AstraZeneca

Statistical Analysis Plan

Study Code	D419EC00001
Edition Number	4.0
Date	19Apr2023

Phase I/II, Open-Label, Multicenter Study to Evaluate the Safety, Tolerability and Preliminary Efficacy of Durvalumab Monotherapy or Durvalumab in Combination with Tremelimumab in Pediatric Patients with Advanced Solid Tumors and Hematological Malignancies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
ADA	Anti-drug antibody
AE	Adverse event
AEPI	Adverse event of possible interest
AESI	Adverse event of special interest
ALT	Alanine transaminase
APF	Alive and progression-free
APF12/18	Proportion of patients alive and progression-free at 12 / 18 months from first dose of study treatment
AST	Aspartate transaminase
AUC	Area under the curve
BMI	Body mass index
BoR	Best objective response
CAR-T	Chimeric antigen receptor T-cell therapy
CI	Confidence interval
C_{max}	Maximum serum concentration
$C_{max\ ss}$	Maximum plasma concentration at steady state
C_{min}	Minimum serum concentration
$C_{min\ ss}$	Minimum plasma concentration at steady state
COVID-19	Coronavirus Disease 2019
CPA	Clinical Pharmacology Alliance
CR	Complete response
CSP	Clinical study protocol
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Event
CTLA-4	Cytotoxic T-lymphocyte-associated antigen 4
CV	Coefficient of variation
DCR	Disease control rate
DCO	Data cut off
DLT	Dose-limiting toxicity
DoR	Duration of response

Abbreviation or special term	Explanation
DRC	Data Review Committee
ECG	Electrocardiogram
eCRF	electronic case report form
FAS	Full analysis set
GeoSD	Geometric standard deviation
Gmean	Geometric mean
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HLT	Higher-level terms
ICs	Immune cells
IHC	Immunohistochemistry
imAE	Immune-mediated adverse event
INRC	International Neuroblastoma Response Criteria
irAE	Immune-related adverse event
irRECIST	Immune-related Response Evaluation Criteria in Solid Tumors
KM	Kaplan-Meier
LD	Longest diameter
LLOQ	Lower limit of quantitation
MedDRA	Medical Dictionary for Regulatory Activities
MR	Minor response
MRI	Magnetic resonance imaging
MTD	Maximum tolerated dose
NA	Not applicable
NB	Neuroblastoma
NC	Not calculable
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
NK	Natural killer
NQ	Non-quantifiable
NTL	Non-target lesion
ORR	Objective response rate

Abbreviation or special term	Explanation
OS	Overall survival
OS12/24	Proportion of patients alive at 12 / 24 months from first dose of study treatment
PCR	Polymerase chain reaction
PD	Progression of disease
PD-L1	Programmed cell death ligand 1
PD-L2	Programmed cell death ligand 2
PFS	Progression-free survival
PK	Pharmacokinetics
PR	Partial response
PT	Preferred term
q4w	Every 4 weeks
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
RNA	Ribonucleic acid
RP2D	Recommended Phase II dose
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SARC-1	Osteosarcoma and Ewing sarcoma
SARC-2	Rhabdomyosarcoma, non-rhabdomyosarcoma, and other sarcomas
SARCOMA	Osteosarcoma and Ewing sarcoma; soft-tissue sarcomas: Rhabdomyosarcoma, non-rhabdomyosarcoma, and other sarcomas
SAS	Safety analysis set
SD	Stable disease
SoA	Schedule of activities
SOC	System organ class
STD	Standard deviation
STO	Other solid tumors
TL	Target lesion
TSH	Thyroid stimulating hormone
ULN	Upper limit of normal

AMENDMENT HISTORY

Date	Brief description of change
08JUN2021	There have been 3 protocol versions since the original SAP was created, updates as follows: Section 1, updated to align objectives with protocol version 5. Study design and patient numbers updated to reflect cohorts removed and revised patient numbers in protocol version 5. Section 3.1 updated to reflect new tumour assessment schedule and to remove references to cohorts removed in protocol version 5. Section 3.3 removed references to ECG interval data as this is not being collected. Section 3.5.1 and 3.5.2, references to the HL cohort were removed. Section 4.2 layout of outputs for dose-finding part amended to include separate summaries by weight group. Section 4.2.3.2 clarified that dose modification in AE outputs means interruption or discontinuation, as increases and decreases are not permitted. Section 4.2.2, removed references to cohorts removed in protocol version 5. Clarified efficacy outputs are for dose-expansion only and accordingly removed references to INRC which is for the NB cohort enrolled in dose-finding only. Section 4.2.5, COVID-19 pandemic related outputs have been added. Section 4.2.3.5 removed reference to listing of physical examination data, as any physical examination findings are reported as adverse events. Section 4.2.3.6 borderline ECG results added to reporting as a separate category. CCI [redacted] Section 5, patient numbers for the DRC updated in line with protocol version 5.
24JUN2021	Section 2.1, amended the Evaluable for response analysis set and added the sensitivity analysis for unconfirmed ORR. Section 3, removed INRC references throughout. Section 3.2.2, amended the missed visits to account for the change to 16 weekly efficacy assessment being optional and clarified a RECIST visit response of NE does not count as a missed visit. Section 3.3.4, changed BMI to BSA. Section 4.2.2.1 and 4.2.2.6, clarified the mid-P approach is only required if the cohorts are expanded. Section 4.2.3.4, removed the vital signs shift table. Section 6, updated to include any relevant change detailed above.

Date	Brief description of change
17Nov2022	Removed signature pages. Section 2.1, more clarification provided for Evaluable for response analysis set, and only derive for Dose expansion phase. Section 3.3.1 and 4.2.3.2, added AEPI. Section 3.3.2 and 4.2.3.3, updated list of parameters with CTCAE grades and added detail for parameters with both low and high direction CTCAE grading. Section 3.4.1 and 4.2.4.1. removed PK steady state. Section 4.1.2, clarified derivation of baseline for questionnaires. Section 4.2.7.2, removed efficacy repeat outputs[redacted]. Section 4.2.4.3, clarified evaluable ADA samples for combination treatment for summaries. Section 4.1.3 and 6, clarified upper limit of visit windowing for safety data.
19Apr2023	Section 1.3, updated number of patients in STO cohort following CSP v6. Section 2.2, removed the need for programmed important protocol deviations, these will come straight from the CTMS log. Section 4.2.2.10 and 6, added analysis for dose-finding efficacy listing. Section 4.2.3.2, removed imAE summary table of the Starting steroid dose and Time to first steroid dose, these are produced as listings Section 4.2.3.4, removed VS change from baseline table to align with previous updates. Section 4.2.4.1, reinstated text regarding PK figures. Throughout, removed Non Hodgkin lymphoma (NHL) cohort from dose-expansion analysis following CSP v6.

1 STUDY DETAILS

This is the statistical analysis plan (SAP) for study D419EC00001. The SAP describes the statistical analyses specified in the clinical study protocol (CSP) version 6.0 in more detail; any changes with regards to what is already specified in the CSP will be described in Section **Error! Reference source not found.**.

1.1 Study objectives

1.1.1 Primary objectives

Dose-finding phase

Primary Objective:	**Outcome Measure:**
To determine the adult equivalent exposure/ maximum tolerated dose (MTD) /recommended Phase II (pediatric) dose (RP2D) of durvalumab in combination with tremelimumab and durvalumab as monotherapy following combination therapy.	Pharmacokinetics (PK) parameters (including maximum serum concentration (C_{max}), minimum serum concentration (C_{min}), area under the curve (AUC), and others). Time of RP2D assessment will be at the end of the dose-finding phase, when sufficient numbers of evaluable samples have been accrued.
To determine the safety profile of durvalumab in combination with tremelimumab and durvalumab as monotherapy following combination therapy.	Adverse events (AEs), vital signs, physical examinations, electrocardiograms (ECGs), and laboratory evaluations.

Dose-expansion phase

Primary Objective:	**Outcome Measure:**
To determine the preliminary antitumor activity of durvalumab in combination with tremelimumab and durvalumab as monotherapy following combination therapy at the recommended dose, using cohort-specific response criteria (e.g. Response Evaluation Criteria in Solid Tumors (RECIST) 1.1).	Objective response rate (ORR) as determined by the Investigator assessed RECIST 1.1 or alternative pre-specified tumor-specific response rates for different scoring systems. Assessment of antitumor activity will be specific to tumor cohort. Additional efficacy endpoints that will be collected include DoR, BoR, DCR, PFS,

	APF12, and APF18 based on RECIST 1.1assessed by the investigator, and OS, OS12, and OS24 as appropriate to each individual cohort.

1.1.2 Secondary objectives

Dose-finding and Dose-expansion Phase

Secondary Objective:	**Outcome Measure:**
To describe the PK of durvalumab and tremelimumab in combination and durvalumab as monotherapy following combination therapy in children and young adults with solid tumors.	Individual durvalumab and tremelimumab concentrations in serum, and PK parameters including C_{max}, C_{min}, AUC and other parameters where appropriate.
To determine the immunogenicity of durvalumab and tremelimumab in combination and durvalumab as monotherapy following combination therapy in children and young adults with solid tumors.	CCI [redacted]
To measure effects on immune checkpoint inhibition in response to routine immunizations (dose-expansion phase only).	
To evaluate immune activation and counts of natural killer (NK)-, B- and T-cells	

1.1.3 Safety objectives

Safety Objective:	Outcome Measure:
To determine the safety profile and tolerability of patients from dose-expansion cohort(s) treated with durvalumab in combination with tremelimumab every 4 weeks (q4w).	AEs, vital signs, physical examinations, ECGs, and laboratory evaluations.

1.1.4 Exploratory objectives

Exploratory Objective:	Outcome Measure:
CCI	

1.2 Study design

This is an open-label, non-randomized, international, multicenter study investigating durvalumab in combination with tremelimumab (q4w for 4 cycles only) followed by durvalumab monotherapy (q4w) in pediatric patients from birth to <18 years of age with relapsed or refractory malignant solid tumors. Each treatment cycle is 28 days.

The study will be conducted in 2 sequential phases; a dose-finding phase (Phase I) followed by a disease specific dose-expansion phase (Phase II). The purpose of this study is to identify the recommended Phase II dose (RP2D) that will be taken forward to determine preliminary antitumor activity of durvalumab in combination with tremelimumab in solid malignant tumors, using disease specific response criteria.

Dose-finding phase

The dose-finding phase of the study will be conducted using a 3 + 3 design to determine whether durvalumab and tremelimumab can be administered safely in children and if adult exposures can be achieved. Pediatric patients with relapsed or refractory solid malignant tumors including osteosarcoma and Ewing sarcoma (SARC-1), rhabdomyosarcoma, non-rhabdomyosarcoma, and other sarcomas (SARC-2), neuroblastoma (NB), and other solid tumors (STO), will be enrolled in 2 arms based on patient weight:

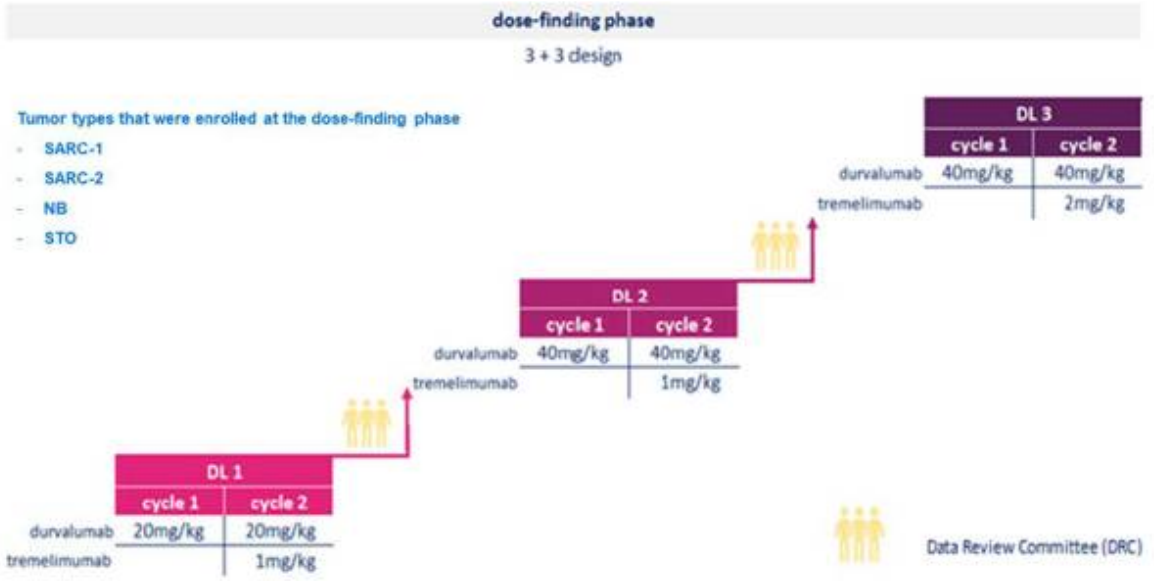
- Arm A: patients weighing ≥35kg
- Arm B: patients weighing <35kg

Three dose levels may be explored in each of the arms as outlined in Figure 1. Dose level 1 will be 100% of the recommended adult dose of both durvalumab and tremelimumab administered as weight-adjusted doses. The doses and schedules are as follows:

1 Cycle 1 (durvalumab monotherapy): durvalumab 20 mg/kg
2 Cycles 2 to 5 (durvalumab in combination with tremelimumab): durvalumab 20 mg/kg and tremelimumab 1 mg/kg [4 cycles administered every 28 days])
3 From Cycle 6 onwards, treatment will continue until a discontinuation criterion is met, with durvalumab monotherapy administered at 20 mg/kg q4w.

At least one day between first patient and subsequent patients dosed at each dose-level is required.

Figure 1 Dose levels – Dose-finding phase



This figure shows the maximum doses for durvalumab at DLs 2 and 3 (40 mg/kg) and tremelimumab at DL 3 (2 mg/kg). Lower doses may be tested, depending on the obtained PK modeling data.
DL Dose level; NB Neuroblastoma; SARC-1 Osteosarcoma and Ewing sarcoma; SARC-2 Rhabdomyosarcoma, non-rhabdomyosarcoma, and other sarcomas; STO Solid tumor other.

A Data Review Committee (DRC) will be established prior to the initiation of the study to make decisions on dose escalations and determination of a recommended phase 2 dose (RP2D) to be further explored in the dose-expansion phase.

Dose escalation decisions will aim to identify the RP2D for durvalumab and tremelimumab in combination. Dose escalation decisions and determination of RP2D will mainly consider safety and PK information obtained from Cycle 1 (durvalumab monotherapy) and Cycle 2 (first administration of durvalumab and tremelimumab combined). This includes assessment of pre-defined dose-limiting toxicities (DLTs) in the first 2 cycles.

A de-escalation step may occur for durvalumab or/and tremelimumab, with reduced doses of 15 mg/kg and 0.75 mg/kg, respectively, if dose level 1 is considered not tolerated. Furthermore, if exposure of durvalumab or/and tremelimumab proves inadequate, intermediate dose levels may also be tested.

At least 1 patient representing each of the following pediatric age subsets will be enrolled: birth to 5 years, 6 to 11 years and 12 to <18 years. In addition, 3 patients older than 2 years of age must be treated with durvalumab in combination with tremelimumab and clear the DLT period and have been evaluated by the DRC before children less than 2 years old may be enrolled in the dose-finding phase.

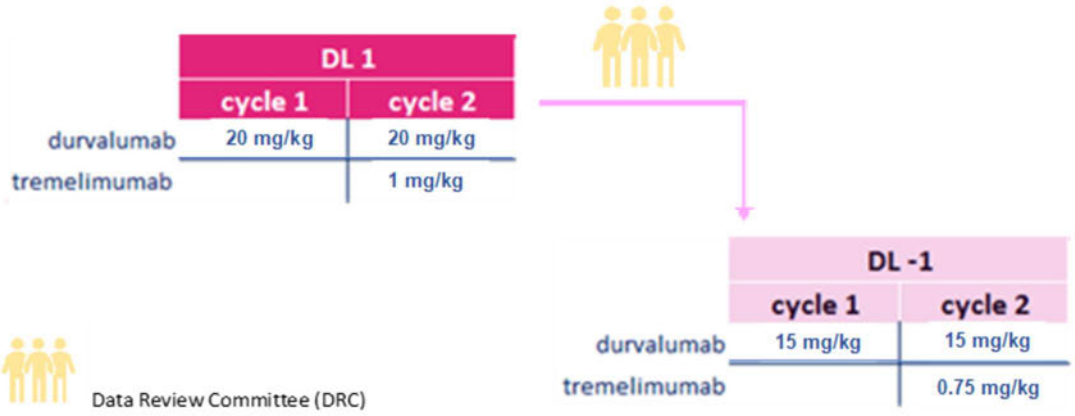
The key considerations to be followed by the DRC during the dose-finding phase of the study are outlined in Table 1:

Table 1 Key considerations for DRC

#DLTs	PK	Decision/Outcome
0 out of 3-5 OR **1** out of 6	Equivalent to adult exposure (5 out of 6 patients)	Declare tolerated and observed for 4 cycles, and if <2 patients with DLTs are observed during the first 4 cycles of study drug, enter dose-expansion stage
	<80% of adult exposure (≤2 patients)	Declare tolerated and move to next dose level.
1 out of 3		Increase number of patients to 6.
≥2		Declare not tolerated; if observed on dose level 1, move to dose level -1 (see Figure 2).

DLT dose limiting toxicity; DRC data review committee; PK pharmacokinetics.

Figure 2 Dose-finding: de-escalation schema



DL dose level.

A separate document (a DRC charter) will be created to detail all key tasks, working rules and timing of meetings of the DRC committee.

Dose-expansion phase

Once the RP2D has been established, recruitment into the dose-expansion study will be conducted using a Simon 2-stage optimal design with an additional provision to include 1 mixed disease cohort (other malignant solid tumors). Patients will be recruited into the cohorts outlined in Table 2 .

During dose-expansion, patients will be monitored for safety according to the same criteria employed during the dose-finding phase (although events will not be considered DLTs during dose-expansion). Additionally, 4 cycles of safety information are required prior to opening the initial dose-expansion cohorts. If during the treatment period, ≥33% of patients experience safety events meeting the DLT criteria defined for the dose-finding phase (see CSP section 4.1.2), even if outside of the DLT evaluation period, enrollment may be paused, and study data will be reviewed to determine whether additional monitoring or alternative dose levels or schedules should be evaluated prior to further enrollment.

Table 2 Malignant solid tumor cohorts

	Dose-finding phase
Cohort	**Type of cancer**
SARC-1	Osteo sarcoma or Ewing sarcoma

SARC-2	Rhabdomyosarcoma, non-rhabdomyosarcoma soft-tissue sarcoma, other sarcomas
NB	Neuroblastoma
STO	Other solid tumors (not represented in the other solid tumor cohorts)

Dose-expansion phase

Cohort	Type of cancer
SARCOMA	bone sarcomas: osteosarcoma or Ewing Sarcoma; soft-tissue sarcomas [≥ 40% of enrollment]: Rhabdomyosarcoma, non-rhabdomyosarcoma and other sarcomas
STO	Other solid tumors (not represented in the other solid tumor cohorts)

STO Solid tumors other.

Disease specific cohorts will not be separated into weight-based arms, but patients within the expansion cohorts will be dosed according to the appropriate weight-based RP2D determined in the arms of dose-finding cohorts.

For the SARCOMA cohort, the initial stage will allow 11 evaluable patients, to be dosed in this disease-specific cohort. If 2 or more responses are observed, additional patients will be accrued, as part of the second stage, into the corresponding expansion cohort; 15 patients will be accrued into the SARCOMA expansion cohort for a total of 26 evaluable patients. If there has been ≤1 objective response in either cohort, at the time that evaluable patients (11 patients) dosed in the initial stage who have been followed for at least 3 cycles and have been assessed, the cohort will be discontinued for lack of benefit.

For the STO cohort, 10 evaluable patients will be enrolled into this mixed-disease cohort; Simon rules are not applicable for this cohort.

Duration of treatment and criteria for restarting treatment with durvalumab in combination with tremelimumab

All treatment will be administered beginning on Day 1 until clinical progression or confirmed radiological progression (refer to CSP Appendix E for solid tumors, CSP Appendix F for NB,; other malignancies will be analyzed based on the best response assessed by the Investigator) unless there is unacceptable toxicity, withdrawal of consent, or another discontinuation criterion is met.

If during the durvalumab monotherapy component of the treatment (after the 4 cycles of durvalumab + tremelimumab, where applicable), a patient's disease progresses, re-administration of tremelimumab (in combination with durvalumab) for another 4 cycles may be considered, following the conditions below:

- Patients who complete the 4 dosing cycles of the combination of durvalumab and tremelimumab portion of the regimen (with clinical benefit as per Investigator judgement) but subsequently have evidence of defined PD, with or without confirmation, during the durvalumab monotherapy portion of the combination regimen may restart treatment with the combination once (4 cycles only). After this, durvalumab monotherapy will resume, if the Investigator judges there is clinical benefit to the patient.

1.3 Number of patients

Dose-finding phase

The final sample size in the dose-finding phase will depend on the number of DLTs and on the number of dose levels to be explored. In addition, any patients not evaluable (NE) will be replaced in order to have the required number of patients evaluable at each dose level. A minimum of 12 patients is anticipated to be dosed, to enable exploration of 1 dose level in both patients weighing 35 kg and greater, and patients weighing under 35 kg. If all 3 dose levels are used at both weight-based arms, then 36 evaluable patients would be necessary to complete the dose-finding phase. If approximately 20% of patients need to be replaced due to non-evaluability, then a maximum of 45 patients will be enrolled on the dose-finding cohorts.

Dose-expansion phase

For the dose-expansion phase, the SARCOMA cohort will follow a Simon 2-stage optimal design. An ORR ≤10% will not be considered clinically meaningful.

In the SARCOMA cohort, the Simon 2-stage design will have a type I error rate set at 0.1 (1-sided) and power of at least 85%. Based on the assumption that the ORR for the null hypothesis is 10% and the true ORR for the SARCOMA cohort is 30%, then the SARCOMA cohort will require a total of 26 evaluable patients (with the requirement that ≥ 40% of the patients enrolled must have soft-tissue sarcomas), with 11 evaluable patients required in the first stage. If fewer than 2 confirmed responses are observed in the first 11 evaluable patients, then the cohort will be closed for lack of benefit. If 2 or more responses are observed out of the first 11 evaluable patients, then an additional 15 evaluable patients will be enrolled, for a total of 26 evaluable patients.

For a cohort that enrolls the maximum evaluable patients for that particular cohort (26 in the SARCOMA cohort), if fewer than 5 responses are observed out of the maximum total of evaluable patients, then no further investigation of that cohort will be done; if 5 or more responses are observed out of the maximum total of evaluable patients, then the null hypothesis will be rejected for this cohort and further investigation will be warranted.

The STO cohort is planned to recruit 10 evaluable patients. This sample size is based on practical, rather than formal considerations, as no formal hypothesis testing is planned for this cohort.

With the above considerations, the sample size for evaluable patients in the dose-expansion phase is expected to range from 21 to 36. Assuming that about 20% of patients may not be evaluable for ORR, the total number of patients dosed in the dose-expansion phase may range from approximately 25 to 43.

2 ANALYSIS SETS

2.1 Definition of analysis sets

Details of the analysis sets are presented in Table 3 and Table 4.

Table 3 Analysis sets

Analysis Set	Definition	Phase derived for
Full analysis set (FAS)	The FAS will include all patients who were assigned to treatment and received at least 1 dose of study treatment. The FAS (or subset of the FAS specified below) will be used for all efficacy analyses.	Dose-finding and Dose-expansion

Evaluable for response analysis set	The subset of patients in the FAS who had
	• measurable disease (per RECIST 1.1) at baseline
	  ○ and had at least one follow-up scan measuring all required target lesions
	  ○ and have been followed for at least 3 cycles (to allow for a confirmatory scan at 4 weeks after the first assessment scan)
	OR
	• measurable disease (per RECIST 1.1) at baseline
	  ○ and progressed or died in the absence of a follow-up scan.
		Dose-expansion

Safety analysis set (SAS)	The SAS will consist of all patients who received any amount of study treatment. Safety data will be summarized using the SAS according to the treatment received.	Dose-finding and Dose-expansion

Table 3 Analysis sets

Analysis Set	Definition	Phase derived for
Dose-limiting toxicity (DLT) evaluable analysis set	The DLT evaluable analysis set is a subset of the SAS for the dose-finding phase of the study. It includes all patients enrolled in the dose-finding phase of the study who receive the protocol-assigned treatment with durvalumab + tremelimumab and complete the safety follow-up through the DLT evaluation period (Cycle 1 + Cycle 2) or experience a DLT during the DLT evaluation period.	Dose-finding
PK analysis set	All patients who receive at least 1 dose of study treatment per the CSP for whom any post dose data are available and who do not violate or deviate from the CSP in ways that would significantly affect the PK analyses will be included in the PK analysis set. The population will be defined by the Study Physician, Pharmacokineticist, and Statistician prior to any analyses being performed.	Dose-finding and Dose-expansion
ADA analysis set	All patients who receive at least 1 dose of study treatment per the CSP for whom baseline, and any post dose data are available will be included in the ADA analysis set.	Dose-finding and Dose-expansion

Table 4 Summary outcome variables and analysis sets

Outcome variable	Analysis set
Efficacy Data	
ORR (primary for dose-expansion phase [confirmed responses])	Evaluable for response analysis set (Sensitivity analysis on FAS)
ORR (unconfirmed responses)	(Sensitivity analysis on Evaluable for response analysis set)
PFS, OS, BoR (confirmed responses), APF12, APF18, OS12, OS24	FAS
DoR	Evaluable for response analysis set (restricted to confirmed responders only)
DCR	Evaluable for response analysis set (Sensitivity analysis on FAS)

Table 4 Summary outcome variables and analysis sets

Outcome variable	Analysis set
Demography	FAS
PK data	PK analysis set
ADA data	ADA analysis set
Safety Data	
Exposure	SAS
AEs	SAS
Laboratory measurements	SAS
Vital Signs	SAS
DLTs (primary for dose-finding phase)	DLT evaluable analysis set

AE Adverse event; APF12 / APF18 Proportion of patients alive and progression-free at 12 / 18 months from first dose of study treatment; BoR Best objective response; DCR Disease control rate; DLT Dose-limiting toxicity; DoR Duration of response; FAS Full analysis set; OS Overall survival; ORR Objective response rate; OS12 / OS24 Proportion of patients alive at 12 / 24 months from first dose of study treatment; PFS Progression-free survival; PK Pharmacokinetics; SAS Safety analysis set.

2.2 Violations and deviations

The following general categories will be considered important protocol deviations and will be raised in the Clinical Trial Management System (CTMS) log.. These will be listed and discussed in the clinical study report (CSR) as appropriate:

Dose-finding phase

- Deviation 1: Patients who received a dose of study treatment other than the dose that was intended

    1.1 Patient either did not receive any dose of planned study treatment or did not receive the correct dose of planned study treatment (i.e. less than 75% of the planned dose at the time of infusion) (as per Figure 1 in the CSP).

- Deviation 2: Patients who deviate from key entry criteria as per the CSP. These are inclusion criteria 2, 3, and 8 and exclusion criteria 2, 3, 5, 6, 7 and 11:

    2.1 Patient, older than 18 years, received study treatment (inclusion criteria 2)

2.2 Patient received study treatment without a histopathologic confirmation of a malignancy or/and have not progressed or were refractory to standard therapies or/and have not explored standard of care treatments before participation on this protocol. (inclusion criteria 3)

2.3 Patient with a hematological condition listed in the protocol received treatment

2.4 Inadequate organ and marrow function, defined as any of the following (inclusion criteria 8):

- Hemoglobin < 9.0 g/dL
- Absolute neutrophil count < 1.0×10^9/L
- Platelet count < 75×10^9/L
- Serum bilirubin > 1.5 × the ULN
- Creatinine clearance ≤ 60 mL/min
- Alanine transaminase (ALT) or aspartate transaminase (AST) >3 × upper limit of normal (ULN) (for patients without liver metastases)
- ALT or AST > 5 × ULN (for patients with liver metastases)

2.5 Patients with a history of primary immunodeficiency an uncontrolled intercurrent illness that would limit compliance with study requirements, substantially increase risk of incurring AEs from IP, or compromise the ability of the patient to give written informed consent received study treatment.

2.6 Patients with unresolved toxicities with a severity of grade ≥2 (as determined by NCI CTCAE version 5.0) from previous anticancer therapy (excluding alopecia, vitiligo, lymphopenia, the laboratory values defined in the inclusion criteria and toxicities not reasonably expected to be exacerbated by treatment with durvalumab or tremelimumab) received study treatment. (excluding criteria 6).

2.7 Patients with known or clinically suspected active brain metastases or spinal cord compression that have not been treated or/and were considered unstable or/and were dependent on steroids above physiologic replacement dose, received study treatment (exclusion criteria 7).

2.8 Hypothyroidism before the start of treatment and worsening at any stage before the start of cycle 3. (inclusion criteria 2)

2.9 Received the last dose of an anticancer therapy (chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumor embolization, or

mAbs) ≤28 days or less than 5 half-lives of the drug, whichever is shorter (and minimum of 7 days), prior to the first dose of study treatment (exclusion criteria 11).

2.10 Received any anti-cancer therapy (chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumor embolization, or mAbs) between Cycle 1, Day 1 and Cycle 3, Day 1 (exclusion criteria 11).

Patient was dosed with study treatment, but fulfilled any of the following criteria:

2.11 Have or have had documented autoimmune or inflammatory disorders (exclusion criteria 2) excluding:
- Patients with hypothyroidism not controlled by endocrine therapy, where 'controlled' is defined as having a thyroid stimulating hormone (TSH) measurement within normal range at baseline
- Known psoriasis that requires systemic therapy
- Celiac disease, not controlled by diet alone.

2.12 Have an active infection (exclusion criteria 3).

2.13 Have tuberculosis (defined by clinical diagnosis or radiographic findings and/or a positive skin test), hepatitis B (defined by a positive surface antigen result), active hepatitis C (defined by a positive polymerase chain reaction (PCR) for hepatitis C virus (HCV) ribonucleic acid (RNA)) or human immunodeficiency virus (HIV) (defined as positive HIV 1/2 antibodies) (exclusion criteria 5).

- Deviation 3: Received prohibited concomitant medications (including other anticancer agents). Please refer to the CSP section 6.4 for those medications that are detailed as being 'excluded' from permitted use during the study. This will be used as a guiding principle for the physician review of all medications prior to database lock to identify those likely have an impact on efficacy.

3.1 Prior treatment with immune-mediated therapy of any of the following
- Anti-PD-L1
- Anti-PD-1
- Anti- Programmed cell death ligand 2(PD-L2)
- Anti- Cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4)
- Antibodies of chimeric antigen receptor T-cell therapy (CAR-T) or other cell therapies

3.2 Receive a live attenuated vaccine within 30 days prior to the first dose of study treatment

3.3 Receive an immunosuppressive medication within 14 days prior to the first dose of study excluding;
- Intranasal, inhaled, topical steroids, or local steroid injections (e.g. intra-articular injection).
- Systemic corticosteroids at physiologic doses as replacement therapy, not to exceed 10 mg/m2/day of prednisone or its equivalent.
- Steroids as premedication for hypersensitivity reactions (e.g. computed tomography (CT) scan premedication).

Dose-expansion phase

- Deviation 1: Patients who received a dose of study treatment other than the dose that was intended

  1.1 Patient either did not receive any dose of planned study treatment or did not receive the correct dose of planned study treatment (i.e. less than 75% of the planned dose at the time of infusion) (as per Figure 1 in the CSP).

- Deviation 2: Patients who deviate from key entry criteria as per the CSP. These are inclusion criteria 2, 3, 6 and 8 and exclusion criteria 11:

  2.1 Patient, older than 18 years, received study treatment (inclusion criteria 2).

  2.2 Patient received study treatment without a histopathologic confirmation of a malignancy or/and have not progressed or were refractory to standard therapies or/and have not explored standard of care treatments before participation on this protocol. (inclusion criteria 3).

  2.3 Selection of a previously radiated tumor as a target lesion and tumor did not show disease progression prior to study entry (inclusion criteria 6).

  2.4 Inadequate organ and marrow function, defined as any of the following (inclusion criteria 8):

  - Hemoglobin < 9.0 g/dL
  - Absolute neutrophil count < 1.0×10^9/L
  - Platelet count < 75×10^9/L
  - Serum bilirubin > 1.5 × the ULN

- Creatinine clearance ≤ 60 mL/min
- Alanine transaminase (ALT) or aspartate transaminase (AST) >3 × upper limit of normal (ULN) (for patients without liver metastases)
- ALT or AST > 5 × ULN (for patients with liver metastases)

2.5 Received the last dose of an anticancer therapy (chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumor embolization, or mAbs) ≤28 days or less than 5 half-lives of the drug, whichever is shorter (and minimum of 7 days), prior to the first dose of study treatment (exclusion criteria 11).

2.6 Received any anti-cancer therapy (chemotherapy, immunotherapy, endocrine therapy, targeted therapy, biologic therapy, tumor embolization, or mAbs) between Cycle 1, Day 1 and Cycle 3, Day 1(exclusion criteria 11).

- Deviation 3: Baseline assessment > 21 days before start date of first dose of study treatment.

- Deviation 4: Received prohibited concomitant medications (including other anticancer agents). Please refer to the CSP section 6.4 for those medications that are detailed as being 'excluded' from permitted use during the study. This will be used as a guiding principle for the physician review of all medications prior to database lock to identify those likely have an impact on efficacy.

4.1 Prior treatment with immune-mediated therapy of any of the following
- Anti-PD-L1
- Anti-PD-1
- Anti- Programmed cell death ligand 2(PD-L2)
- Anti- Cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4)
- Antibodies of chimeric antigen receptor T-cell therapy (CAR-T) or other cell therapies

- Deviation 5: Had non-measurable disease at baseline

- Deviation 6: Had a solitary measurable lesion that was biopsied prior to study entry

- Deviation 7: Imaging modalities for imaging acquired at baseline and first subsequent scan were not consistent (e.g. bone scan used for baseline, followed by CT scan post baseline)

3 PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of visit tumor responses

The analysis of the primary endpoint, ORR, and the analyses of the secondary endpoints, PFS, DoR, and DCR, will be based on the site Investigator assessments using the following criteria

Table 5 Criteria for Investigator assessments

Type of cancer	Criteria
Solid Tumors	
SARCOMA and STO	RECIST 1.1 (see CSP Appendix E)

All RECIST 1.1, whether scheduled or unscheduled, will be included in the calculations. This is also regardless of whether a patient discontinues study treatment or receives another anticancer therapy.

Baseline will be assessed within the 21 days prior to first dose of study treatment and ideally as close as possible to the start of study treatment. Assessments are performed every 8 weeks (relative to the date of first dose of study treatment) in the first instance. For patients who achieve an objective response or stable disease, and for whom the response is still evident at 12 months, subsequent tumor assessments are to occur every 16 weeks ± 1 week until disease progression.

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 minimise any unintentional bias caused by some patients being assessed at a different frequency than other patients.

3.1.1 Derivation of RECIST 1.1 visit responses

For all patients with solid tumors excluding NB (SARC-1, SARC-2 and STO (dose-finding), SARCOMA and STO (dose-expansion)), the RECIST tumor response data will be used to determine each patient's visit response according to RECIST 1.1. It will also be used to determine if and when a patient has progressed in accordance with RECIST 1.1 and their BoR to study treatment.

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 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 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 NE, (unless there is evidence of progression in which case the response will be assigned as PD).

RECIST outcomes (i.e. PFS, ORR etc.) will be calculated programmatically for the site Investigator data (see Section 3.2) from the overall visit responses.

3.1.1.1 Target lesions (TLs) – site Investigator data

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (LD), (except lymph nodes which must have short axis ≥ 15 mm) with CT or magnetic resonance imaging (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 (representative of all lesions involved and suitable for accurate repeated measurement) and these are referred to as TLs. If more than one baseline scan is recorded, then measurements from the one that is closest and prior to first dose will be used to define the baseline sum of TLs. 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.

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.

Note: For patients who do not have measurable disease at entry (i.e. no TLs) but have non-measurable disease, evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions (see Section 3.1.1.3 for further details). If a patient does not have measurable disease at baseline, then the TL visit response will be not applicable (NA).

Table 6 TL Visit Responses (RECIST 1.1)

Visit Response	Description
CR	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10 mm.
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.
SD	Neither sufficient decrease in sum of diameters to qualify for PR nor sufficient increase to qualify for PD.
PD	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest previous sum of diameters since treatment started including the baseline sum of diameters.
NE	Only relevant if any of the TLs were not assessed or NE or had a lesion intervention at this visit. Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides NE as a TL response.
NA	No TLs are recorded at baseline.

CR Complete response; PD Progressive disease; PR Partial response; SD Stable disease; NA Not applicable; NE Not estimable.

Rounding of TL data

For calculation of PD and PR for TLs percentage changes from baseline and previous minimum should be rounded to one 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

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

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 0mm 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

A CR can only be followed by CR, PD or NE. 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. 0mm or < 10mm for lymph nodes) then response will be set to CR irrespective of whether the criteria for PD of TL is also met i.e. if a lymph node LD increases by 20% but remains < 10mm.
- Step 2: If some lesion measurements are missing but all other lesions meet the CR criteria (i.e. 0mm or < 10mm 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 and the sum of lesions meets the criteria for PD 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 measured 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 blinded to treatment assignment. 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 eCRF and a value of 5mm 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 eCRF and has entered a smaller value that can be reliably measured. If a TL response of PD results, 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 / embolization), should be handled in the following way and

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

At subsequent visits, the above steps will be repeated to determine the TL and overall visit response. 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 has still increased by 20% or more compared to nadir and the sum of TLs has increased by ≥5mm 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	Longest diameter at nadir visit	Longest diameter at follow-up visit
1	7.2	7.1
2	6.7	6.4
3	4.3	4.0
4	8.6	8.5
5	2.5	Intervention
Sum	**29.3**	**26**

Lesion 5 is missing at the follow-up visit; there was a baseline TL sum of 29.3cm.

The sum of lesions 1-4 at the follow-up is 26 cm. The sum of the corresponding lesions at the nadir visit is 26.8 cm.

Scale up as follows to give an estimated TL sum of 28.4cm:

$$26/26.8 \times 29.3 = 28.4\ cm$$

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

Lesions that split in two

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

Lesions that merge

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

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. The TL visit response may still be evaluable if the number of missing TL measurements at a visit is ≤1/3 of the total number of TLs.

3.1.1.2 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 investigational site at each visit.

NTL response will be derived based on the Investigator's overall assessment of NTLs as follows

Table 7 NTL Visit Responses (RECIST 1.1)

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

CR Complete response; NTL Non-target lesion; PD Progressive disease; PR Partial response; SD Stable disease; TL Target lesions; NA Not applicable; NE Not estimable.

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 one 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 one or more new lesions is assessed as progression.

A lesion identified at a follow up assessment in an anatomical location that was not scanned at baseline is considered a new lesion and 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 the 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 any assessment of NTLs.

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

3.1.1.3 Overall visit response – site Investigator data

Table 8 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 8 Overall visit responses

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

CR Complete response; PD Progressive disease; PR Partial response; SD Stable disease; NA Not applicable; NE Not estimable; NED No evidence of disease.

3.2 Efficacy Variables

3.2.1 Objective response rate (ORR)

ORR (per RECIST 1.1, using Investigator assessments) is defined as the percentage of patients with at least 1 investigator-assessed visit response of CR or PR that is subsequently confirmed on another scan not less than 4 weeks after visit observed response. Data obtained up until progression, or the last evaluable assessment in the absence of progression, will be included in the assessment of ORR. Patients who go off treatment without progression, receive a subsequent therapy, and then respond will not be included as responders in the ORR.

ORR is primary endpoint of dose-expansion phase.

3.2.2 Progression-free survival (PFS)

Progression free survival (per RECIST 1.1, as assessed by the site Investigator) will be defined as the time from the date of first dose of study treatment until the date of objective disease progression or death (by any cause in the absence of progression) regardless of

whether the patient withdraws from study therapy or receives another anti-cancer therapy prior to progression (i.e. date of PFS event or censoring – date of first dose + 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 after 2 or more 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 scheduled visit assessment scheme (i.e. eight-weekly (+/-1 week) for the first 12 months with the option of moving to every 16 weeks ± 1 week thereafter (for those with a sustained objective response), until clinical progression/deterioration) the definition of 2 missed visits will change over time on a per patient basis.

If the previous (RECIST 1.1) assessment is between study day 49 (i.e. week 7) and less than the last assessment on an 8-weekly schedule then two missed visits will equate to 18 weeks since the previous RECIST 1.1 assessment, allowing for early and late visits (i.e. 2 x 8 weeks + 1 week for an early assessment + 1 week for a late assessment = 18 weeks). If the two missed visits occur over the period when the scheduled frequency of RECIST 1.1 assessments changes from eight-weekly to sixteen-weekly this will equate to 26 weeks (i.e. take the average of 8 and 16 weeks which gives 12 weeks and then apply same rationale, hence 2 x 12 weeks + 1 week for an early assessment + 1 week for a late assessment = 26 weeks). Once the scheduling changes to sixteen-weekly assessments, two missing visits will equate to 34 weeks (i.e. 2 x 16 weeks + 1 week for an early assessment + 1 week for a late assessment = 34 weeks).

If a 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 (16 weeks plus 1 week allowing for a late assessment within the visit window) then they will be treated as an event with date of death as the event date.

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

RECIST 1.1 assessments/scans contributing towards a particular visit may be performed on different dates. The following rules will be applied:

- For Investigator assessments, date of progression will be determined based on the earliest of the assessment/scan dates of the component that triggered the progression.
- 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.

3.2.3 Overall Survival (OS)

OS is defined as the time from the date of first dose of study treatment until death due to any cause regardless of whether the patient withdraws from study treatment or receives another anti-cancer therapy (i.e. date of death or censoring – date of first dose + 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. ('Date subject confirmed to be alive', recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made following the date of data cut off (DCO) for the analysis (these contacts should generally occur within 7 days of the DCO) and if patients are confirmed to be alive or if the death date is post 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 the final OS analysis should be obtained by the site personnel by checking the patient's notes, hospital records, contacting the patient's general practitioner and checking publicly-available death registries. In the event that the patient has actively withdrawn consent to the processing of their personal data, the vital status of the patient can be obtained by site personnel from publicly available resources where it is possible to do so under applicable local laws.

Note that 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 eCRF fields to determine the last recorded date on which the patient was known to be alive for those patients still on treatment. The last date for each individual patient is defined as the latest among the following dates recorded on the eCRF:

- AE start and stop dates
- Admission and discharge dates of hospitalization
- Study treatment date
- End of treatment date
- Laboratory test dates
- Date of vital signs
- Date of ECG

- Disease assessment dates on RECIST/non-RECIST eCRF
- Start and stop dates of alternative anticancer treatment
- Date last known alive on survival status eCRF
- End of study date
- Date of any other safety data

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 the last date known to be alive +1 from the database and the death date using the available information provided:

a. For Missing day only – using the 1st of the month

b. For Missing day and Month – using the 1st of January

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.

3.2.4 Duration of response (DoR)

DoR (per RECIST 1.1, using Investigator assessment) will be calculated for all patients with a confirmed overall response of CR or PR, and will be defined as the time from the date of first documented response (which is subsequently confirmed) until the first date of documented progression or death in the absence of disease progression (i.e. date of PFS event or censoring – date of first response + 1). The end of response should coincide with the date of progression or death from any cause used for the (RECIST 1.1) PFS endpoint.

The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of PR or CR that was subsequently confirmed. If a patient does not progress following a response, then their DoR will be censored at the PFS censoring time.

DoR will not be defined for those patients who do not have documented response.

3.2.5 Best objective response (BoR)

BoR is calculated based on the overall visit responses from each (RECIST 1.1, using Investigator assessment) assessment. It is the best response a patient has had during their time in the study, but prior to starting any subsequent cancer therapy and up until (RECIST 1.1) progression or the last evaluable assessment in the absence of (RECIST 1.1) progression.

Categorization of BoR for solid tumors will be based on RECIST 1.1 using the following response categories: CR, PR, SD, PD, and NE.

CR or PR must be confirmed (confirmed is defined as 2 consecutive assessments with the same outcome, not less than 4 weeks apart). BoR will be determined programmatically (based on RECIST 1.1) using all Investigator assessments up until the first progression event. For patients whose progression event is death, BoR will be calculated based upon all evaluable (RECIST 1.1) assessments prior to death.

For patients who die with no evaluable (RECIST 1.1) assessments, if the death occurs ≤119 days (i.e. 2 × 8 weeks + 7 days) after first dose of study treatment, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable (RECIST 1.1) assessments, if the death occurs >119 days (i.e. 2 × 8 weeks + 7 days) after the date of first dose of study treatment, then BoR will be assigned to the NE category.

Progression events that have been censored due to them being >119 days after the last evaluable assessment will not contribute to the BoR derivation.

3.2.6 Disease control rate (DCR)

DCR at 16 or 24 weeks will be defined as the proportion of patients who achieve a BoR of unconfirmed CR or PR, respectively, or who have SD (without subsequent cancer therapy) for at least 16 weeks − 7 days or 24 weeks - 7 days, respectively, after start of treatment. For NB, MR will be included in the DCR count.

DCR will be determined programmatically (based on RECIST 1.1, using Investigator assessment) from the overall visit response using all data up until the first progression event.

3.2.7 Proportion alive and progression free at 12 (APF12) and 18 (APF18) months

APF12 will be defined as the Kaplan-Meier (KM) estimate of PFS (per RECIST 1.1, using Investigator assessment) at 12 months. Similarly, APF18 will be defined as KM estimate of PFS at 18 months.

3.2.8 Overall survival at 12 (OS12) and 24 (OS24) months

OS12 will be defined as the KM estimate of OS at 12 months. Similarly, OS24 will be defined as the KM estimate of OS at 24 months.

3.3 Safety Variables

Safety and tolerability will be assessed in terms of AEs (including serious adverse events [SAEs]), deaths, laboratory data, vital signs, ECGs, physical examinations and exposure. These will be collected for all patients.

The Safety analysis set (SAS) will be used for reporting of safety data.

On-treatment is defined as assessment where the date and time of the assessment is the same as or greater than the date and time of first treatment with either tremelimumab or durvalumab and up to (and including) 90 days after the last dose date of tremelimumab or durvalumab. For vital signs, On-treatment is defined as assessment where the date of first treatment with either tremelimumab or durvalumab and update (including) 30 days after the last dose date of tremelimumab or durvalumab.

3.3.1 Adverse Events

AEs will be assigned to the following periods related to date and time of onset of the AEs relative to first and last dose dates of study treatment and the date of starting any subsequent anti-cancer therapy post-discontinuation of study treatment.

AE occurrence	Period Assigned
Date before first treatment with either durvalumab or tremelimumab	Pre-treatment
Date the same as or greater than the date of first treatment with durvalumab or tremelimumab, either as monotherapy or combination therapy (or date before first treatment, but worsening following first dose) and up to (and including) 90 days after the last dose date of durvalumab or tremelimumab, or initiation of the first subsequent anti-cancer therapy following discontinuation of study treatment (whichever occurs first)	On-treatment, prior to subsequent anti-cancer therapy
Date the same as or greater than the date of first treatment with durvalumab or tremelimumab, either as monotherapy or combination therapy (or date before first treatment, but worsening following first dose) and up to (and including) 90 days after the last dose date of durvalumab or tremelimumab	On-treatment
Date after initiation of the first subsequent anti-cancer therapy following discontinuation of study treatment up to (and including) 90 days after the last dose date of durvalumab or tremelimumab	On-treatment Post-subsequent anti-cancer therapy
Date more than 90 days after the last dose date of either durvalumab or tremelimumab	Post-study treatment

AE Adverse event.

AEs that have missing causality (after data querying) will be assumed to be related to study treatment.

The latest version of the Medical Dictionary for Regulatory Activities (MedDRA) will be used to code the AEs. AEs will be graded by the Investigator according to the latest version of the National Cancer Institute Common Terminology Criteria for AEs (CTCAE).

AEs of special or possible interest

Some clinical concepts (including some selected individual preferred terms [PTs] and higher-level terms [HTLs]) have been considered "AEs of special interest" (AESI) or "AEs of possible interest" (AEPI) to the durvalumab program. AESI/AEPIs represent pre-specified risks that are considered to be of importance to a clinical development program.

These AESI/AEPIs have been identified as a list of categories provided by the patient safety team.

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

AESI/AEPIs observed with durvalumab ± tremelimumab include but are not limited to:

- Diarrhea/colitis and intestinal perforation
- Pneumonitis/ interstitial lung disease
- Endocrinopathies (i.e. events of hypophysitis/hypopituitarism, adrenal insufficiency, hyper and hypothyroidism and type I diabetes mellitus)
- Hepatitis/transaminase increases
- Nephritis/blood creatinine increases
- Pancreatitis/serum lipase and amylase increases
- Rash/dermatitis
- Myocarditis
- Myositis/polymyositis
- Other inflammatory responses that are rare/less frequent with a potential immune-mediated etiology include, but are not limited to, pericarditis, sarcoidosis, uveitis, and other events involving the eye, skin, hematological, and rheumatological events.

In addition, infusion-related reactions and hypersensitivity/anaphylactic reactions with a different underlying pharmacological etiology are also considered AESIs.

Other categories may be added as necessary or existing terms may be merged. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which HLTs and which PTs contribute to each AESI/AEPI. Further reviews may take place prior to database lock to ensure any further terms not already included are captured within the categories.

Immune-Mediated Adverse Events (imAE)

Sponsor will perform medical review of those AESIs and classify them as immune- mediated AEs (imAEs) or not imAEs. Further details are provided in an imAE Charter.

3.3.2 Laboratory data

Blood and urine samples for determination of clinical chemistry, hematology, and urinalysis will be taken at the times indicated in the assessment schedules and as clinically indicated Sections 1.1 and 8.2.1 of the CSP.

Change from baseline in hematology and clinical chemistry variables will be calculated for each post-dose visit on treatment. CTCAE grades will be defined at each visit according to the CTCAE grade criteria using local ranges as required, after conversion of lab result to corresponding preferred unit. The following parameters have NCI CTCAE grades defined for both high and low values: Hemoglobin, Leukocytes, Lymphocytes, Potassium, Sodium, Magnesium and Corrected calcium so high and low CTC grades will be calculated.

Absolute values will be compared to the local laboratory reference ranges and classified as low (below range), normal (within range or on limits of range) and high (above range).

The maximum or minimum on-treatment value (depending on the direction of an adverse effect) will be defined for each laboratory parameter as the maximum (or minimum) post-baseline value at any time.

Corrected calcium product will be derived during creation of the reporting database using the following formula:

Corrected calcium (mmol/L) = Total calcium (mmol/L) + ([40 – albumin (G/L)] × 0.02)

Project reference ranges will be used for the primary interpretation of laboratory data. Local reference ranges will be used for analysing laboratory data at the local laboratory. The denominator used in laboratory summaries of CTCAE grades will only include evaluable patients, in other words those who had sufficient data to have the possibility of an abnormality.

For example:

- If a NCI CTCAE criterion involves a change from baseline, evaluable patients would have both a pre-dose and at least 1 post dose value recorded.
- If a NCI CTCAE criterion does not consider changes from baseline to be evaluable, the patient need only have 1 post dose value recorded.

3.3.3 ECG data

ECGs will be recorded at screening and as clinically indicated throughout the study (see Section 1.1 of the CSP).

3.3.4 Vital signs

Changes in vital signs variables between baseline and each subsequent scheduled visit will be calculated.

Body surface area (BSA) will be calculated from the height (in centimeters) and weight (in kilograms) as follows:

$$\text{BSA} = \sqrt{((\text{Weight (kg)} * \text{height(cm)})/3600}$$

or (Du Bois formula)

$$BSA = 0.007184 \times W^{0.425} \times H^{0.725}$$

3.3.5 Exposure and dose intensity

Duration of exposure is defined as:

- Total treatment duration (in days) = min (date of last dose>0mg +27 days, date of death, date of DCO) - date of first dose +1.
- Actual treatment duration = total treatment duration, excluding the duration of dose delays.

Exposure will be defined separately for durvalumab monotherapy, and for combination treatment.

For the combination therapy, date of first dose refers to first dose of tremelimumab and date of last dose refers to the latest last dose of either durvalumab or tremelimumab, when

durvalumab is not being received as monotherapy, and a dose delay is any day(s) when neither of these treatments was received.

If a patient has more than one period of combination therapy, or monotherapy, then each period will be summed when deriving the exposure to durvalumab monotherapy and combination treatment.

In addition, the total number of dosing occasions will be calculated per patient. The calculation of actual exposure makes no adjustment for any dose reductions that may have occurred.

The relative dose intensity (RDI) will be calculated for each patient for both durvalumab and tremelimumab. RDI is the percentage of the actual dose delivered relative to the intended dose through to treatment discontinuation. RDI will be defined as follows:

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

The actual cumulative dose (mg/kg) per patient in a time period is the sum of the actual dose levels that the patient received within that period (i.e. total dose administered mg/kg).

Treatment Delays

Delays will be derived based on infusion date and will be based on the deviation of the actual to the planned treatment administration day (relative to the previous treatment administration date). A delay is defined as more than 1 day of delay between the actual and the planned treatment administration day. For example, if one patient receives the study drug on day 1, then the next study drug administration date will be on day 29 (Day 1 of next cycle); however, if the patient receives the study drug at day 31 (Day 3 of next cycle), this is considered as a delay of 2 days. For all dosing dates, the total duration of dose delays = Sum of (date of the dose - date of the previous dose – 28 days). Thus, if no delays were encountered, the duration would sum up to 0, since infusions were done every 4 weeks.

3.3.6 Dose Limiting Toxicities (DLTs)

Dose-limiting toxicities will be determined during Cycles 1 and 2 of the dose-finding phase using CTCAE version 5. Toxicities that occur during the first 28-day cycle and meet the criteria below will be assessed for attribution to durvalumab monotherapy. Any of the below listed AEs encountered between Cycle 1 Day 1 and Cycle 2 Day 28 (ie, 56 day DLT monitoring period for combination therapy) and in the opinion of the Investigator is thought to

be attributable to durvalumab or/and tremelimumab, given a reasonable possibility based on temporal exposure to IP and for which an alternate etiology does not exist or cannot be identified, will be classified as DLTs:

- Grade 5 toxicity

- Hematological toxicity: Any Grade ≥3 (including febrile neutropenia Grade ≥3) except for the following:
  - Grade 3 thrombocytopenia that does not result in a bleeding event or does not require platelet transfusion
  - Grade 3 neutropenia that resolves to Grade ≤2 within 7 days.
- Non-hematological toxicity:
  - Any Grade 4 immune-related AE (irAE)
  - Any ≥ Grade 3 colitis
  - Any Grade 3 or 4 noninfectious pneumonitis
  - Any Grade 2 pneumonitis that does not resolve within 5 days to ≤ Grade 1 after initiation of maximal supportive care (including use of steroids)
  - Any Grade 3 irAE, excluding colitis or pneumonitis, that does not downgrade to Grade 2 within 3 days after onset of the event despite optimal medical management including systemic corticosteroids and does not downgrade to ≤ Grade 1 or baseline within 14 days of onset
  - Liver transaminase elevation >5 × ULN or total bilirubin >3 × ULN. For patients with liver metastases at baseline, transaminase elevation >8 × ULN will constitute a DLT
  - Grade 4 vomiting, diarrhea, electrolyte abnormality, or systemic reaction
  - Grade 3 toxicity lasing >7 days despite optimal supportive care
  - Any ≥ Grade 3 non-irAE, except for the exclusions listed below
  - Any Grade 3 immune-mediated peripheral neuropathy or other immune-mediated neurotoxicity
  - Grade ≥3 inflammatory reactions.

The definition of DLTs excludes the following conditions:

- Grade 3 fatigue lasting ≤7 days.

- Grade 3 endocrine disorder (pituitary and/or adrenal insufficiency) that is managed with or without systemic corticosteroid therapy and/or hormone replacement therapy and the patient is asymptomatic.

- Grade 4 hypothyroidism adequately treatment with hormone replacement.

- Grade 3 inflammatory reaction attributed to a local antitumor response (eg, inflammatory reaction at sites of metastatic disease, lymph nodes, etc.).

- Concurrent vitiligo or alopecia of any AE grade.
- Grade 3 infusion-related reaction (first occurrence and in the absence of steroid prophylaxis) that resolves within 6 hours with appropriate clinical management.
- Grade 3 fever (fever greater than 40ºF for ≤24h in duration).
- Any grade lymphopenia.
- Isolated Grade 3 electrolyte abnormalities that are not associated with clinical signs or symptoms and are reversed with appropriate maximal medical intervention (including supplementation) within 7 days of initiating supplementation.

irAEs are defined as AEs of immune nature (ie, inflammatory) in the absence of a clear alternative etiology. In the absence of clinical abnormality, repeat laboratory testing must be conducted to confirm significant laboratory findings prior to designation as a DLT.

Guiding principles:

1. While the rules outlined above dictate the classification of AEs as DLTs, any other AEs that are observed during the observation period may be defined as a DLT in consultation with the DRC; after having taken into consideration the emerging safety profile of durvalumab and tremelimumab.
2. The NCI CTCAE (version 5.0) will be used for AEs reporting and its severity will be graded on a scale from 1 to 5 provided for each AE term.
3. An immune-mediated AE is defined as an AE that is associated with exposure to durvalumab or/and tremelimumab and is consistent with an immune-mediated mechanism of action after alternate etiologies have been considered and excluded.
4. In the absence of clinical findings, repeat laboratory testing must be conducted to confirm the clinical significance of abnormal laboratory test result before it is classified as a DLT.

For reporting purposes, the DLTs are as approved and documented at the DRC meetings.

3.4 Pharmacokinetic and Immunogenicity Variables

3.4.1 PK parameters

PK parameters (including C_{max}, C_{min}, AUC, and others) is primary endpoint of dose-finding phase and will be used to facilitate the determination of the RP2D.

Individual durvalumab and tremelimumab concentrations in serum (C_{max} and C_{min}), time to C_{max} and AUC will be used to describe the PK of durvalumab and tremelimumab.

Pharmacokinetic non-compartmental analysis

The actual sampling times will be used in the PK calculations. PK parameters will be determined using standard non-compartmental methods.

3.4.2 ADAs

The presence of ADA will be assessed in serum samples for patients with solid tumors. Samples will be measured for the presence of ADAs and ADA-neutralizing antibodies for both durvalumab and tremelimumab using validated assays.

3.5 Biomarker Variables

The following exploratory biomarkers will be collected

CCI

CCI

[redacted]

4 ANALYSIS METHODS

4.1 General principles

The primary analysis cut off will be defined at approximately 6 months after enrollment of the last patient. At this point, data analysis will be performed, and the CSR will be written. After this, patients still in the study will continue to be followed and a CSR addendum will be issued for OS and safety data collection.

For PK and other biomarkers, the sample bioanalysis will be performed by AstraZeneca or a contracted laboratory. PK parameters will be derived by Labcorp. The merging of PK concentration data with actual elapsed time from dose for PK sampling times will be performed by IQVIA Programming. The PK summaries and data listings will be the responsibility of IQVIA Biostatistics.

Statistical summaries and analyses of non-PK data will be performed by IQVIA under the direction of the Biostatistics Group, AstraZeneca using SAS® version 9.4 or higher and, where appropriate, additional validated software.

There is no formal statistical testing for the dose-finding phase of the study.

The dose-expansion phase of the study will formally test the following hypothesis:

- H0: ORR ≤10%
- H1: ORR >10%.

The test will be done at the 1-sided 5% level and it will be performed for the SARCOMA cohort.

No formal statistical testing will be performed for the STO cohort.

For the dose-finding phase, and dose-expansion phase where relevant, dose levels will be labelled in summary tables as:

- durvalumab xx mg/kg + tremelimumab xx mg/kg

with xx depending on the dose levels used.

For the dose-expansion phase, cohorts will be labelled in summary tables as:

- SARCOMA
- STO

Total columns will be included for the dose-finding and dose-expansion phases separately for outputs where a total column is required.

Visits will be present as cycle and day (e.g. Cycle 1 Day 1). Study day will be calculated as follows:

Days prior to first dose: Study day = date – first dose date

Days on or after first dose: Study day = date – first dose date + 1

First dose refers to either durvalumab or tremelimumab, whichever was administered first. Dates of progression will not be imputed.

The end of study is defined as the last expected visit/contact of the last patient undergoing the study. A patient is considered to have completed the study when he/she has completed his/her last scheduled visit shown in the schedule of activities (CSP section 1.1).

Categorical data will be summarized by the number and percentage of patients in each category. Continuous data will be summarized using descriptive statistics, including n, arithmetic mean, STD, median, minimum, and maximum values. Additionally, geometric means and coefficient of variation (CV)% will be reported for PK variables (concentrations and all PK parameters. The geometric mean is calculated as the exponential of the arithmetic mean calculated from data on a log scale. All data will be listed.

4.1.1 Rounding and precision

- For continuous data, the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation (STD) 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 rounded to 1 decimal place.

If further guidance is needed the methods described in Section 5.2 of 'Statistical guidelines for Clinical Pharmacology Alliance studies' will be applied.

4.1.2 Baseline measurements and change from baseline

Baseline will be the last non-missing value obtained prior to the first dose/administration of study treatment (durvalumab or tremelimumab) and any information taken after first dose/administration of study treatment is regarded as post baseline information in the relevant phase. If two visits are equally eligible to assess patient status at baseline (e.g., screening and baseline assessments both on the same date prior to first dose/administration with no washout or other intervention in the screening period) and time of assessment is not available, the average should be taken as the baseline value. For non-numeric laboratory tests (i.e. some of the urinalysis parameters) where taking an average is not possible then the best value would be taken as baseline as this is the most conservative. In the scenario where there are two assessments on the same day one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. If no value exists before the first dose/administration, then the baseline value will be treated as missing. For questionnaires where multiple records are eligible to be used as the baseline value and time of assessment is not available, the best scenario will be flagged as this is the most conservative.

Absolute change from baseline outcome variables is computed as

(post baseline value – baseline value).

Percent change from baseline is computed as

((post baseline value – baseline value) / baseline value) × 100%.

If either the post-baseline value or the baseline value is missing, then the absolute or percent change from baseline value will also be set to missing.

4.1.3 Time Windows

Safety assessments

Time windows will need defining for any presentations that summarize values by visit. The following conventions apply:

- The time windows should be exhaustive so that data recorded at any time point has the potential to be summarized. Inclusion within the time window should be based on the actual date and not the intended date of the visit.
- All unscheduled visit data should have the potential to be included in the summaries.
- 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 two visits (the lower limit of the first post-baseline visit will be Day 2). If an even number of days exists between two 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 in the dose-expansion phase for patients' treatment with durvalumab in combination with tremelimumab are:
    - Cycle 1 Day 8 visit window day 2-11
    - Cycle 1 Day 15 visit window day 12 - 21
    - Cycle 2 Day 1 visit window day 22 - 35
    - Cycle 2 Day 15 visit window day 36 - 49
    - Cycle 3 Day 1 visit window day 50 – 63

      etc
- The upper limit for the last record on treatment will consider end of treatment date and not date of disease progression.
- For summaries showing the maximum or minimum values, the maximum/minimum value recorded on treatment will be used (regardless of where it falls in an interval).
- Listings should display all values contributing to a time point for a patient.
- For visit based summaries:
    - If there is more than one value per patient within a time window then the closest value to the scheduled visit date should be used, or the earlier in the event the values are equidistant from the nominal visit date. The listings should highlight the value for that patient that went into the summary table, wherever feasible. Note: in summaries of extreme values all on-treatment values collected are used including those collected at unscheduled visits.
    - To prevent very large tables or plots being produced that contain many cells with meaningless data for outputs in the dose expansion phase, for each treatment group visit data should only be summarized if the number of observations is greater than the minimum of 5 of patients dosed.

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

4.2 Analysis methods

The two phases of the study will be summarized and analyzed separately. For the dose-finding phase, summaries will be presented by dose level. For the dose-expansion phase, summaries and analyses will be presented by cohort (see section 4.1).

For the dose-finding phase, the arms based on patient weight (Arm A: patients ≥35 kg and Arm B: patients <35 kg) will be summarized separately, along with an overall summary by dose level.

4.2.1 Multiplicity

Not applicable for this study.

4.2.2 Efficacy analysis

The following table (Table 9) presents the analyses planned for the efficacy endpoints. All tumor-related endpoints will be analyzed using Investigator RECIST 1.1 assessments. The efficacy analyses will be produced for the dose-expansion patients only.

Table 9 Pre-planned statistical analysis to be conducted

Endpoint analyzed	Notes
ORR	Exact 90% 2-sided CI (Mid-P)
DoR	Median estimated from KM curve
DCR	Exact 90% 2-sided CI (Mid-P)
BoR	n (%) of patients in each response category
PFS	Median estimated from KM curve
OS	Median estimated from KM curve
APF12	90% 2-sided CI estimated from KM curve
APF18	90% 2-sided CI estimated from KM curve
OS12	90% 2-sided CI estimated from KM curve
OS24	90% 2-sided CI estimated from KM curve

CI Confidence interval; KM Kaplan-Meier.

4.2.2.1 Objective response rate (ORR)

The ORR will be based on the programmatically derived (RECIST 1.1) outcome. The point estimate and 90% confidence interval (CI) (2-sided) of the ORR will be presented for each if the three analysis cohorts. If the cohort is expanded after the initial DRC review, the CI will be calculated using an exact Mid-P method (Porcher and Desseaux 2012). If the cohort is not expanded an exact CI will be presented.

90% Mid-p confidence limits are the values that have a one-sided Mid-p value of 0.05. (A Mid-p value for a one-sided test is obtained by including in the tail only one-half of the probability of the observed sample).

The exact Mid-p confidence limits θ_L (lower limit) and θ_U (upper limit) are those values which satisfy the equations:

$$\sum_{i=0}^{n} p_i(\theta_L) - \frac{1}{2}p_n(\theta_L) = 0.95$$

$$\sum_{i=0}^{n} p_i(\theta_U) - \frac{1}{2}p_n(\theta_U) = 0.05$$

This analysis will be performed on the evaluable for response analysis set.

Summaries will be produced that present the number and percentage of patients with a confirmed tumor response (CR/PR).

Sensitivity analyses will be performed for tumor response (CR/PR) on FAS and "unconfirmed CR/PR" on the evaluable for response analysis set.

4.2.2.2 Progression-free survival (PFS)

KM plots of PFS (based on the programmatically derived [RECIST 1.1] outcome) will be presented. Median PFS will be presented (calculated by the KM method).

Summaries of the number of percentage of patients experiencing a PFS event and the type of event (RECIST 1.1, or death) will also be provided. The analysis will be performed on the FAS.

The treatment status at the time of analysis will be summarized. This will include the number (%) of patients who were on treatment at the time of progression, the number (%) of patients

who discontinued study treatment prior to progression, the number (%) of patients who have not progressed and were on treatment or discontinued treatment.

4.2.2.3 Overall survival (OS)

KM plots as well as summaries of the number and percentage of patients who have died, those still in survival follow-up, those lost to follow-up, and those who have withdrawn consent will be provided along with the median OS (calculated using KM method) for each cohort.

4.2.2.4 Duration of response (DoR)

Descriptive data and KM plots of DoR will be presented. The median DoR will also be summarized, calculated by the KM method. Only patients who have a confirmed response will be included in this summary. This analysis will be performed on the evaluable for response analysis set.

Swimmer plots that clearly show the profile of each patient who responds will also be produced. This depicts each duration of response for each patient as a separate bar (horizontally) over time. Each cohort will appear on a separate plot.

Sensitivity analyses will be performed on FAS.

4.2.2.5 Best objective response (BoR)

Overall visit response data will be listed for all patients in the FAS, BoR will be summarized by n (%) for each category (CR, PR, SD, PD, and NE categories section 3.2.5).

4.2.2.6 Disease control rate (DCR)

For DCR, the point estimate and 90% CI (2-sided) will be presented at each timepoint (16 and 24 weeks). If the cohorts are expanded after the initial DRC, the CI will be calculated using an exact Mid-P method, otherwise an exact CI will be presented. This analysis will be performed on the evaluable for response analysis set.

Sensitivity analyses will be performed on FAS.

4.2.2.7 Proportion alive and progression-free at 12 and 18 months

APF12, and its 90% CI will be summarized (using KM methodology) and presented.

APF18 will be summarized as for APF12.

4.2.2.8 Overall survival at 12 and 24 months

OS12, and its 90% CI will be summarized (using KM methodology) and presented. OS24 will be summarized as per OS12.

4.2.2.9 Change in tumor size

The best change in TL tumor size from baseline, (where best change in TL size is the maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction) will be summarized using descriptive statistics for each solid tumor cohort (SARCOMA and STO).

Tumor size will also be presented graphically using waterfall plots for each solid tumor cohort, to present each subject's best percentage change in tumor size as a separate bar, with the bars ordered from the largest increase to the largest decrease. Reference lines at the +20% and – 30% change in tumor size levels will be added to the plots, which correspond with the definitions of progression and 'partial' response respectively. All progressions will be marked with a '●' or designated with patterns or colors for ORR categories. The scale in 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 are marked with '#'.

Additionally, 'spider' plots will be produced for each solid tumor cohort. This depicts each patient's percentage change in TL tumor size as a line over time and progression due to nontarget and/or new lesions will be indicated.

4.2.2.10 Overall RECIST response (Dose-finding)

No formal efficacy analysis will be performed for dose-finding cohorts, however a listing detailing overall RECIST response based on investigators assessment will be presented.

4.2.3 Safety analysis

Data from all cycles of treatment will be combined in the presentation of safety data. Safety summaries will be descriptive only. No formal statistical analyses will be performed on the safety variables.

4.2.3.1 DLT

The number and percentage of subjects with DLT during the dose finding phase will be presented by dose level based on DLT evaluable population.

4.2.3.2 Adverse events

AEs will be summarized separately for the periods defined in Section 3.3.1. AEs occurring pre-treatment or occurring post study treatment, or post-subsequent anti-cancer therapy will be listed, but not summarized separately.

An overall summary table will be produced showing the number and percentage of patients with at least 1 AE in any of the following categories: AEs, SAEs, deaths due to AE, AEs causing discontinuation of study treatment, and AESI/AEPI. The total number of AEs in the different AE categories in terms of AE counts will also be presented (i.e. accounting for multiple occurrences of the same event in a patient).

AEs will be summarized by system organ class (SOC) and PT assigned to the event using MedDRA. For each PT, the number and percentage of patients reporting at least one occurrence will be presented i.e. for 1 patient, multiple occurrences of an AE will only be counted once.

AEs (by SOC and PT) will be summarized by causality and maximum reported CTCAE grade. If a patient reports multiple occurrences of the same AE, maximum reported CTCAE grade will be taken as the highest recorded CTCAE grade for each SOC and PT.

Summary information (the number and percent of patients by SOC and PT, for each phase, will be tabulated for:

- All AEs
- Most common AEs
- All AEs by maximum reported CTCAE grade
- AEs of CTCAE grade 3 or 4
- Most common AEs of CTCAE grade 3 or 4
- Causally related AEs (to any study treatment)
- Causally related AEs of CTCAE grade 3 or 4
- AEs leading to dose interruption / discontinuation
- Dose limiting toxicities (based on DRC decision)

Death summaries and listings:

- All deaths
- AEs with an outcome of death
- AEs with an outcome of death – key patient information
- Listing of deaths

SAE summaries and listings:

- SAEs
- Causally related SAEs
- SAEs – listing of key information

AEs leading to discontinuation summaries and listings:

- AEs leading to discontinuation of any study treatment
- Causally related AEs leading to discontinuation of any study treatment
- AEs leading to discontinuation of study treatment – key patient information

AEs of special or possible interest summaries:

- AESI/AEPI list of PTs [a]
- AESI/AEPI by CTCAE grade
- AESI/AEPI by outcome
- AESI/AEPI with an outcome of no longer present, by time of resolution
- AESI/AEPI leading to dose interruptions or discontinuation

[a] – AESI/AEPI PTs will be obtained from the AZ team (see section 3.3.1)

imAE:

The imAEs will be summarized in the same manner as for the summaries for AESI described above. The additional analyses include but not limited to, time to first onset imAE and resolution of imAE of Grade 3 or 4. See further details in the imAE Charter with respect to derivation rules.

In addition, the following analyses regarding systemic steroid use for imAE will be provided by imAE group. See further details in the imAE Charter with respect to derivation rules associated with duration and time to systemic steroid use for imAE. The duration of steroid use will be produced by imAE group for both all systemic steroids and high dose systemic steroids.

4.2.3.3 Clinical laboratory variables (clinical chemistry, hematology and urinalysis)

All laboratory safety data will be listed for each patient and summarized where appropriate. Clinical chemistry and hematology values outside the reference ranges will be highlighted in the listings. Data summaries will be provided in preferred units.

For continuous laboratory data absolute and change from baseline will be summarized by visit.

Box plots of change from baseline will be presented for haematology and clinical chemistry variables.

Shift tables for laboratory values by worst CTCAE grade will be produced, and for specific parameters separate shift tables indicating hyper- and hypo- directionality of change will be produced. The laboratory parameters for which CTCAE grade shift outputs will be produced are:

- Hematology: Hemoglobin (low and high), Leukocytes (low and high), Lymphocytes (low and high – absolute count), Neutrophils (low – absolute count), Platelets.

- Clinical chemistry: ALT (high), AST (high), ALP (high), Total bilirubin (high), Albumin (low), Magnesium (hypo and hyper), Sodium (hypo and hyper), Potassium (hypo and hyper), Corrected calcium (hypo and hyper), Glucose (low), Creatinine (high).

For urinalysis a shift table comparing baseline to the maximum on-treatment value will be presented.

Hy's law

The following summaries will include the number (%) of patients who have:

- Elevated ALT, AST, and Total bilirubin during the study
  - ALT ≥3× - ≤5×, > 5× - ≤8×, >8× - ≤10×, >10× - ≤20×, and >20× Upper Limit of Normal (ULN) during the study
  - AST ≥3× - ≤5×, >5× - ≤8×, >8× - ≤10×, >10× - ≤20×, and >20× ULN during the study
  - Total bilirubin ≥2× - ≤3×, >3× - ≤5×, >5× ULN during the study
  - ALT or AST ≥3× - ≤5×, >5× - ≤8×, >8× - ≤10×, >10× - ≤20× and >20× ULN during the study
  - ALT or AST ≥3× ULN and Total bilirubin ≥2× ULN during the study (potential Hy's law). Note: Note: The onset date of ALT or AST elevation should be prior to or on the date of total bilirubin elevation.
- Narratives will be provided in the CSR for patients who have ALT ≥3× ULN plus Total bilirubin ≥2× ULN or AST ≥3× ULN plus Total bilirubin ≥2× ULN at any visit.

Liver biochemistry test results over time for patients with elevated ALT (i.e. ≥3× ULN) or AST (i.e. ≥3× ULN), and elevated Total bilirubin (i.e. ≥2× ULN) (at any time) will be plotted. Individual patient data where ALT or AST (i.e. ≥3× ULN) plus Total bilirubin (i.e. ≥2× ULN) are elevated at any time will be listed also.

Plots of ALT and AST versus total bilirubin by cohort will also be produced with reference lines at 3×ULN for ALT, AST, and 2×ULN for total bilirubin. In each plot, total bilirubin will be in the vertical axis.

4.2.3.4 Vital signs

Height, weight, BMI, Pulse rate, respiratory rate, systolic and diastolic blood pressure, peripheral oxygen saturation and temperature will be listed by patient and summarized for the absolute value at each visit.

Box plots for absolute values and change from baseline by visit may be presented for certain vital signs parameters if warranted after data review.

4.2.3.5 Physical examination

Physical examination findings will be reported as adverse events.

4.2.3.6 Electrocardiogram

ECG details will be listed individually by patient.

The overall ECG evaluations, classified as normal, borderline, abnormal not clinically significant and abnormal clinically significant will be summarized using shift tables summarising the change from baseline to last on-treatment observation, a shift table of baseline evaluation to worst evaluation on-treatment will be produced if sufficient assessments are recorded to allow this.

4.2.3.7 Exposure and compliance

Study treatment exposure data will be listed for each patient.

The total and actual duration of durvalumab monotherapy and combination treatment exposure will be summarized.

Number of patients with at least one delay (and reason for delay), the number of patients with no delay, the number of dosing occasions and the RDI will also be summarized.

4.2.3.8 Other safety data

Other safety data will be listed and/or summarized:

- Tanner stage
- Performance status (Karnofsky or Lansky – the scale appropriate to the patient at the time of screening will be used through the study)

4.2.4 Pharmacokinetic and Immunogenicity analysis

4.2.4.1 Pharmacokinetics - PK parameters

PK data will be summarized and analyzed based on the PK analysis set.

PK concentration data will be listed for each patient and each dosing day, and a summary will be provided for all evaluable patients at each planned time point.

Serum concentrations at each time point will be summarized using the following summary statistics:

- The geometric mean (gmean), calculated as exp [μ], where μ is the mean of the data on a logarithmic scale.
- CV, calculated as 100 √ [exp(s^2)-1], where s is the STD of the data on a log scale
- Arithmetic mean calculated using untransformed data
- STD calculated using untransformed data
- Minimum
- Median
- Maximum
- Number of observations

The following summary statistics will be presented for C_{max} (μg/mL), C_{min} (μg/mL), time to Cmax, and AUC after the first dose, if the data allow

- Gmean, calculated as exp[μ], where μ is the mean of the data on a logarithmic scale
- CV, calculated as 100 √ [exp(s2)-1], where s is the STD of the data on a log scale
- Arithmetic mean calculated using untransformed data
- STD calculated using untransformed data
- Minimum
- Median
- Maximum
- Number of observations

The following conventions will be applied to values below the lower limit of quantitation (LLOQ) for summarising PK data:

- If, at a given time point, 50% or less of the blood concentrations are non-quantifiable (NQ), the gmean, CV, geometric standard deviation (geoSD), arithmetic mean and STD are calculated by substituting the limit of quantification (LOQ) for values which are NQ.
- If more than 50%, but not all, of the concentrations are NQ, the gmean, CV, geoSD, arithmetic mean and STD are reported as not calculable (NC)
- If all the concentrations are NQ, the gmean and arithmetic mean are reported as NQ and the CV, geoSD and STD as NC
- The number of values above LLOQ are reported for each time point along with the total number of collected values

Samples below the lower limit of quantification will be treated as zero in predose samples and as missing subsequently in the analyses.

The PK data for after a single-dose and separately will also be displayed graphically. Displays will include serum concentration patient profiles (on the linear and log-scale) versus time and gmean concentration (±STD) versus time, stratified by dose.

4.2.4.2 Pharmacokinetic relationships

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated.

A population PK model will be developed using a non-linear mixed-effects modeling approach. The impact of physiologically-relevant patient characteristics (covariates) and disease on PK will be evaluated. The relationship between the PK exposure and the effect on safety and efficacy endpoints will be evaluated if the data allow. The results of such an analysis will be reported in a separate report. The PK, demographic, safety, and efficacy data collected in this study may also be combined with similar data from other studies and explored using population PK and/or PK-pharmacodynamic methods.

4.2.4.3 Immunogenicity - ADAs

A summary will be provided of number and percentage of patients who develop detectable ADAs. The immunogenicity titer and neutralizing ADA data will be listed for samples confirmed positive for the presence of anti-durvalumab and anti-tremelimumab antibodies.

The summaries will consider durvalumab and tremelimumab evaluability of ADA samples separately. For instance, summaries considering ADA+ to durvalumab and tremelimumab will need to first consider if a patient was both evaluable to durvalumab and evaluable to tremelimumab. Subsequently, if summarizing for ADA+ to durvalumab only, we only consider if a patient was evaluable to durvalumab.

The effect of immunogenicity as well as the effect of its neutralizing properties on PK, efficacy, and safety will be evaluated, if the data allow. A detailed plan will be written by the AstraZeneca Clinical Pharmacology group or designee.

4.2.5 Demographic, patient disposition and baseline characteristics data

The following will be summarized for all patients in the FAS.

Disposition of patients

A summary will be provided for the number of patients enrolled, who received treatment, who completed the study, and who terminated prematurely. Reasons for premature termination will be summarized.

The number of patients in each analysis set, by country and centre will also be summarized.

Demographic data

Descriptive statistics will be presented for age, height, weight and BMI. In addition, frequencies and percentage of patients will be tabulated for the categorical variables age group (0-<6, 6-<12, 12-<18), sex, race and ethnic group.

Other baseline data

Other demographic and baseline data will be listed and/or summarized:

- Medical and surgical history
- Nicotine use categorized (never, current, former)
- Extent of disease at baseline
- Pregnancy information
- Important protocol deviations (both overall, and separate summaries for those related and unrelated to the Coronavirus Disease 2019 (COVID-19) pandemic)
- COVID-19 study disruptions (number and percentage of patients with the following: visits impacted, study drug impacted and concomitant medications not started)
- Inclusion in analysis populations.

Any COVID-19 related summaries will only be produced if there are a minimum of 5 patients with data to present.

4.2.6 Concomitant and other treatment

All medications will be coded using the latest AstraZeneca Drug Dictionary at the time of database lock.

Prior medication is defined as medication with a stop date before the date of the first dose of study treatment

A medication will be regarded as concomitant if the start date is on or after the date of first dose, or if it started prior to the date of first dose and was ongoing after the date of first dose. Separate tables will be presented according to whether the medication is allowed or disallowed. Disallowed medications are detailed in section 6.4 of the CSP and will identified by the physician throughout the study and merged into the reporting database by IQVIA programming.

Start and stop dates for all concomitant medications are collected on the eCRF. However, in case of missing or partial information in these dates, the following rules will be used:
If start date is missing or partial:

- if month is missing, use January (01)
- if day is missing, use the 1st (01)
- if year is missing, use year of the consent date
- if entire date is missing, use consent date

If stop date is missing, partial or “continuing:”

- if month is missing, use December (12)
- if day is missing, use the last day of the month under consideration
- if year or the entire date is missing or if “continuing”, set to missing.

A patient is only counted once if receiving the medication more than once.

4.2.7 Biomarker data

4.2.7.1 Antibody titer measurements before and after routine immunization

Vaccine antibody titer data will be listed by patient.

4.2.7.2 CCI [redacted]

[redacted] will be listed by patient.

CCI [redacted]

[redacted] will be reported in the CSR. Summaries and analyses for exploratory biomarkers will be documented in a separate analysis plan and will be reported outside the CSR in a separate report.

Based on availability of data, additional exploratory biomarker analyses may be performed.

5 INTERIM ANALYSES

No formal interim analysis for efficacy is planned for this study. However, in dose-expansion phase, ORR will be evaluated on an ongoing basis for the first 11 evaluable patients (SARCOMA cohort) to complete Stage 1 of the Simon 2-Stage optimal design (see Section 4.1.4 of the CSP). The methodology described in this SAP for the derivation of visit responses and efficacy variables will also be used for the DRC analyses. Further information around the analysis and outputs required to facilitate this will be described in the DRC charter.

6 CHANGES OF ANALYSIS FROM PROTOCOL

As there is no current plan to incorporate irRECIST, this has not been referenced in the SAP.

The definition of the evaluable for response analysis set has been updated to align with the definition in the DRC charter.

An additional sensitivity analysis of ORR using unconfirmed responses has been added.

Clarification has been added that the mid-P methodology is applicable only if the cohorts are expanded.

Visit windowing of safety domains will be considered up to end of treatment, rather than up to progression as mentioned in the protocol.

An additional efficacy listing detailing dose-finding cohort overall RECIST response per investigator assessment has been added.

7 REFERENCES

Porcher and Desseaux 2012

Porcher R, Desseaux K. What inference for two-stage phase II trials? BMC Med Res Methodol. 2012;12:117. doi: 10.1186/1471-2288-12-117

8 APPENDIX (NOT APPLICABLE)

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: d419ec00001-sap-ed-4		
Document Title:	Statistical Analysis Plan Edition 4	
Document ID:	CCI [redacted]	
Version Label:	2.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
20-Apr-2023 09:54 UTC	PPD [redacted]	Author Approval
19-Apr-2023 14:35 UTC	PPD [redacted]	Author Approval
19-Apr-2023 13:59 UTC	PPD [redacted] IQVIA	Author Approval

CCI [redacted]