repeated measurements. These may include overlapping or different TLs than the ones from RECIST.

Further details can be found in the CNS RECIST 1.1 and CCI Independent Review Charter.

4.3.9 CCI

CCI

CCI

CCI

CCI

CCI

CCI

Table 8 Response Criteria for CNS Metastases per CCI

	CR	PR	SD	PD
TL [d]	CCI			
NTL [d]				
NL [b, d]				
CCI				
CCI				
CCI				

CNS = Central nervous system; CR = complete response; LD = longest diameter; NE = not evaluable; NL = new lesion; NTL = non-target lesion; PD = progression of disease; PR = partial response; SD = stable disease; TL = target lesion.

[a] Progression occurs when this criterion is met.

[b] CCI

[c] CCI

[d] Results as provided by central readers.

[e] If one of the required variables is missing, then the response will be NE.

The overall visit response is derived and provided by the ICR.

Further details can be found in the CNS RECIST 1.1 and CCI Independent Review Charter.

5 STATISTICAL ANALYSIS

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

5.1 Study Population

The domain study population covers patient disposition, analysis sets, protocol deviations, demographics, baseline characteristics, disease characteristics, medical history, and prior

and concomitant medication. All study population outputs will be presented on the FAS unless otherwise indicated.

5.1.1 Patient Disposition and Completion Status

5.1.1.1 Definitions and Derivations (Patient Disposition and Completion Status)

A clear account of the disposition of all patients who enter the study will be provided, from screening to study completion. A patient is considered to have completed the study when he/she has completed all phases of the study including the last visit.

Patients who die during the study or withdraw from study will be considered as non-completers. Patients ongoing at study closure will be considered completers.

5.1.1.2 Presentation (Patient Disposition and Completion Status)

Patient disposition will be listed and summarized for all subjects. Summaries will include the number and percentage of patients:

- Enrolled (informed consent received)
- Treated
- Ongoing study intervention at DCO
- Ongoing study at data cut-off (DCO)
- Completed study

In addition, the number and percentage of patients who discontinued study intervention and who discontinued the study, including a breakdown of the reason for discontinuation will be presented for all patients. Discontinuation of study intervention and/or withdrawal from study due to COVID-19 will be presented as a separate category.

The number of patients in the FAS population for each country and each center will be presented.

The number and percentage of patients with confirmed or suspected COVID-19 infection will be presented separately, including details on COVID-19 related interruptions impacting on visits and investigational product administration. Listings of patients affected by the COVID-19 pandemic will be presented detailing any affect and impact on the study. Issues reported in the Clinical Trial Management System will be considered for presented in listings as well.

5.1.2 Analysis Sets

5.1.2.1 Definitions and Derivations (Analysis Sets)

Analysis sets are defined in Section 4.2.

5.1.2.2 Presentation (Analysis Sets)

The number of patients included in each analysis set will be summarized.

Reasons for exclusions from any analysis set will be summarized in tables and provided in listings.

5.1.3 Protocol Deviations

5.1.3.1 Definitions and Derivations (Protocol Deviations)

Protocol deviations are defined in Section 4.3.6.

5.1.3.2 Presentation (Protocol Deviations)

The important protocol deviations will be listed and summarized, including:

- Number of patients with at least 1 important protocol deviation
- Number of patients with at least 1 COVID-19 important protocol deviation
- Number of patients with at least 1 important protocol deviation, excluding COVID-19 deviations

These will be further summarized into categories, e.g. failed to meet inclusion criteria, met exclusion criteria, other protocol deviations, etc.

5.1.4 Demographics

5.1.4.1 Presentation (Demographics)

- Demographics will be summarized descriptively and listed for all patients in the FAS including age (years) (as entered on CRF), age group (< 65 and ≥ 65 years), sex, race and ethnicity. Demographics will further be summarized separately for all patients in the FAS who had confirmed or suspected COVID-19 infection, and patients in cohort 2 with stable vs. active (untreated or progressing) BM at baseline, as defined in 5.2.4.4.

5.1.5 Baseline Characteristics

5.1.5.1 Definitions and Derivations (Baseline Characteristics)

Body mass index (BMI) will be derived as:

$$\text{BMI}\left(\frac{\text{kg}}{\text{m}^2}\right) = \frac{\text{weight (kg)}}{\text{height (m)}^2}$$

5.1.5.2 Presentation (Baseline Characteristics)

Patient characteristics at baseline (height, weight and BMI) will be summarized and listed for all patients in the FAS. Categorized alcohol, drug, caffeine and nicotine usage (never, current, former), and number of pack-years will also be listed. Patient characteristics will further be summarized separately for all patients in the FAS who had confirmed or suspected COVID-19 infection.

5.1.6 Disease Characteristics

5.1.6.1 Presentation (Disease Characteristics)

Disease characteristics will be summarized and listed for all patients in the FAS including:

- Previous disease related treatment modalities (immunotherapy, hormonal therapy, cytotoxic chemotherapy, systemic therapy, radiotherapy, and other)
- Disease characteristics at baseline (Eastern Cooperative Oncology Group [ECOG] performance status, primary tumor location, histology type, tumor grade and time from diagnosis to first dose of study intervention)
- Extent of disease at baseline (metastatic or locally advanced disease)
- Tumor, node and metastasis (TNM) classification and American Joint Committee on Cancer (AJCC) stage at time of diagnosis
- Biomarkers results (i.e. ER, PR and HER2 status)

5.1.7 Medical and Surgical History

5.1.7.1 Definitions and Derivations (Medical History)

Medical and surgical history will be coded using Medical Dictionary for Regulatory Activities (MedDRA).

5.1.7.2 Presentation (Medical History)

All medical history will be listed, and the number and percentage of patients will be summarized for the FAS by system organ class (SOC) and preferred term (PT). A separate summary will be presented for patients who had confirmed or suspected COVID-19 infection.

A listing on surgical histories will also be presented.

5.1.8 Prior, Concomitant and Post Medications

5.1.8.1 Definitions and Derivations (Prior, Concomitant and Post Medications)

Treatments received prior to, concomitantly or post-study intervention will be coded using the World Health Organization – Drug Dictionary (WHODD) Anatomical Therapeutic Chemical (ATC) classification codes.

For the purpose of inclusion in prior and/or concomitant medication or therapy summaries, incomplete dates will be imputed as detailed in Section 4.3.5.

Medications will be categorized into multiple mutually exclusive categories. Prior medications, concomitant medications and post-study intervention medications will be defined based on imputed start and stop dates as follows:

- Prior medications are those taken prior to study intervention with a stop date prior to the first dose of study intervention
- Concomitant medications are those with a stop date on or after the first dose of study intervention (and could have started prior to or during study intervention), or with a start date within the 28 days after last dose of study intervention
- Post-study medications are those with a start date at least 29 days after the last dose of study intervention

5.1.8.2 Presentation (Prior, Concomitant and Post Medications)

The following summaries will be produced:

- Previous cancer therapies prior to this study
- Disallowed concomitant medications (as identified during physician review described in Section 4.3.6, deviation 4)
- Allowed concomitant medications (all medications not identified as being disallowed)
- Cancer treatments and therapies subsequent to T-DXd
- Number of regimens of previous anti-cancer therapies in metastatic breast cancer

All medication data will be listed.

Missing coding terms will be listed and summarized as “Not coded”.

5.2 Efficacy Analyses

This section covers details related to the endpoint analyses such as primary, secondary, exploratory and additional endpoints including sensitivity and supportive analyses. All efficacy analyses will be conducted using the FAS unless otherwise stated.

5.2.1 Primary Endpoint for Cohort 1 – ORR by RECIST 1.1 per ICR

The primary endpoint for Cohort 1 is ORR by RECIST 1.1 per ICR, and will be analyzed for all patients in Cohort 1 in the FAS.

5.2.1.1 Definitions and Derivations (ORR by RECIST 1.1 per ICR)

ORR is defined as the percentage of patients who have a confirmed CR or confirmed PR, as determined per ICR by RECIST 1.1.

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.

Data obtained up until progression, or last evaluable assessment in the absence of progression, will be included in the assessment of ORR, regardless of whether the patient withdraws from therapy. Patients who stop study intervention without a response or progression, receive a subsequent anti-cancer therapy, and then respond will not be included as responders in the ORR (i.e., both visits contributing to a response must be prior to subsequent therapy for the patient to be considered as a responder).

In the case where a patient has 2 non-consecutive visits responses of PR, then as long as the time between the 2 visits of PR is greater than 4 weeks and there are no PD between the PR visits, the patient will be 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 will be assigned.

Best Objective Response (BoR)

As a supportive endpoint for ORR, BoR according to RECIST 1.1 will be defined. The denominators will be consistent with those used in the ORR analysis.

BoR is calculated based on the overall visit responses from each RECIST assessment. It is the best response a patient has had following first treatment of study intervention, but prior to starting any subsequent anticancer therapy and up to and including RECIST progression or the last evaluable assessment in the absence of RECIST progression.

Categorization of BoR will be based on RECIST using the following response categories: CR, PR, SD, PD and NE.

Confirmation of CR or PR is required for determination of response. For ongoing analyses, unconfirmed response will also be provided.

For determination of a best response of SD, the earliest of the dates contributing towards a particular visit will be used. SD should be recorded at least 6 weeks minus 1 week. i.e., at least 35 days (to allow for an early assessment within the assessment window), after first dose of study intervention. For CR/PR, the initial overall visit assessment that showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

For patients who die with no evaluable RECIST assessment, if death occurs ≤ 13 weeks (i.e., 2 x 6 weeks + 1 week for a late assessment within the assessment window) after first dose of study intervention, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST assessment, if death occurs > 13 weeks after first dose of study intervention then BoR will be assigned the category of NE.

5.2.1.2 Primary Analysis of Endpoint (ORR by RECIST 1.1 per ICR)

The ORR will be based on ICR RECIST 1.1 data as primary analysis. All scans regardless of whether they were scheduled or not will be used in the analysis.

Summaries will be produced that present the number and percentage of patients with a confirmed tumor response (CR/PR). The estimate of ORR and its 95% CI using a normal approximation will be provided. Additionally, ORR estimate and its 95% CI at selected intervals (e.g., 6, 9, 12 months (i.e. every 3 months starting from 6-months)) from the first date of dosing will be summarized.

The BoR will be summarized by number and percentage of patients in each category (CR, PR, SD, PD and NE).

5.2.1.3 Sensitivity Analyses of Endpoint (ORR by RECIST 1.1 per ICR)

To assess the sensitivity of the results, the ORR analysis will be repeated on the site investigator RECIST 1.1 data.

5.2.1.4 Subgroup Analyses (ORR by RECIST 1.1 per ICR)

Subgroup analyses will be conducted in the following subgroups of the FAS:

CCI

The subgroup analysis will be based on values as recorded on the CRF or from the third -party vendor data.

Other baseline variables may also be assessed if there is clinical or biological justification.

Subgroup analysis will be done by repeating the primary analysis separately for each component within each subgroup.

In addition, the ORR estimate with its 95% CI will be presented graphically on a forest plot for each component of each subgroup.

5.2.2 Primary Endpoint for Cohort 2 – PFS by RECIST 1.1 per ICR

The primary endpoint for Cohort 2 is PFS by RECIST 1.1 per ICR and will be analyzed for all patients in Cohort 2 in the FAS.

5.2.2.1 Definitions and Derivations (PFS by RECIST 1.1 per ICR)

PFS is defined as the time from first dose of study intervention until date of objective disease progression per RECIST 1.1 or death (by any cause in the absence of progression) regardless of whether the patient withdraws from therapy, receives another anticancer therapy prior to progression or clinically progresses prior to RECIST 1.1 progression (i.e., date of PFS event or censoring – date of first dose of study intervention + 1). Patients who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment. However, if the patient progresses or dies immediately after 2 or more consecutive missed visits, the patient will be censored at the time of the latest evaluable RECIST 1.1 assessment prior to the 2 missed visits (Note: NE visit is not considered a missed visit).

Given the schedule of RECIST 1.1 assessments (i.e., 6-weekly for the first 48 weeks then 9-weekly thereafter) the definition of 2 missed visits will change. If the previous RECIST assessment is less than study day 288 (i.e., week 41) then 2 missing visits will equate to 14 weeks since the previous RECIST assessment, allowing for early and late visits (i.e., 2 x 6 weeks + 1 week for an early assessment + 1 week for a late assessment = 14 weeks). If the 2 missed visits occur over the period when the scheduled frequency of RECIST assessments change from 6-weekly to 9-weekly this will equate to 17 weeks (i.e., take the average of 6 and 9 weeks which gives 7.5 weeks and then apply same rationale, hence 2 x 7.5 weeks + 1 week for early assessment + 1 week for late assessment = 17 weeks). The time period for the previous RECIST assessment will be from study days 288 to 330 (i.e., week 41 to 47). From week 49 onwards (when scheduling changes to 9-weekly assessments), 2 missing visits will equate to 20 weeks (i.e., 2 x 9 weeks + 1 week for an early assessment + 1 week for a late assessment = 20 weeks).

If the patient has no evaluable visits or does not have baseline data, they will be censored at Day 1, unless they die within 2 visits of baseline (plus 1 week allowing for a late assessment with the visit window [13 weeks]).

The PFS time will always be derived based on scan/assessment dates and not visit dates.

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

- For ICR assessments, the date of progression will be determined based on the **earliest** of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or reviewer who read baseline first if there is no adjudication for ICR data.
- For investigator assessments, the date of progression will be determined based on the **earliest** of dates of the component that triggered the progression.
- For both ICR and investigator assessments, when censoring a patient for PFS the patient will be 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.

5.2.2.2 Primary Analysis of Primary Endpoint (PFS by RECIST 1.1 per ICR)

The treatment status at progression of patients at the time of analysis will be summarized. This will include the number (%) of patients who were on study intervention at the time of progression, the number (%) of patients who discontinued study intervention prior to progression, the number (%) of patients who have not progressed and were on study intervention or discontinued study intervention.

A KM plot of PFS will be presented. Summaries of the number and percentage of patients experiencing a PFS event, and the type of event (progression or death) will be provided along with the median PFS and respective CIs (calculated using the Brookmeyer and Crowley method) and the proportion of patients that are alive and progression-free at 3-months, 6-months, 9-months, 12-months, etc. (i.e. every 3 months starting from 3-months) calculated using the KM technique.

In addition, the number of patients prematurely censored will be summarized. A patient would be defined as prematurely censored if they have not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than 1 scheduled tumor assessment interval plus 2 weeks prior to the DCO date.

A KM plot of the time to censoring of PFS data will be presented similarly as described above but where the censoring indicator of the PFS analysis is reversed.

Additionally, summary statistics will be given for the number of weeks from censoring to DCO for all censored patients. The number and percentage of patients censored with a time from last tumor assessment greater than or lesser or equal to 1 scheduled tumor assessment

interval plus 2 weeks prior to the DCO date will be presented. A distinction will be made between those patients censored before and after week 48, as that is when the schedule of tumor assessments changes.

The duration of follow-up will be summarized using median time from date of first study intervention to date of censoring (date last known to have not progressed) in censored (not progressed) patients only.

Additionally, summary statistics for the number of weeks between the time of progression and the last evaluable RECIST assessment prior to progression will be presented.

Summaries of the number and percentage of patients who miss 2 or more consecutive RECIST assessments will be presented. The average days between RECIST assessment per patient will also be summarized.

A summary of location of CNS specific lesions will be presented.

All collected RECIST data will be listed for all patients. In addition, a summary of new lesions (i.e., site of new lesions) will be produced.

5.2.2.3 Sensitivity Analyses of the Primary Endpoint (PFS by RECIST 1.1 per ICR)

To assess the sensitivity of the results, the PFS analysis will be repeated on the site investigator RECIST 1.1 data.

A sensitivity analysis will be conducted to assess the potential impact of COVID-19 related deaths on PFS based on the ICR and the site investigator RECIST 1.1 data. That is, patients who had a PFS event due to death where the primary or secondary cause of death was COVID-19 infection or COVID-19 infection was reported as a fatal AE, will be censored at the last evaluable RECIST 1.1 assessment prior to COVID-19 infection related death.

5.2.2.4 Subgroup Analyses (PFS by RECIST 1.1 per ICR)

Subgroup analyses will be performed on the same subgroups defined in Section 5.2.1.4.

Subgroup analysis will be done by repeating the primary analysis separately for each component within each subgroup.

In addition, the PFS estimate with its 95% CI will be presented graphically on a forest plot for each component of each subgroup.

5.2.3 Secondary Endpoint – ORR by RECIST 1.1 for Cohort 2

ORR by RECIST 1.1 is defined as a secondary endpoint for patients in Cohort 2 and will be derived and analyzed for Cohort 2, the same way it was done for Cohort 1 as described in Section 5.2.1, and including all sensitivity and subgroup analyses as well.

5.2.4 Secondary Endpoint – Overall Survival

Overall survival (OS) is defined as a secondary endpoint for patients in both cohorts.

5.2.4.1 Definitions and Derivations (Overall Survival)

OS is defined as the time from the date of first dose of study intervention until death due to any cause regardless of whether the patient withdraws from study intervention or receives another anticancer therapy (i.e., date of death or censoring – date of first dose of study intervention +1). Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive.

Survival calls will be made in the week following the date of DCO for the analysis, and if patients are confirmed to be alive or if the death date is after the DCO date these patients will be censored at the date of DCO. The status of ongoing, withdrawn (from the study) and "lost to follow-up" patients at the time of 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 laws.

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 study intervention. The last date of each individual patient is defined as the latest among the following dates recorded on the CRF:

- AE start and stop dates
- Admission and discharge dates of hospitalization
- Study intervention date
- End of study intervention date
- Laboratory test date
- Date of vital signs
- Disease assessment dates on RECIST CRF
- Start and stop dates of alternative anti-cancer therapy
- Date last known to be alive on survival status CRF

- End of study 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 as the latest of either the last date known to be alive + 1 from the database or the death date using the available information after the below imputations:

- For missing day only – using the 1st of the month, unless the month and year is the same as the date of last visit, then use the date of last visit instead.
- For missing day and month – using the 1st of January, unless the year is the same as the date of last visit, then use the date of last visit instead.

If there is evidence of death but the date is entirely missing, it will be treated as missing, i.e., censored at the last known alive date.

5.2.4.2 Primary Analysis of Secondary Endpoint (Overall Survival)

A KM plot of OS will be presented. The proportion of patients alive will be presented for the 6-month, 9-month, 12-month, etc. (i.e. every 3 months starting from 6-months) milestones, besides their 95% CIs.

The number of patients prematurely censored will be summarized. A patient would be defined as prematurely censored if there is no indication that the patient has died, but there is no survival status available in the 10 weeks prior to the DCO.

Time from first dose of study intervention to the date of death (i.e., OS) or censoring (date last known to be alive) for censored patients will be summarized using median and respective 95% CIs (calculated using the Brookmeyer and Crowley method).

5.2.4.3 Sensitivity Analyses of the Secondary Endpoint (Overall Survival)

A sensitivity analysis will be conducted to assess the potential impact of COVID-19 related deaths on OS. That is, patients who had a death event where the primary or secondary cause of death was COVID-19 infection or COVID-19 infection was reported as a fatal AE, will be censored at the date of their COVID-19 infection related death.

5.2.4.4 Subgroup Analyses (Overall Survival)

For patients in Cohort 2, subgroup analyses will be performed on the following subgroup for the FAS:

CCI

Subgroup analysis will be done by repeating the primary analysis separately for each component within each subgroup.

In addition, the median OS with its 95% CI will be presented graphically on a forest plot for each component of each subgroup.

5.2.5 Secondary Endpoint – Duration of Response by RECIST 1.1 per ICR

DoR is defined as a secondary endpoint for patients in both cohorts.

5.2.5.1 Definitions and Derivations (Duration of Response by RECIST 1.1 per ICR)

DoR will be defined as the time from the date of first documented response (which is subsequently confirmed) until date of documented progression or death in the absence of disease progression (i.e., date of progression event or censoring – date of first response + 1). The end of response should coincide with the date of progression or death from any cause. The time of initial response will be defined as the latest of the dates contributing towards the first visit that was CR or PR that was subsequently confirmed.

If a patient does not progress following a response, then their DoR will use the PFS censoring time.

5.2.5.2 Primary Analysis of Secondary Endpoint (Duration of Response by RECIST 1.1 per ICR)

KM plots of DoR based on the ICR RECIST 1.1 assessment will be presented. Summary statistics about the DoR calculated from the KM curve will also be presented. Only patients who have a confirmed response will be included in the summary table. Swimmer plots that clearly show the profile of each patient who responds will also be produced.

5.2.6 Secondary Endpoint – Time to Progression by RECIST 1.1 per ICR

Time to progression is defined as a secondary endpoint for patients in both cohorts.

5.2.6.1 Definitions and Derivations (Time to Progression by RECIST 1.1 per ICR)

Time to progression by RECIST 1.1 per ICR is defined as the time from the date of first dose of study intervention to the date of documented disease progression. Patients who have not progressed at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment regardless of whether or not the patient died subsequent to this assessment.

5.2.6.2 Primary Analysis of Secondary Endpoint (Time to Progression by RECIST 1.1 per ICR)

Time to progression will be summarized (i.e., number of patients [%] that progressed). KM curves will be produced where patients without progressing will be censored. Median time

to progression with respective 95% CIs as calculated from the KM curve will also be presented.

5.2.6.3 Sensitivity Analysis of Secondary Endpoint (Time to Progression by RECIST 1.1 per ICR)

To assess the sensitivity of the results, the time to progression analysis will be repeated on the site investigator RECIST 1.1 data.

5.2.7 Secondary Endpoint – Duration of Treatment on Subsequent Lines of Therapy

Duration of treatment (DoT) on the subsequent lines of therapy after T-DXd is defined as a secondary endpoint for patients in both cohorts. It will be analyzed using all patients who started a subsequent line of therapy i.e. the FAS-ST. An additional analysis will be performed for patients who completed subsequent lines of therapy.

5.2.7.1 Definitions and Derivations (Duration of Treatment on Subsequent Lines of Therapy)

DoT of subsequent therapy will be defined as the time from the date of first dose of a subsequent therapy until date of the last dose of that therapy.

DoT will be calculated separately for each subsequent line after T-DXd until the latest line.

Any patient not known to have died or discontinued during a subsequent line of therapy will be censored on the last recorded date known to be alive for the respective line.

5.2.7.2 Primary Analysis of Secondary Endpoint (Duration of Treatment on Subsequent Lines of Therapy)

Only lines with 10 or more patients will be summarized. KM plots of DoT on subsequent therapy will be presented, by line. Median DoT and respective 95% CIs on subsequent therapy derived from the KM curve will also be summarized.

All lines will be listed.

5.2.8 Secondary Endpoint – Time from First Dose of Study Intervention to Second Progression or Death (PFS2)

Time from first dose of study intervention to second progression or death (PFS2) is defined as a secondary endpoint for patients in both cohorts.

5.2.8.1 Definitions and Derivations (Time from First Dose of Study Intervention to Second Progression or Death [PFS2])

PFS2 is defined as time from the first dose of study intervention until the date of tumor progression on next line treatment (the earliest of the progression event subsequent to first

subsequent anti-cancer therapy after the first progression) or death from any cause. The date of second progression according to site investigator RECIST 1.1 will be recorded in the CRF and defined according to local standard clinical practice. Following discontinuation of study intervention due to disease progression, as determined by the investigator according to RECIST 1.1 assessment, patients who started on subsequent cancer therapy post progression will continue to be followed at the 40 day (+ up to 7 days) follow-up visit and every 3 months (± 14 days) thereafter for documentation of progression on subsequent anticancer therapy. Patients alive and for whom a second disease progression has not been observed should be censored at the earliest of: date of study termination, date last known alive, DCO or if a patient has not had a first subsequent therapy, the date last known not to have received a first subsequent therapy. However, if the patient experiences a second progression or dies immediately after 2 or more consecutive missed visits, the patient will be censored at the time of the last PFS2 assessment prior to the 2 missed visits.

For patients who develop isolated CNS progression, receive local therapy, and continue on protocol therapy, second progression will correspond to the third progression according to site investigator RECIST 1.1 recorded in the CRF.

5.2.8.2 Primary Analysis of Secondary Endpoint (Time from First Dose of Study Intervention to Second Progression or Death [PFS2])

PFS2 will be analyzed using similar methods as outlined for PFS (refer to Section 5.2.2). Medians and KM plots will be presented to support the analysis.

The number and percentage of patients experiencing a PFS2 event will be summarized as well as summaries of death in the absence of second progression, and categories of PFS2 censoring. Time from first dose of study intervention to second progression will also be summarized.

A summary of the number of patients, events and corresponding percentages in PFS2 by line, by treatment after T-DXd and overall will be provided.

5.2.9 Secondary Endpoint – Incidence of New Symptomatic CNS Metastasis During Study Intervention

Incidence of new symptomatic CNS metastasis during study intervention is defined as a secondary endpoint for patients in Cohort 1.

5.2.9.1 Definitions and Derivations (Incidence of New Symptomatic CNS Metastasis During Study Intervention)

The incidence rate of new symptomatic metastases will be derived as

$$\frac{\text{Number of new symptomatic CNS metastases during study intervention period}}{\text{Total number of patients without symptomatic CNS at beginning of study}}.$$

5.2.9.2 Primary Analysis of Secondary Endpoint (Incidence of New Symptomatic CNS Metastasis During Study Intervention)

The number and percentage of patients observing new symptomatic CNS metastases will be summarized. The incidence rate together with its 95% CI based on exact methods (e.g. Poisson) will be presented.

5.2.10 Secondary Endpoint – Time to next progression (CNS or Extracranial) or Death – RECIST 1.1

Time to next progression (CNS or extracranial) or death is defined as a secondary endpoint for patients in both cohorts who develop isolated CNS progression, receive local therapy, and continue on protocol therapy. It will be analyzed for the FAS-CNS population, by cohort.

5.2.10.1 Definitions and Derivations (Time to Next Progression [CNS or Extracranial] or Death – RECIST 1.1)

Time to next progression is defined as the time from the date of the first documented isolated CNS progression to the date of the next documented disease progression (CNS or extracranial) per RECIST 1.1 by investigator assessment or death.

5.2.10.2 Primary Analysis of Secondary Endpoint (Time to Next Progression [CNS or Extracranial] or Death – RECIST 1.1)

Time to next progression (CNS or extracranial) per RECIST 1.1 by ICR or death will be summarized descriptively (i.e., number of patients [%] that progressed or died). KM curves will be produced with median time to next progression (CNS or extracranial) or death with respective 95% CIs presented as calculated from the KM curve.

5.2.11 Secondary Endpoint – Site (CNS vs Extracranial vs Both) of Next Progression – RECIST 1.1

Site (CNS vs extracranial vs both) of next progression is defined as a secondary endpoint for patients in both cohorts who develop isolated CNS progression, receive local therapy, and continue on protocol therapy. It will be analyzed for the FAS-CNS population, by cohort.

5.2.11.1 Definitions and Derivations (Site [CNS vs Extracranial vs Both] of Next Progression – RECIST 1.1)

The site of next progression in patients who have a subsequent documented disease progression (CNS or extracranial) per RECIST 1.1 by investigator assessment will be documented.

5.2.11.2 Primary Analysis of Secondary Endpoint (Site [CNS vs Extracranial vs Both] of Next Progression – RECIST 1.1)

Descriptive statistics presenting the site of next progression will be provided.

5.2.12 Secondary Endpoint – CNS PFS by CNS RECIST 1.1 per ICR

CNS PFS by CNS RECIST 1.1 per ICR is defined as a secondary endpoint for patients in Cohort 2.

5.2.12.1 Definitions and Derivations (CNS PFS by CNS RECIST 1.1 per ICR)

CNS PFS is defined from the first dose of study intervention to disease progression in the brain per CNS RECIST 1.1 or death resulting from any cause, whichever occurs first. For the purpose of this endpoint, progressions not in the brain will not be considered a CNS progression and thus not contribute to CNS progression derivation.

Defining the date of endpoint or censoring will be conducted similarly as described for PFS in Section 5.2.2. Patients who have RECIST progression but no CNS progression will be censored at the time of the RECIST progression assessment. Patients who have not progressed or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST 1.1 assessment.

5.2.12.2 Primary Analysis of Secondary Endpoint (CNS PFS by CNS RECIST 1.1 per ICR)

A KM plot of CNS PFS will be presented. Summaries of the number and percentage of patients experiencing a CNS PFS event, and the type of event (progression or death) will be provided along with the median CNS PFS and respective CIs (calculated using the Brookmeyer and Crowley method) and the proportion of patients that are alive and progression-free at 3-months, 6-months, 9-months, 12-months, etc. (i.e. every 3 months starting from 3-months) calculated using the KM technique.

In addition, the number of patients prematurely censored will be summarized. A patient would be defined as prematurely censored if they have not CNS progression (or died in the absence of progression) and the latest scan prior to DCO was more than 1 scheduled tumor assessment interval plus 2 weeks prior to the DCO date.

The duration of follow-up will be summarized using median time from date of first study intervention to date of censoring (date last known to have not CNS progressed) in censored (not CNS progressed) patients only.

5.2.12.3 Subgroup Analyses (CNS PFS by CNS RECIST 1.1 per ICR)

Subgroup analyses will be performed on the same subgroups using the same methodology as defined in Section 5.2.4.4.

5.2.13 Secondary Endpoint – Time to New CNS Lesions

Time to new CNS lesions is defined as a secondary endpoint for patients in Cohort 2.

5.2.13.1 Definitions and Derivations (Time to New CNS Lesions)

Time to new CNS lesions is defined as the time from the date of the first dose of study intervention to the date of documented new CNS lesions by RECIST 1.1 per ICR, and will only be defined for patients in Cohort 2.

5.2.13.2 Primary Analysis of Secondary Endpoint (Time to New CNS Lesions)

Time to new CNS lesions will be summarized descriptively (i.e., number of patients [%] that had new CNS lesions). KM curves will be produced with median time to new CNS lesions with respective 95% CIs presented as calculated from the KM curve.

5.2.14 Secondary Endpoint – CNS ORR by CNS RECIST 1.1 per ICR

CNS ORR by CNS RECIST 1.1 per ICR is defined as a secondary endpoint for patients in Cohort 2.

5.2.14.1 Definitions and Derivations (CNS ORR by CNS RECIST 1.1 per ICR)

CNS ORR is defined as the proportion of patients with measurable BM at baseline who have a confirmed CR or confirmed PR of CNS lesions, as determined by ICR per CNS RECIST 1.1.

Data obtained from first dose until CNS progression, or the last evaluable assessment in the absence of CNS progression, will be included in the assessment of ORR, regardless of whether the patients withdraws from therapy. Patients who stop study intervention without a response or progression, receive a subsequent therapy, and then respond will not be included as responders in the ORR.

Definitions of confirmed response and situations where PR and CR are from non-consecutive visits will follow the same principles as described for ORR in Section 5.2.1.

5.2.14.2 Primary Analysis of Secondary Endpoint (CNS ORR by CNS RECIST 1.1 per ICR)

CNS ORR will be analyzed using identical methods as outlined for ORR (refer to Section 5.2.1.2).

CNS ORR will only be analyzed for patients in Cohort 2.

5.2.14.3 Subgroup Analyses (CNS ORR by CNS RECIST 1.1 per ICR)

Subgroup analyses will be performed on the same subgroups and following the same methodology defined in Section 5.2.4.4.

5.2.15 Secondary Endpoint – CNS Duration of Response by CNS RECIST 1.1 per ICR

CNS DoR by CNS RECIST 1.1 per ICR is defined as a secondary endpoint for patients in Cohort 2.

5.2.15.1 Definitions and Derivations (CNS Duration of Response by CNS RECIST 1.1 per ICR)

CNS DoR will be defined as the time from the date of first documented confirmed CNS response until date of documented CNS progression per CNS RECIST 1.1 as assessed by ICR or death due to any cause.

Only patients in Cohort 2 who have a confirmed response, regardless of whether the patient withdraws from therapy, receives another anticancer therapy, or clinically progresses prior to CNS RECIST 1.1 progression will have CNS DoR defined.

5.2.15.2 Primary Analysis of Secondary Endpoint (CNS Duration of Response by CNS RECIST 1.1 per ICR)

CNS DoR will be analyzed using identical methods as outlined for DoR (refer to Section 5.2.5.2).

CNS DoR will only be analyzed for patients in Cohort 2.

5.2.15.3 Subgroup Analyses (CNS Duration of Response by CNS RECIST 1.1 per ICR)

Subgroup analyses will be performed on the same subgroups and following the same methodology defined in Section 5.2.4.4.

5.2.16 Secondary Endpoint – Clinical Outcome Assessments

European Organization for Research and Treatment of Cancer (EORTC) 30 item core quality of life questionnaire (QLQ-C30), Neurologic Assessment in Neuro-oncology (NANO) scale and cognitive tests are defined as a secondary endpoint for all patients. MD Anderson Symptom Inventory (MDASI) brain tumor-specific items are defined as secondary endpoint for patients in Cohort 2. The St. George's Respiratory Questionnaire – idiopathic pulmonary fibrosis-specific version (SGRQ-I) is defined as a secondary endpoint for patients that develop ILD/pneumonitis.

PRO questionnaires will only be completed by patients if a linguistically validated version is available in the language of their country of residence. Should such a version not exist, the patient will be exempt from completing the PRO questionnaires.

In case of handheld device failure, the site should follow the ePRO device failure mitigation plan which has been agreed by AstraZeneca. If duplicated questionnaires result from this,

the set of questions fully completed in the handheld device will have priority for the analysis.

5.2.16.1 Definitions and Derivations (Clinical Outcome Assessments)

5.2.16.1.1 EORTC QLQ-C30

The EORTC QLQ-C30 was developed to assess health-related quality of life (HRQoL), functioning and symptoms in cancer clinical trials.

The EORTC QLQ-C30 consists of 30 questions which are grouped to produce 5 multi-item functional scales, 3 multi-item symptom scales, 6 individual items (5 items assessing additional symptoms commonly reported by cancer patients and 1 on the financial impact of the disease) and a 2-item global measure of health status/quality of life (QoL):

- Functional scales:
  - Physical functioning
  - Role functioning
  - Emotional functioning
  - Cognitive functioning
  - Social functioning
- Multi-item symptom scales:
  - Fatigue
  - Nausea and vomiting
  - Pain
- Individual items:
  - Dyspnea
  - Insomnia
  - Appetitive loss
  - Constipation
  - Diarrhea
  - Financial difficulties
- Global health status/QoL

The EORTC QLQ-C30 will be completed by all patients in this study.

An outcome variable consisting of a score from 0 to 100 will be derived for each of the functional scales, symptom scales, individual items and global health status/QoL according to the EORTC QLQ-C30 Scoring Manual (Fayers et al 2001).

Higher scores on the global health status/QoL and functioning scales indicate better health status/functioning. Higher scores on the symptom scales indicate greater symptom burden.

For all scales, if at least half the components of a scale are present for a timepoint then the score will be calculated, otherwise the score will be set to missing.

The principle for scoring the scales of the QLQ-C30 are:

- Estimate the average of the items that contribute to the scale (raw score [RS])
- Use a linear transformation to standardize the RS, so that the scores range from 0 to 100

In practical terms, if terms $I_1, I_2, \ldots, I_n$ are included in a scale then:

$$RS = \frac{\sum_{i=1}^{n} I_i}{n}$$

where RS denotes the raw score.

All functional scales are then calculated as $S = \left\{1 - \frac{RS-1}{\text{range}}\right\} \times 100$ and symptom scales, items and global health status as $S = \left\{\frac{RS-1}{\text{range}}\right\} \times 100$.

Range is defined as the difference between the maximum possible value of RS and minimum possible value of RS.

The global health status/QoL, functional scales and symptom scales, including the items included in each of these scales are presented in Table 9.

Table 9 Scoring the QLQ-C30

	Scale	Number of Items	Item Range [a]	Item Numbers
Global Health Status				
Global Health Status	QL2	2	6	29, 30
Functional Scales				
Physical functioning	PF2	5	3	1 to 5
Role functioning	RF2	2	3	6, 7
Emotional functioning	EF	4	3	21 to 24
Cognitive functioning	CF	2	3	20, 25
Social functioning	SF	2	3	26, 27
Symptom Scales/ Items				
Fatigue	FA	3	3	10, 12, 18
Nausea and vomiting	NV	2	3	14, 15

Table 9 Scoring the QLQ-C30

	Scale	Number of Items	Item Range [a]	Item Numbers
Pain	PA	2	3	9, 19
Dyspnea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite loss	AP	1	3	13
Constipation	CO	1	3	16
Diarrhea	DI	1	3	17
Financial difficulties	FI	1	3	28

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

The EORTC QLQ-C30 summary score will also be calculated from the mean of 13 of the 15 QLQ-C30 scales (the Global Health Status scale and the Financial difficulties are not included). Prior to calculating the mean, the symptom scales need to be reversed to obtain a uniform direction of all scales (i.e. 100 - scale score). The summary score should only be calculated if all of the required 13 scales are available.

For EORTC QLQ-C30, a clinical meaningful change or difference in score will be defined as the corresponding change in points for each scale from Table 10 (Musoro et al, 2019), or if a scale-specific value is not available, a change of at least 10 points (Osoba et al, 1998) will be used. Specifically, once the symptom scores are reversed to follow the function scales interpretation, a clinically meaningful improvement will be defined as a clinically meaningful increase in the score from baseline, whereas a clinical meaningful deterioration will be defined as a clinically meaningful decrease in the score from baseline. At each post-baseline assessment, change in symptoms/functioning from baseline will be categorized as improvement, deterioration, or no change as shown in Table 10.

Table 10 Clinical Meaningful Changes in EORTC QLQ-C30 Scale Scores

Score	Change from Baseline	Visit Response
Physical functioning	≥ + 9	Improvement
	≤ - 10	Deterioration
	Otherwise	No change
Role functioning *	≤ - 6	Deterioration
	Otherwise	No change
Social functioning	≥ + 8	Improvement
	≤ - 7	Deterioration
	Otherwise	No change
Cognitive functioning	≥ + 5	Improvement
	≤ - 4	Deterioration
	Otherwise	No change
Global Health Status	≥ + 12	Improvement
	≤ - 8	Deterioration
	Otherwise	No change
Fatigue	≥ + 8	Improvement
	≤ - 8	Deterioration
	Otherwise	No change
Nausea and vomiting *	≤ - 11	Deterioration
	Otherwise	No change
Appetite loss *	≤ - 14	Deterioration
	Otherwise	No change
Rest of scales	≥ + 10	Improvement
	≤ - 10	Deterioration
	Otherwise	No change

* No minimally important difference (MID) available for improvement.

For each EORTC QLQ-C30 scale, time to deterioration will be defined as the time from first dose of study intervention until the date of first clinically meaningful deterioration (as defined in Table 10) that is confirmed at a subsequent assessment, regardless of whether the patients withdraws from study intervention or receives another anticancer therapy prior to deterioration (i.e., date of deterioration event or censoring – date of first administration of study intervention + 1). Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Patients with no clinical meaningful deterioration will be censored at the time of their last PRO assessment where the score could be evaluated. Also, in the case of deterioration after 2 or more missed PRO assessment visits (i.e., 42 days) the patients will be censored at the time of the last PRO assessment where the score could be evaluated. If the patient has no evaluable visits, does not have baseline data, or has a baseline score that will not allow for a

10-point deterioration (baseline score ≥ 90 for pain scale and ≤ 10 for physical and role functioning and global health status/QoL scales), they will be censored at Day 1 (date of first study intervention administration).

In this analysis, RECIST 1.1 progression will not be considered as deterioration and data will not be affected by RECIST 1.1 progression.

5.2.16.1.2 NANO Scale

The NANO scale is a clinician-reported assessment of neurologic functioning in neuro-oncology patients. The instrument captures 9 domains (gait, strength, ataxia, sensation, visual fields, facial strength, language, level of consciousness, and behavior) and was developed to provide a simple, objective assessment of neurologic function that would be combined with radiographic assessment to provide an overall outcome assessment for neuro-oncology patients in clinical trials and in daily practice.

Each domain is subdivided into 3 or 4 levels of function with scores based on discrete quantifiable measures. Thus, levels of function for each domain range from 0 to 2 or 0 to 3. A score of 0 indicates normal function, while the highest score indicates the most severe level of deficit in that domain.

An overall NANO score per patient and visit will be determined following assessment of each domain and will include 1 of 5 possible outcomes defined in Table 11.

Table 11 NANO Score Outcomes

Outcome	Description
Neurologic response	A ≥ 2 level improvement in at least 1 domain without worsening in other domains from baseline or best level of function
Neurologic progression	A ≥ 2 level worsening from baseline or best level of function within ≥ 1 domain or worsening to the highest score within ≥ 1 domain
Neurologic stability	Indicates a score that does not meet criteria for neurologic response, neurologic progression, non-evaluable or not assessed
Not assessed	Indicates all 9 domains were not evaluated by investigator
Non-evaluable	Indicates investigator opinion that all 9 domains are not evaluable due to other underlying factors, or if measurement of all selected domain is not feasible due to an alteration in another domain

5.2.16.1.3 Cognitive Tests

Computerized, self-completed cognitive tests will be used to capture cognitive function, including attention, memory, and executive function. These gamified tests include Reaction

Time, Paired Associates Learning and Spatial Working Memory, and are low-burden, highly sensitive, precise measures of cognitive function. The following measurements provided by Cambridge Cognition will be included in the analyses:

- Reaction Time Task (RTI):
  - RTI Median Five-Choice Reaction Time (RTIFMDRT): the median duration it took for a patient to release the response button after the presentation of the target stimulus. Calculated across correct, assessed trials in which the stimulus could appear in any one of five locations. Measured in milliseconds.
  - RTI Median Five-Choice Movement Time (RTIFMDMT): the median time taken for a patient to release the response button and select the target stimulus after it flashed yellow on screen. Calculated across correct, assessed trials in which the stimulus could appear in any one of five locations. Measured in milliseconds.

- Paired Associates Learning (PAL):
  - PAL Total Errors Adjusted (PALTEA): the number of times the patient chose the incorrect box for a stimulus on assessment problems, plus an adjustment for the estimated number of errors they would have made on any problems, attempts and recalls they did not reach. This measure allows you to compare performance on errors made across all patients, regardless of those who terminated early versus those completing the final stage of the task.
  - PAL First Attempt Memory Score (PALFAMS): the number of times a patient chose the correct box on their first attempt when recalling the pattern locations, calculated across all assessed trials.

- Spatial Working Memory (SWM):
  - SWM Between Errors (SWMBE): the number of times the patient incorrectly revisits a box in which a token has previously been found. Calculated across all assessed four, six and eight token trials.
  - SWM Total Errors (SWMTE): the number of times a box is selected that is certain not to contain a token and therefore should not have been visited by the patient. Calculated across all assessed four, six and eight token trials.
  - SWM Strategy (SWMS): The number of times a patient begins a new search pattern from the same box they started with previously. If they always begin a search from the same starting point, we infer that the patient is employing a planned strategy for finding the tokens. Therefore, a low score indicates high strategy use (1 = they always begin the search from the same box), a high score indicates they are beginning their searches from many different boxes. Calculated across assessed trials with 6 tokens or more.
  - SWM Between Errors 4 Boxes (SWMBE4), SWM Between Errors 6 Boxes (SWMBE6) and SWM Between Errors 8 Boxes (SWMBE8): The number of times

a patient revisits a box in which a token has previously been found. Calculated across all trials with 4 tokens only for SWMBE4, 6 tokens only for SWMBE6 and 8 tokens only for SWMBE8.

5.2.16.1.4 MDASI

The MDASI symptom diary is a validated, multi-item cancer specific PRO questionnaire capturing symptom severity and interference.

MDASI brain tumor-specific items

The MDASI brain tumor module includes 9 symptoms specific to brain tumors (weakness on 1 side of the body, difficulty understanding, difficulty speaking, seizures, difficulty concentrating, problems with vision, changes in appearance, change in bowel pattern [diarrhea or constipation], and irritability). These 9 items will be used to capture the symptoms associated with BM for those diagnosed with BM. Each item is rated on an 11-point numeric rating scale (0 – 10), with higher scores indicating greater symptom severity.

5.2.16.1.5 SGRQ-I

The original SGRQ is a commonly used, validated instrument capturing HRQoL for patients with chronic respiratory disease. The SGRQ-I is an idiopathic pulmonary fibrosis-specific version of the instrument developed and validated for use among patients with idiopathic pulmonary fibrosis, a type of interstitial lung disease (ILD). The SGRQ-I will be used to assess the HRQoL among patients who have been diagnosed with ILD/pneumonitis. It includes 34 of the original SGRQ items determined to be most reliable for assessing the HRQoL of patients with idiopathic pulmonary fibrosis. The instrument yields 3 domain scores (symptoms, activity, and impact) as well as a total score, with scores ranging from 0 to 100. Higher scores indicate greater impairment in HRQoL.

Scores for each of the 3 domains and total questionnaire are derived by assigning weights to each response within the questionnaire. A score is then derived as:

$$100 \times \frac{\text{sum of weights (in that component or in total)}}{\text{maximum possible sum of weights (in that component or in total)}}.$$

Weights for each item within SGRQ-I are provided in Section 9.1.

The maximum possible sum of weights to be used in denominators are:

- Symptoms: 490.3
- Activity: 744.5
- Impact: 1483.5

- Total: 2718.3

Scores can only be calculated if at least half the components of that domain are present. Components with a missing result will be removed from the numerator and denominator when calculating the score for that domain.

5.2.16.1.6 Compliance

Summary measures of overall compliance and compliance over time will be derived for each PRO, respectively. These will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least 1 individual item completed.
- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time, e.g., a questionnaire from a patient who has not withdrawn from the study at the scheduled assessment time but excluding patients who have been exempted from completing questionnaires, such as those who are blind or illiterate.
- Evaluable questionnaire = a questionnaire with a completion date and at least 1 subscale that is non-missing.

Compliance over time will be calculated separately for each visit, including baseline, as the number of patients with an evaluable questionnaire at the time point (as defined above), divided by the number of patients still expected to complete questionnaires. Similarly, the evaluability rate over time will be calculated separately for each visit, including baseline, as the number of evaluable questionnaires (per definition above), divided by the number of received questionnaires.

5.2.16.2 Primary Analysis of Secondary Endpoint (Clinical Outcome Assessments)

Descriptive statistics including change from baseline scores/results (where applicable) will be calculated and presented for each scheduled visit/timepoint in the study for EORTC QLQ-C30 scores, individual cognitive test results, MDASI-BM item scores (patients in Cohort 2 only) and SGRQ-I (patients with ILD/pneumonitis only i.e. using the SAF-ILD) scores. The baseline assessment for SGRQ-I will be the scores obtained after diagnosis of ILD/pneumonitis. Plots of mean change from baseline over time will also be presented with indicators for the number of patients in each visit/timepoint. Standard error bars will be presented on the plot.

A shift table from first result obtained after start of study intervention to the worst on-study intervention result for the NANO Scale will be presented.

Categorical results presenting whether there was an improvement, no change or worsening in the particular PRO will also be presented descriptively for all EORTC QLQ-C30.

For EORTC QLQ-C30 time to deterioration will be presented using KM plots. Summaries of the number and percentage of patients experiencing a clinical meaningful deterioration and the median time to deterioration will be provided.

Summaries for the best overall on-treatment visit response and response at each visit for EORTC QLQ-C30 will be provided. Best overall on-treatment visit will use all assessments between the start of study treatment and 47 days following the date of last dose of study treatment, and prior to starting any subsequent anticancer therapy.

Compliance with the completion of EORTC QLQ-C30, NANO scale, cognitive tests and MDASI-BM specific items will also be presented descriptively.

5.2.16.3 Subgroup Analyses (Clinical Outcomes Assessments)

CCI [redacted]

[redacted]

5.2.17 CCI [redacted]

CCI [redacted]

5.2.17.1 CCI [redacted]

CCI [redacted]

CCI [redacted]

5.2.17.2 CCI [redacted]

CCI [redacted]

5.2.18

CCI

CCI

5.2.18.1

CCI

CCI

CCI

CCI

CCI

5.2.18.2

CCI

CCI

5.2.19

CCI

CCI

5.2.19.1

CCI

CCI

5.2.19.2 CCI
CCI

5.2.20 CCI
CCI

5.2.20.1 CCI
CCI

5.2.20.2 CCI
CCI

5.2.21 CCI
CCI

5.2.21.1 CCI
CCI

5.2.21.2 CCI
CCI

5.2.22 CCI
CCI

5.2.22.1 CCI
CCI

5.2.22.2 CCI
CCI

CCI

5.2.23 CCI

CCI

5.2.23.1 CCI

CCI

5.2.23.2 CCI

CCI

5.2.24 CCI

CCI

5.2.24.1 CCI

5.2.24.1.1 CCI

CCI

CCI

CCI

5.2.24.1.2 CCI

CCI

CCI

CCI

5.2.24.1.3 CCI

CCI

5.2.24.2 CCI

CCI

CCI

5.2.25 CCI

CCI

5.2.25.1 CCI

CCI

- CCI
- CCI
- CCI
- CCI
- CCI

CCI

CCI [redacted]

5.2.25.2 CCI [redacted]

CCI [redacted]

5.2.26 Supportive Endpoint – Tumor Size

While tumor size is not a formal endpoint, tumor size will be analyzed as supportive to RECIST 1.1 specific analyses.

5.2.26.1 Definitions and Derivations (Tumor Size)

Tumor size is defined as the sum of LDs of the target lesions (or short axis measurements for lymph nodes).

The best percentage change in tumor size is the largest decrease from baseline or the smallest increase from baseline in the absence of a reduction and will include all assessments prior to the earliest of death in the absence of progression, any evidence of progression, the start of subsequent anticancer therapy, or the latest evaluable RECIST 1.1 assessment if the patient has not died, progressed or started subsequent anticancer therapy. Change in tumor size at progression or the latest evaluable RECIST 1.1 assessment (as applicable) should be included in the determination of the best percentage change in tumor size.

If the best percentage change cannot be calculated due to missing data (including if the patient has no TLs at baseline), a value of +20% will be imputed as best percentage change from baseline in the following situations (otherwise best percentage change will be left as missing):

- If a patient has no post-baseline assessment and has died
- If a patient has a new lesion or progression of NTLs or TLs

- If a patient has withdrawn due to PD and has no evaluable TL data before or at PD

5.2.26.2 Primary Analysis of Supportive Endpoint (Tumor Size)

The absolute values, change in TL tumor size from baseline and percentage change in TL tumor size from baseline will be summarized using descriptive statistics and presented at each timepoint.

The best percentage change in TL tumor size from baseline will also be summarized. The number and percentage of patients whose best percentage change was imputed will also be presented. Imputed data will be included in descriptive summaries.

Best percentage change in tumor size will also be presented graphically using waterfall plots. Each patient will be presented as a separate bar with the bars ordered from the largest increase to the largest decrease. A reference line at the -30% change in TL tumor size level will be added to the plots, which correspond with the definition of a 'partial response' (RECIST 1.1). All progressions will be marked with a '●' or designated with patterns or colors for ORR categories. Flagged progressions on the percentage change in TL tumor size at a particular timepoint will be based upon the NTL or new lesion progression at that timepoint. The scale of these plots will be fixed to be from -100% to 100% to avoid presenting extreme values. Values that are capped as a result of this restriction to the scale will be marked with '#'. Values are ordered in descending order with imputations due to death appearing first followed by a gap followed by all other patients. Imputed values will be clearly marked with '*' and patients with imputation where there was a death or evidence of progression will have different shading to each other and the other patients to make it clear that these are different.

5.2.27 Additional Analysis – Disagreements between ICR and Site Investigator Reviews

The comparison between ICR and site investigator reviews will be analyzed as supportive to RECIST 1.1 specific analyses. It will be analyzed for the FAS population, by cohort.

5.2.27.1 Definitions and Derivations (Disagreements between ICR and Site Investigator Reviews)

Imaging data is reviewed and assessed using RECIST 1.1 by the site investigator on the first instance and then independently by the ICR. The data from the ICR is used mainly on the different study endpoints, although sensitivity analyses are performed on the site

investigator data. A comparison of the RECIST progression indicated in both sources will be presented.

5.2.27.2 Primary Additional Analysis (Disagreements between ICR and Site Investigator Reviews)

The agreement between ICR and site investigator reviews will be assessed by presenting the proportion of patients who had RECIST progression according to both assessments, neither assessment and either assessment. The proportion of patients who had RECIST progression according to both assessments, will be further subdivided into progression identified by investigator assessment >2 weeks after ICR, >2 weeks before ICR and within 2 weeks of ICR.

5.3 Safety Analyses

The domain safety covers exposure, AEs, deaths, clinical laboratory, vital signs, electrocardiograms (ECGs), echocardiograms, pulmonary assessments, ILD/pneumonitis investigation, ECOG performance status and ophthalmologic assessments.

Tables and listings are provided for the SAF depending on the availability of data.

5.3.1 Exposure

5.3.1.1 Definitions and Derivations (Exposure)

Total (or intended) Exposure

The total (or intended) exposure (i.e., duration of study intervention) of a patient to an intervention is calculated using the start and stop date of the intervention and the intended dosing interval. For a dosing period of the intervention, the total (or intended) exposure is calculated as the number of days from date A to date B (i.e., B – A + 1) where:

- A is the date of first dose of study intervention in the dosing period
- B is the earliest of:
  - The date of death
  - The date of DCO
  - The date when the last non-zero dose of study intervention was received (e.g., > 0 mg T-DXd) plus 20 (where 20 is the dosing interval in days)

Actual Exposure

Actual exposure is defined as the actual study intervention duration in days.

Actual exposure = intended exposure (days) – total duration of dose interruptions (days), where intended exposure will be calculated as described above and a dose interruption is defined as any length of time where the patient has not taken any of the planned dose.

Since patients will receive T-DXd via IV infusion q3w (± 2 days), the total duration of dose interruptions (for deriving actual exposure) will be calculated as follows:

Total duration of dose interruptions = sum for all positive values of [date of the dose – date of the previous dose – (21 + 2) days].

Dose modifications are permitted, and the calculation of actual study intervention duration makes no adjustment for any dose modifications that may have occurred.

Number of Study Intervention Cycles Received

Exposure will also be measured by the number of cycles received. A cycle corresponds to a period of 21 days. If a cycle is prolonged due to toxicity, this should still be counted as 1 cycle. A cycle will be counted if study intervention is started even if the full dose is not administered.

Safety Follow-up

The total safety follow-up of a patient is calculated using the start date of the intervention and the intended safety follow-up interval. It is calculated as the number of days from date A to date B (i.e., B – A + 1) where:

- A is the date of first dose of study intervention in the dosing period
- B is the earliest of:
  - The date when the last non-zero dose of study intervention was received (e.g., > 0 mg T-DXd) plus 47 (where 47 is the safety follow-up period in days)
  - The date of withdrawal of main informed consent
  - The date of death
  - The date of DCO

Dose Intensity

Dose intensity (DI) is defined as the actual cumulative dose delivered up to the actual last day of dosing relative to the exposure period and derived as:

DI = d/total exposure/21

where d is the actual cumulative dose delivered up to the actual last day of dosing.

When deriving actual dose administered the volume before and after infusion will be considered.

Planned Dose Intensity

Planned dose intensity (PDI) is the planned cumulative dose that should've been delivered up to the actual last day of dosing relative to the exposure period derived as:

PDI = D/total exposure/21

where 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 modifications to the dose or schedule.

When accounting for the calculation of intended cumulative dose, 2 days should be added to the date of last dose of T-DXd to reflect the protocol window for dosing.

Relative Dose Intensity

Relative dose intensity (RDI) is the percentage of the actual dose delivered relative to the intended dose through to study intervention discontinuation. RDI will be defined as follows:

RDI = 100% * DI/PDI.

5.3.1.2 Presentation (Exposure)

Exposure will be analyzed for the safety analysis set. The following summaries will be produced:

- Summary of total exposure
- Summary of actual exposure
- Summary of number of cycles received
- Summary of dose interruptions, reductions and cycle delays
- Summary of RDI

The exposure of each individual patient will be listed as well.

5.3.2 Adverse Events

5.3.2.1 Definitions and Derivations (Adverse Events)

AEs and serious AEs (SAEs) will be collected throughout the study, from date of informed consent until 47 days after the last dose of study intervention.

Events will be defined as treatment-emergent AEs (TEAEs) if they onset or worsen (by investigator report of a change in intensity), after initiating study intervention until 47 days after last dose of study intervention. SAEs with an onset or worsening 48 days or more after the last dose of study treatment, if considered related to the study treatment, are also TEAEs.

MedDRA (using the latest or current version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute CTCAE version 5.0.

Missing start and stop dates for AEs and missing causality will be handled using the rules described in Section 4.3.5.

AEs of special interest (AESI)

The following events are considered to be AESIs:

- ILD/Pneumonitis
- Left Ventricular Ejection Fraction (LVEF) decrease

Other categories may be added 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 database lock to ensure any further terms not already included are captured within the categories.

Time to onset for AESIs will be derived as [AESI start date – study intervention start date] + 1 while duration will be derived as [AE end date – AE start date] + 1. If the AESI is ongoing at the DCO, the DCO date will be imputed as the end date for duration calculations. Time to onset and duration will be presented in days.

Other significant adverse events (OAE)

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs and 'Discontinuation of Investigational Product due to Adverse Events'. Based on the expert's judgement, significant adverse events that are considered to be of clinical relevance or importance may, after consultation with the Global Patient Safety Physician, be considered OAEs and reported as such in the CSR. A similar review of laboratory/vital signs/ECG data (i.e. clinical relevance) will be performed for identification of OAEs.

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

5.3.2.2 Presentation (Adverse Events)

In general, all AE summary tables will include only TEAEs prior to the start of new anticancer therapies. AEs occurring prior to dosing or starting more than 47 days after discontinuation of study intervention or that are a TEAE but started after the start of a new anticancer therapy, unless they are SAEs with an onset or worsening 48 days or more after

the last dose of study treatment and considered related to study treatment, will be flagged in listings and will not be included in any summaries.

All reported AEs will be listed along with date of onset, date of resolution (if AE is resolved), investigator's assessment of CTCAE grade, relationship to study intervention, action taken and outcome. Frequencies and percentages of patients reporting each preferred term will be presented (i.e., multiple events per patient will not be accounted for, except in event level summaries).Summary information (the number and percentage of patients) by MedDRA SOC and PT will be tabulated for:

- All TEAEs
- All TEAEs, excluding TEAEs associated with COVID-19 infection
- All TEAEs associated with COVID-19 infection
- All TEAEs occurring in at least 5% of patients
- All non-serious TEAEs occurring in at least 5% of patients.
- All TEAEs possibly related to study intervention
- TEAEs by maximum CTCAE grade
- TEAEs of CTCAE grade 3 or higher
- TEAEs of CTCAE grade 3 or higher, possibly related to study intervention
- TEAEs of CTCAE grade 3 or higher occurring in at least 5% of patients
- TEAEs with outcome of death
- TEAEs with outcome of death, excluding TEAEs associated with COVID-19 infection
- TEAEs associated with COVID-19 infection with outcome of death
- TEAEs with outcome of death, possibly related to study intervention
- TEAEs leading to dose reduction and interruption (separately)
- All treatment-emergent SAEs
- All treatment-emergent SAEs possibly related to study intervention
- TEAEs leading to discontinuation of study intervention
- TEAEs leading to discontinuation of study intervention, excluding TEAEs associated with COVID-19 infection
- TEAEs associated with COVID-19 infection leading to discontinuation of study intervention
- TEAEs leading to discontinuation of study intervention, possibly related to study intervention
- Treatment-emergent SAEs leading to discontinuation of study intervention
- Treatment-emergent SAEs leading to discontinuation of study intervention, possibly related to study intervention
- Treatment-emergent OAEs
- Treatment-emergent AESIs

- TEAEs among patients with baseline BM who are treated with concurrent high-dose steroid

An overall summary of the number and percentage of patients in each of the above categories (where applicable) will be presented, as well as an overall summary of the number of events in each of the above categories.

In addition, an event level summary will be presented for all TEAEs by PT.

All AESI PTs, if applicable, searched for in this study will be presented. In addition, summaries of treatment-emergent AESIs will be presented by maximum reported CTCAE grade, by AE outcome, including time of resolution (on-treatment or survival follow-up [as defined in Section 4.3.1]), whether treatment was received (yes/no) and action taken.

Descriptive statistics for time to onset and duration of first AESI for patients experiencing AESIs will be presented.

In addition, TEAEs with outcome death, treatment-emergent SAEs, TEAEs leading to discontinuation of study intervention, treatment-emergent OAEs, TEAEs with CTCAE grade 3 or higher, TEAEs leading to dose reduction or dose interruption and AESIs will be listed separately.

5.3.2.3 Subgroup Analyses (Adverse Events)

Subgroup analyses will be presented as followed:

CCI

CCI

CCI

5.3.3 Deaths

5.3.3.1 Presentation (Deaths)

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

- Related to disease under investigation only
- AE outcome = death only
- Both related to disease under investigation and with AE outcome = death
- AE with outcome = death > 47 days after last study intervention
- Other deaths

A corresponding listing will also be produced.

5.3.4 Laboratory Safety Variables

5.3.4.1 Definitions and Derivations (Laboratory Safety Variables)

Safety laboratory results will be obtained from local laboratories.

Laboratory variables that will be measured are detailed in Table 12.

Table 12 Laboratory Safety Variables

Hematology/hemostasis (whole blood)	**Clinical chemistry (serum or plasma)**
Hemoglobin	Creatinine
Leukocyte count	Bilirubin, total
Leukocyte differential count (absolute count; neutrophils, lymphocytes, monocytes, eosinophils, basophils)	Alkaline phosphatase
Platelet count	AST
Total white blood cell count	ALT
Total red blood cell count	Albumin
Hematocrit	Potassium
	Calcium, total
Urinalysis	Sodium
Hemoglobin/erythrocytes/blood	Gamma-glutamyl transferase
Protein/albumin	Lactate dehydrogenase
Glucose	Protein, total
	Urea nitrogen/blood urea nitrogen
Coagulation	Troponin
Coagulation variables (aPTT, PTT, and INR)	Magnesium
	Chloride
Other	Serum creatinine

Table 12 Laboratory Safety Variables

Hematology/hemostasis (whole blood)	Clinical chemistry (serum or plasma)
Pregnancy (serum [at screening] or urine [other timepoints])	
Hepatitis B and C serology (optional)	
HIV antibodies (as required)	

ALT = alanine aminotransferase; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; HIV = human immunodeficiency virus; INR = international normalized ratio; PTT = partial thromboplastin time.

Corrected calcium will be calculated as:

Corrected calcium (mmol/L) = 0.02 x (normal albumin [g/dL] – patient's albumin [g/dL]) + total serum calcium.

All values will be classified as low (below range), normal (within range), or high (above range) based on local laboratory reference ranges. Results will be converted to standard units and graded with CTCAE version 5.0

If the same parameter is found as measured in serum and in plasma, then the summaries will not distinguish between them (e.g., values from plasma Albumin and serum Albumin will be summarized under Albumin). If the same parameter is found as measured in serum and in plasma within the same patient, which would be a rare case, then the change from baseline will only be calculated for those post-baseline values using the same source, i.e., only within plasma or serum. If 1 patient has multiple toxicity grades, because they are derived separately from serum and plasma then the maximum value of the 2 will be considered.

5.3.4.2 Presentations (Laboratory Safety Variables)

For all continuous laboratory assessments, absolute value, change from baseline and percentage change from baseline will be summarized using descriptive statistics at each schedule assessment time.

For clinical chemistry and hematology, CTCAE grade changes from baseline to the maximum grade on-study intervention will be provided. Corresponding shift tables ("Negative", "Trace", "Positive", "0", "+", "++", "+++") will be produced for urinalysis. In addition, the number of patients with ≥ 2 CTCAE grade changes and CTCAE grade changes to 3 or 4 will be summarized for clinical chemistry and hematology parameters.

Liver biochemistry test results over time for patients who show elevated ALT or AST (≥ 3 x ULN) and elevated bilirubin (≥ 2 x ULN) (elevated results do not need to be present at the same visit) or ALT or AST ≥ 5 x ULN, will be tabulated and plotted.

Individual laboratory measurements will be listed.

5.3.5 Vital Signs

5.3.5.1 Definitions and Derivations (Vital Signs)

Temperature, pulse rate, respiratory rate and blood pressure will be assessed.

Normal ranges are presented in Table 13.

Table 13 Vital Sign Normal Ranges

Vital Sign	Outside AZ defined reference range lower limit if	Outside AZ defined reference range upper limit if	Treatment emergent decrease if change from baseline	Treatment emergent increase if change from baseline
Systolic blood pressure (mmHg)	< 100	> 160	< -30	> 30
Diastolic blood pressure (mmHg)	< 60	> 100	< -15	> 15
Pulse (bpm)	< 40	> 100	< -20	> 20

5.3.5.2 Presentations (Vital Signs)

Absolute values, change from baseline and percentage change from baseline for vital sign parameters will be summarized by visit.

Shift tables baseline to maximum and minimum value on-treatment, prior to starting any subsequent anticancer therapy, for blood pressure and pulse will also be presented.

5.3.6 Electrocardiogram

5.3.6.1 Definitions and Derivations (Electrocardiogram)

ECGs at screening will be performed in triplicate. Subsequent ECGs will only be performed in triplicate if abnormalities are noted. Where triplicate ECGs are present the mean of the triplicates will be calculated. ECG variables that will be assessed include heart rate, RR, PR, QRs, QT and QTcF intervals.

The following relationship between RR and heart rate (with RR expressed in seconds and heart rate in bpm) will be used to derive programmatically the missing parameter in case only 1 of these variables is available:

$$\mathrm{RR} = \frac{60}{\text{heart rate}}.$$

5.3.6.2 Presentations (Electrocardiogram)

Absolute values, change from baseline and percentage change from baseline for ECG parameters will be summarized by visit. Where readings are taken in triplicate, the mean of the triplicate results will be used for analysis purposes, while the results from individual tracings will be presented in listings.

Overall evaluation of ECG will be summarized descriptively.

Shift tables from baseline to maximum and minimum values on-treatment, prior to starting any subsequent anticancer therapy, will be provided.

QTcF outliers (defined as values following study intervention that are greater than 450 msec or increases from baseline greater than 30 msec) will be summarized using cumulative counts and percentages under the following categories:

- Absolute value > 450 msec
- Absolute value > 480 msec
- Absolute value > 500 msec
- Change from baseline > 30 msec
- Change from baseline > 60 msec
- Absolute value > 450 msec and change from baseline > 30 msec
- Absolute value > 500 msec and change from baseline > 60 msec

5.3.7 Echocardiogram/Multigated Acquisition Scan

5.3.7.1 Definitions and Derivations (Echocardiogram/Multigated Acquisition Scan)

Echocardiogram or multigated acquisition scans will be performed to assess LVEF.

5.3.7.2 Presentations (Echocardiogram/Multigated Acquisition Scan)

Absolute value, change from baseline and percentage change from baseline will be presented for LVEF.

LVEF outliers will be summarized descriptively where outliers are defined as:

- ≥ 10% decrease from baseline and result < 50%; or
- ≥ 15% decrease from baseline and result ≥ 50%.

Box plots of absolute value and change from baseline over time will also be presented.

5.3.8 SpO2 and Pulmonary Assessments

5.3.8.1 Definitions and Derivations (SpO2 and Pulmonary Assessments)

Pulse oximetry (SpO2) and basic spirometry components including forced vital capacity (FVC), FVC% predicted, forced expiratory volume in 1 second (FEV1), FEV1% predicted, FEV1/FVC% and diffusion capacity of the lungs for carbon monoxide (DLCO) for pulmonary function tests will be presented. Optional additional spirometry components may be measured.

CCI [redacted]

5.3.8.2 Presentations (SpO2 and Pulmonary Assessments)

Absolute values will be listed for SpO2, FVC, FVC% predicted, FEV1, FEV1% predicted, FEV1/FVC% and DLCO.

5.3.9 ILD/Pneumonitis Investigation

5.3.9.1 Definitions and Derivations (ILD/Pneumonitis Investigation)

If new or worsening pulmonary symptoms (e.g., dyspnea, cough or fever) or radiological abnormality suggestive of ILD/pneumonitis is observed, study intervention should be interrupted, and a full investigation will be conducted based on investigator's judgement. Evaluations will be reported in the CRF.

5.3.9.2 Presentations (ILD/Pneumonitis Investigation)

Data collected for ILD/pneumonitis investigation will be summarized and listed using the SAF-ILD population. An additional listing will be produced for patients who died with ILD/pneumonitis as cause of death, as assessed by investigator.

Additionally, ILD clinical symptoms resolution among the ILD patients will be summarized for patients who have been treated with high-dose steroids. The table will present the total number of resolved/unresolved cases, and number of resolved/unresolved cases by high-dose steroid treatment duration.

5.3.10 ECOG Performance Status

5.3.10.1 Definitions and Derivations (ECOG Performance Status)

ECOG performance status results will be collected.

5.3.10.2 Presentations (ECOG Performance Status)

Shifts in ECOG performance status will be presented descriptively.

5.3.10.3 Subgroup Analyses (ECOG Performance Status)

CCI [redacted]

5.4 Other Analyses

5.4.1 Impact of COVID-19

Additional analyses will be performed to explore the impact of COVID-19 and implemented contingency measures (e.g., patients discontinued from study intervention and/or study, alternative procedures used to collect critical safety and/or efficacy data, protocol deviations related to COVID-19) on the safety and efficacy results reported for the study. These analyses have been described separately within each section of this SAP.

6 INTERIM ANALYSIS

6.1 Independent Data Monitoring Committee (IDMC)

An Independent Data Monitoring Committee (IDMC) comprised of independent experts will be convened and will meet approximately 6 months after the study has started or after the first 60 patients have been treated, whichever occurs first. The IDMC will review safety data and selected efficacy data, and make recommendations to continue, amend or stop the study based on safety findings. The committee will meet approximately every 6 months thereafter and at each meeting make recommendations to continue, amend or stop the study based on safety findings. ORR summaries using RECIST 1.1 derivations will use the FAS-ORR instead of FAS to ensure all patients in the analysis set have the opportunity to have an unconfirmed response. Listings using RECIT 1.1 derivations will continue to use the FAS.

Full details on the IDMC procedures and processes can be found in the IDMC Charter.

6.2 Interim Analysis

An interim analysis of the primary endpoint of PFS and other pertinent endpoints in Cohort 2 is planned for when approximately 125 (50%) eligible patients are enrolled in Cohort 2.

The interim analysis will only present data on Cohort 2 patients and analyses will follow the same principles as detailed in this SAP. Outputs presented for the interim analysis will concentrate on PFS based on RECIST 1.1, disposition, demographics, baseline and disease characteristics and key safety.

ORR summaries using RECIST 1.1 derivations will use the FAS-ORRc and FAS-ORRi instead of FAS to ensure all patients in the analysis set have the opportunity to have an unconfirmed response. Listings using RECIST 1.1 derivations will continue to use the FAS.

7 COUNTRY SPECIFIC ANALYSES

No country specific analyses are planned for this study.

8 REFERENCES

Fayers et al, 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). European Organization for Research and Treatment of Cancer, Brussels 2001.

Modi et al, 2020
Modi S, Saura C, Yamashita T, et al. Trastuzumab Deruxtecan in Previously Treated HER2-Positive Breast Cancer. N Engl J Med. 2020;382(7):610-21.

Murthy et al, 2020
Murthy RK, Loi S, Okines A, et al. Tucatinib, Trastuzumab, and Capecitabine for HER2-Positive Metastatic Breast Cancer [published correction appears in N Engl J Med. 2020;382(6):586]. N Engl J Med. 2020;382(7):597-609.

Musoro et al, 2019
Musoro JZ, Coens C, Fiteni F, et al. Minimally Important Differences for Interpreting EORTC QLQ-C30 Scores in Patients With Advanced Breast Cancer. JNCI Cancer Spectrum, Volume 3, Issue 3, September 2019.

Osoba et al, 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality of life scores. J Clin. Oncol. 1998; 16:139-144.

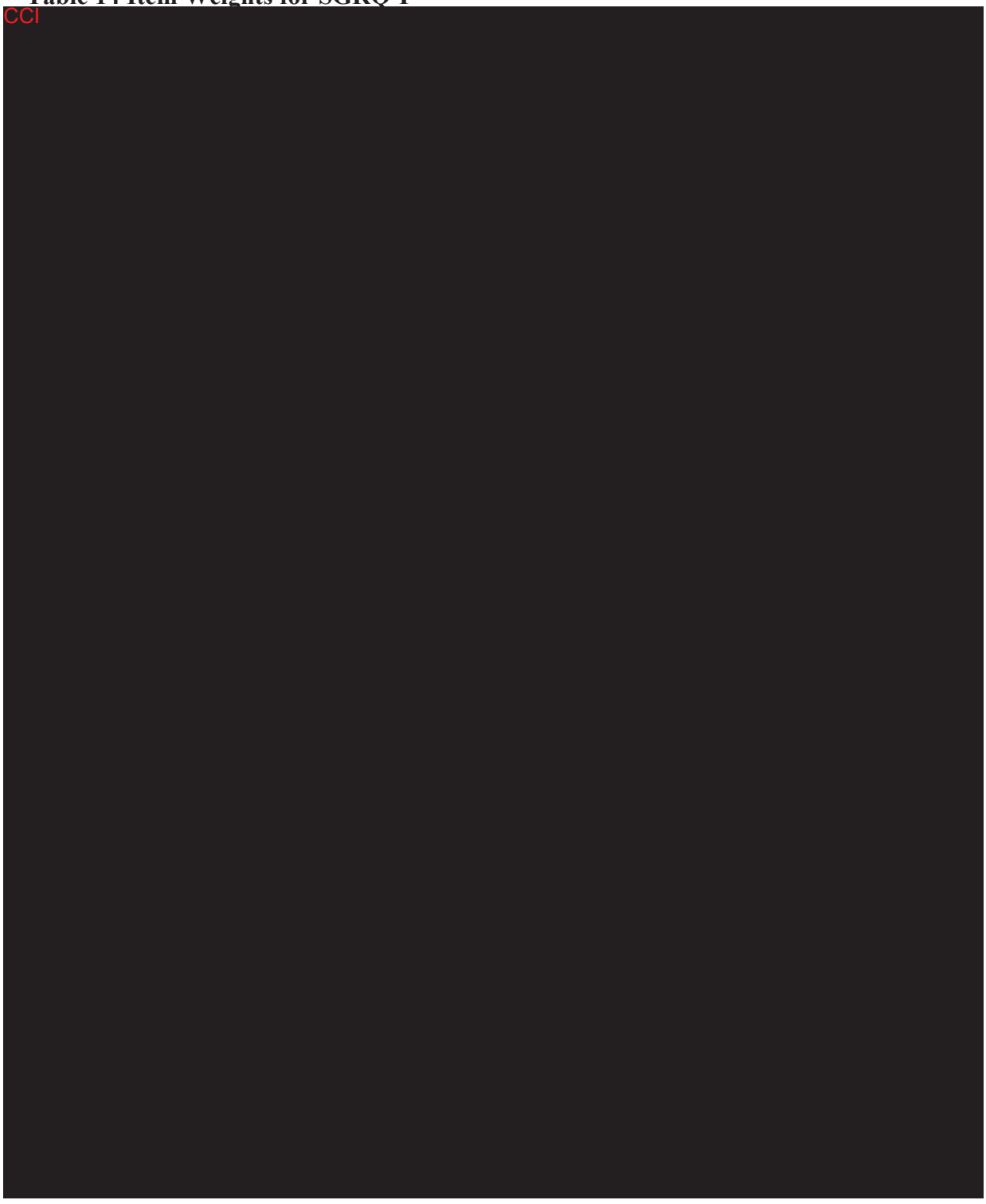
Wolff et al, 2018
Wolff AC, Hammond MEH, Allison KH, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. J Clin Oncol. 2018;36(20):2105-22.

9 APPENDIX

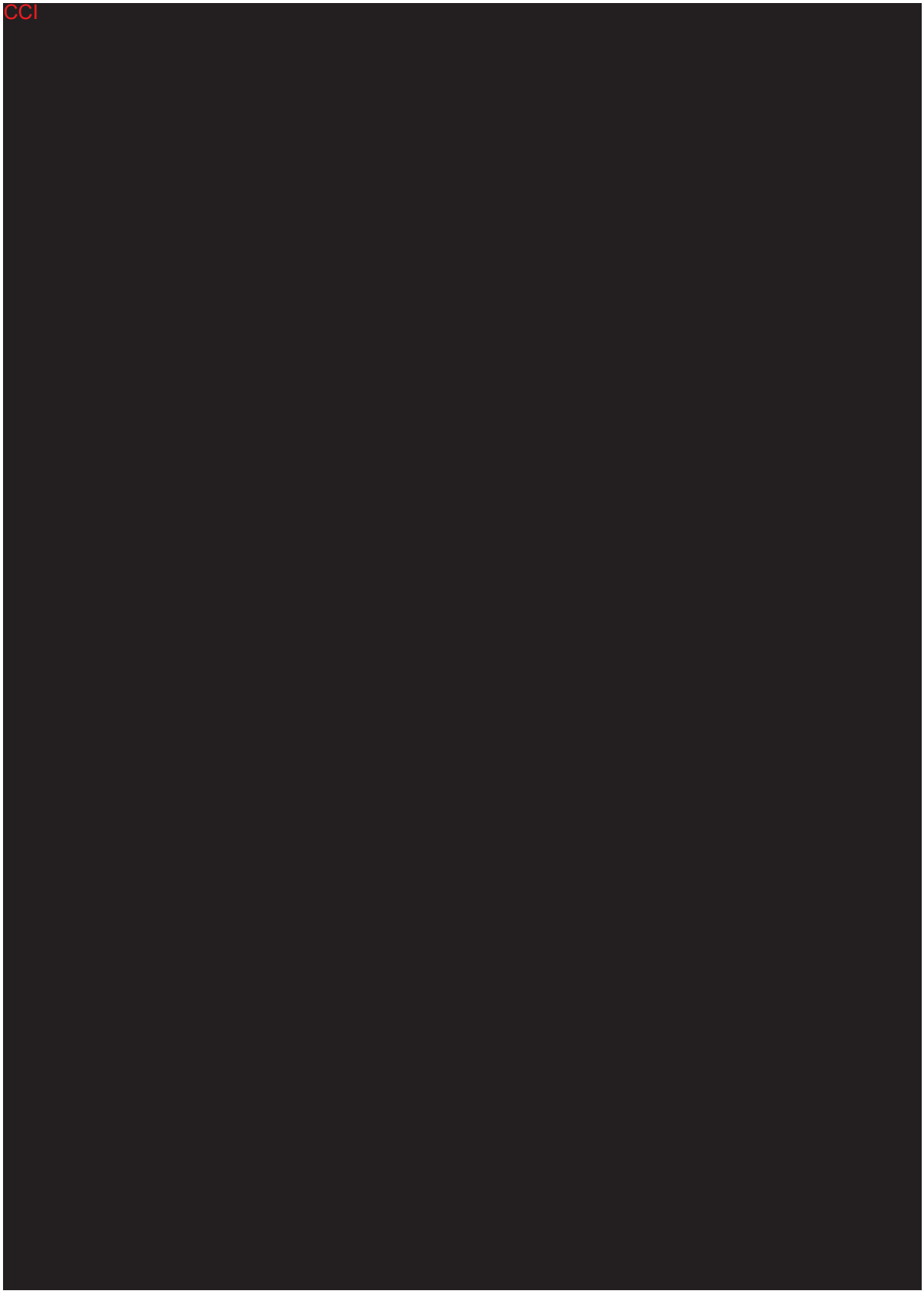
9.1 Item Weights for SGRQ-I

Table 14 Item Weights for SGRQ-I

CCI



CCI



CCI

SIGNATURE PAGE

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature

Document Name: d9673c00007-sap-ed-4		
Document Title:	Statistical Analysis Plan Edition 4	
Document ID:	Doc ID-004586422	
Version Label:	3.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
11-Dec-2023 13:05 UTC	PPD	Author Approval
11-Dec-2023 16:56 UTC	PPD	Qualified Person Approval

Notes: (1) Document details as stored in ANGEL, an AstraZeneca document management system.