SUPPLEMENTARY APPENDIX

This supplement contains the following items.

1. The original and final versions of the Clinical Study Protocol (CSP), plus a summary of the changes between each version.

2. The original and final versions of the Statistical Analysis Plan (SAP), plus a summary of the changes between each version.

Clinical Study Protocol

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

Sponsor: AstraZeneca AB, 151 85 Södertälje, Sweden

EudraCT Number: 2015-001279-39

The following Amendment(s) and Administrative Changes have been made to this protocol since the date of preparation:

Amendment No.	Date of Amendment	Local Amendment No.	Date of Local Amendment

Administrative Change No.	Date of Administrative Change	Local Administrative Change No.	Date of Local Administrative Change

This clinical study protocol has been subject to a peer review according to AstraZeneca standard procedures. The clinical study protocol is publicly registered and the results are disclosed and/or published according to the AstraZeneca Global Policy on Bioethics and in compliance with prevailing laws and regulations.

PROTOCOL SYNOPSIS

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

International Coordinating Investigators

Study site(s) and number of patients planned

The study will enroll approximately 810 patients to identify approximately 675 patients who will be randomized to receive MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or platinum-based Standard of Care (SoC) therapy (225 in each arm, including at a minimum 438 patients with programmed cell death ligand 1 [PD-L1]–negative non-small-cell lung cancer [NSCLC], defined as PD-L1 expression <25%).

Study period	Phase of development
Estimated date of first patient enrolled	III
Estimated date of last patient completed	III

Study design

This is a randomized, open-label, multi-center, global, Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) wild-type locally advanced or metastatic NSCLC.

Patients will provide a tumor tissue sample at screening to determine PD-L1 expression status (defined by an immunohistochemistry assay developed by in which ≥25% PD-L1 membrane–expression in tumoral tissue is considered positive and <25% is considered negative; referred to hereafter as patients with PD-L1-positive or -negative tumors, respectively). Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (as described above) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy.

Tumor assessments will be performed every 6 weeks for the first 48 weeks and then every 8 weeks until confirmed disease progression, with categorization of objective tumor response by Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1).

Objectives

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1-positive and/or PD-L1–negative NSCLC.

Primary Objective:	**Outcome Measure:**
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with NSCLC	PFS using Investigator assessments according to RECIST 1.1

Secondary Objectives:	**Outcome Measures:**
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, ORR, DoR, APF12, PFS2, and OS	PFS in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1 ORR, DoR, and APF12 using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, ORR, PFS2, and OS	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS in patients with PD-L1–positive NSCLC using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice OS
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS and ORR	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS and ORR in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1
To assess disease-related symptoms and HRQoL in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC QLQ-C30 v3 and the LC13 module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.
To assess the PK of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)

Secondary Objectives:	Outcome Measures:
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of ADAs for MEDI4736 and tremelimumab
To explore irRECIST as an assessment methodology for clinical benefit of MEDI4736 + tremelimumab compared to SoC with assessment by BICR	PFS and ORR using BICR assessment according to irRECIST

Safety Objective:	Outcome Measures:
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of locally advanced or metastatic NSCLC patients	AEs, physical examinations, laboratory findings, and vital signs

Target patient population

Adult patients (age ≥18 years) with locally advanced or metastatic (Stage IV) histologically or cytologically documented EGFR and ALK wild-type NSCLC who are treatment naive.

Duration of treatment

Until specific treatment discontinuation criteria are met, treatment will continue for a 12-month period for the MEDI4736 + tremelimumab and MEDI4736 monotherapy groups and with no maximum treatment duration for the SoC group.

Patients in the MEDI4736 + tremelimumab or MEDI4736 monotherapy groups who develop PD after completing 12 months of therapy may restart their assigned treatment with the same treatment guidelines followed previously. Patients in the MEDI4736 + tremelimumab group may also restart treatment if they complete the 4 dosing cycles with MEDI4736 + tremelimumab (with clinical benefit per Investigator judgement) but subsequently have confirmed PD during treatment with MEDI4736 alone.

Patients who have discontinued study treatment will enter follow-up.

Investigational product, dosage, and mode of administration

MEDI4736 + tremelimumab combination therapy

- MEDI4736 20 mg/kg via IV infusion q4w, starting on Week 0, for up to a total of 4 doses/cycles, and then continue MEDI4736 20 mg/kg via IV infusion q4w, starting on Week 16, for up to a total of 8 months (9 doses)
- Tremelimumab 1 mg/kg via IV infusion q4w, starting on Week 0, for up to 4 doses/cycles

MEDI4736 monotherapy

- MEDI4736 20 mg/kg via IV infusion q4w, starting on Week 0, for up to a total 12 months (13 doses)

Standard of Care therapy

Patients randomized to SoC therapy will receive 1 of the following:

- Paclitaxel + carboplatin: Paclitaxel 200 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + cisplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + cisplatin 75 or 80 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + carboplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + carboplatin area under the curve (AUC) 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Pemetrexed + cisplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and cisplatin 75 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.
- Pemetrexed + carboplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.

Statistical methods

The primary objective of this study is to assess the efficacy of MEDI4736 + tremelimumab treatment compared with SoC in terms of progression-free survival (PFS) in patients with EGFR and ALK wild-type locally advanced or metastatic NSCLC. PFS (per RECIST 1.1, using site Investigator tumor assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression. A sensitivity analysis of PFS will be performed based on data from Blinded Independent Central Review (BICR) assessments for a random sample of patients. If bias cannot be excluded based upon the sample BICR assessment, then an independent evaluation of all radiographic images may be required for the assessment of the primary PFS endpoint.

Secondary efficacy variables include PFS for MEDI4736 + tremelimumab versus SoC in patients with PD-L1-negative tumors, PFS for MEDI4736 monotherapy versus SoC, and PFS for MEDI4736 monotherapy versus SoC in patients with PD-L1-positive tumors, PFS for MEDI4736 + tremelimumab versus MEDI4736 monotherapy, PFS for MEDI4736 + tremelimumab versus MEDI4736 monotherapy in patients with PD-L1-negative tumors, as well as objective response rate (ORR), duration of response (DoR), proportion of patients alive and progression free at 12 months from randomization (APF12), time from randomization to second progression (PFS2), and overall survival (OS). All tumor-assessment-related endpoints as assessed by site Investigator.

Efficacy data will be summarized and analyzed on an intent-to-treat (ITT) basis and the treatment arms will be compared on the basis of randomized treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the ITT population.

Approximately 675 patients will be randomized 1:1:1 to MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC. The randomization will be stratified based on PD-L1 expression tumor status (PD-L1-positive and -negative) and histology (squamous versus non-squamous).

A hypothesis of improved PFS will be tested when

- (approximately) 338 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups (75% maturity) AND
- (approximately) 223 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1-negative tumors (76% maturity)

Comparison of MEDI4736 monotherapy versus SoC will also be performed in all randomized patients as well as in patients with PD-L1-positive tumors at this time.

MEDI4736 + tremelimumab versus SoC (primary endpoint)

If PFS at 12 months was 33.5% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.59, the trial will have >90% power to demonstrate statistical significance at the 5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.81. With a 15-month recruitment period and a minimum follow up period of 6.5 months assumed, it is anticipated that this analysis will be performed 21.5 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PD-L1–negative population)

If the boundary is crossed for the primary PFS analysis, then PFS in the PD-L1–negative population of MEDI4736 + tremelimumab versus SoC will be tested at the 5% alpha level (in line with the hierarchical testing strategy). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and SoC treatment groups and a true PFS average HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 90% power to demonstrate statistical significance at the 5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.77.

MEDI4736 + tremelimumab versus MEDI4736 monotherapy (PD-L1–negative population)

If the boundary is crossed for the primary PFS analysis and for PFS of MEDI4736 + tremelimumab versus SoC in the PD-L1–negative population, then a statistical analysis of MEDI4736 + tremelimumab versus MEDI4736 monotherapy in the PD-L1–negative population will be performed (in line with the hierarchical testing strategy). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups and a true average PFS HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 86% power to demonstrate statistical significance at the 2.5% significance level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.74.

PFS, based on the programmatically derived PFS from site Investigator assessments, will be analyzed using a stratified log-rank test (stratified for PD-L1 expression tumor status [PD-L1 positive, PD-L1 negative] [for analysis in ITT population only] and histology [squamous, non-squamous]). The effect of treatment will be estimated by the HR together with corresponding 95% CI and p-value.

Safety data will be summarized descriptively and will not be formally analyzed.

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

The following abbreviations and special terms are used in this study Clinical Study Protocol.

Abbreviation or special term	Explanation
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse events of special interest
ALK	Anaplastic lymphoma kinase
ALT	Alanine aminotransferase
APF12	Proportion of patients alive and progression free at 12 months from randomization
AST	Aspartate aminotransferase
AUC	Area under the curve
$AUC_{0\text{-}28day}$	Area under the plasma drug concentration-time curve from time zero to Day 28 post-dose
AUC_{ss}	Area under the curve at steady state
BICR	Blinded Independent Central Review
BoR	Best objective response
BP	Blood pressure
CD	Cluster of differentiation
CI	Confidence interval
C_{max}	Maximum plasma concentration
$C_{max,ss}$	Maximum plasma concentration at steady state
CR	Complete response
CSA	Clinical study agreement
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Event
CTLA-4	Cytotoxic T–lymphocyte-associated antigen 4
C_{trough}	Trough plasma concentration
$C_{trough,ss}$	Trough plasma concentration at steady state
CXCL	Chemokine (C-X-C motif) ligand
DCR	Disease control rate

Abbreviation or special term	Explanation
DLT	Dose-limiting toxicity
DoR	Duration of response
DSR	Deep sustained response
EC	Ethics Committee, synonymous to Institutional Review Board (IRB) and Independent Ethics Committee (IEC)
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDoR	Expected duration of response
EGFR	Epidermal growth factor receptor
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQol 5-Dimension, 5-Level health state utility index
ESMO	European Society for Medical Oncology
EU	European Union
FAS	Full Analysis Set
FDA	Food and Drug Administration
FU	Fluorouracil
GCP	Good Clinical Practice
GI	Gastrointestinal
GMP	Good Manufacturing Practice
hCG	Human chorionic gonadotropin
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
HR	Hazard ratio
HRQoL	Health-related quality of life
IB	Investigator's Brochure
ICF	Informed consent form
ICH	International Conference on Harmonisation
IDMC	Independent Data Monitoring Committee
IFN	Interferon
IgG	Immunoglobulin G
IHC	Immunohistochemistry

Abbreviation or special term	Explanation
IL	Interleukin
IMT	Immunomodulatory therapy
IP	Investigational product
irAE	Immune-related adverse event
irRECIST	Immune-related response criteria
ITT	Intent-to-Treat
IV	Intravenous
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
mAb	Monoclonal antibody
MDSC	Myeloid-derived suppressor cell
MedDRA	Medical Dictionary for Regulatory Activities
MHLW	Minister of Health, Labor, and Welfare
miRNA	Micro-ribonucleic acid
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NE	Not evaluable
NSCLC	Non–small-cell lung cancer
OAE	Other significant adverse event
ORR	Objective response rate
OS	Overall survival
PBMC	Peripheral blood mononuclear cell
PD	Progressive disease
PD-1	Programmed cell death 1
PD-L1	Programmed cell death ligand 1
PD-L2	Programmed cell death ligand 2
PDx	Pharmacodynamic(s)
PFS	Progression-free survival
PFS2	Progression-free survival after subsequent anticancer therapy
PGIC	Patient’s Global Impression of Change
PGx	Pharmacogenetic research
PK	Pharmacokinetic(s)

Abbreviation or special term	Explanation
PR	Partial response
PRO	Patient-reported outcome
PRO-CTCAE	Patient-reported outcomes version of the CTCAE
q2w	Every 2 weeks
q3w	Every 3 weeks
q4w	Every 4 weeks
q8w	Every 8 weeks
q12w	Every 12 weeks
QLQ-C30 v3	30-item core quality of life questionnaire, version 3
QLQ-LC13	13-item lung cancer quality of life questionnaire
QoL	Quality of life
QTcF	QT interval corrected for heart rate using Fridericia's formula
RECIST 1.1	Response Evaluation Criteria in Solid Tumors, version 1.1
RR	Response rate
RT-QPCR	Reverse transcription quantitative polymerase chain reaction
SAE	Serious adverse event
SAP	Statistical analysis plan
SD	Stable disease
SoC	Standard of Care
sPD-L1	Soluble programmed cell death ligand 1
T_3	Triiodothyronine
T_4	Thyroxine
TEAE	Treatment-emergent adverse event
TKI	Tyrosine-kinase inhibitor
TSH	Thyroid-stimulating hormone
ULN	Upper limit of normal
US	United States
WBDC	Web-Based Data Capture
WHO	World Health Organization

1. INTRODUCTION

1.1 Background and rationale for conducting this study

Lung cancer has been the most common cancer in the world for several decades, with an estimated 1.8 million new cases in 2012 (12.9% of all new cancers), and was also the most common cause of death from cancer in 2012, with 1.59 million deaths (19.4% of the total cancer deaths; GLOBOCAN 2012). Non-small-cell lung cancer (NSCLC) represents approximately 80% to 85% of all lung cancers. Unfortunately, at the time of diagnosis, approximately 70% of patients with NSCLC already have advanced or metastatic disease not amenable to surgical resection. Furthermore, a significant percentage of patients with early stage NSCLC who have undergone surgery subsequently develop distant recurrence and die as a result of their lung cancer (Pisters and LeChevalier 2005).

Despite advances in the diagnosis, imaging, staging, and treatment of NSCLC, the estimated overall 5-year survival for patients in Europe and the United States (US) continues to be low (11% and 17%, respectively; D'Addario et al 2010, Howlander et al 2014). Patients presenting with advanced NSCLC have a median overall survival (OS) of 10 to 12 months (Bonomi 2010). Patients without a targetable mutation (ie, epidermal growth factor receptor [EGFR] or anaplastic lymphoma kinase [ALK] mutation) demonstrate responses to systemic treatment of approximately 20% to 30% and progression-free survival (PFS) of 4 to 5 months (Scagliotti et al 2008, Schiller et al 2002). The durations of responses (DoR) are also limited, and toxicities can be a major limiting factor. The 1-year survival rate is 30% to 40% for patients with a good performance status. Maintenance therapy, with either continuation or switch, has also been recommended for certain histologic subtypes of NSCLC; for example, maintenance with pemetrexed has been shown to improve OS and PFS, particularly in non-squamous histologies (Ciuleanu et al 2009, Paz-Ares et al 2013).

Common first-line treatment regimens for advanced NSCLC in major global markets are typically platinum-based doublets and include carboplatin and paclitaxel, carboplatin and gemcitabine (squamous only), carboplatin and pemetrexed (non-squamous only), cisplatin and gemcitabine (squamous only), and cisplatin and pemetrexed (non-squamous only). Platinum-based doublet chemotherapy regimens vary to some extent with regard to convenience, associated toxicities, and cost, with the selection of a specific regimen often dictated by local practice and individualized on a case-by-case basis.

1.1.1 Immunotherapies

It is increasingly understood that cancers are recognized by the immune system, and, under some circumstances, the immune system may control or even eliminate tumors (Dunn et al 2004). Studies in mouse models of transplantable tumors have demonstrated that manipulation of co-stimulatory or co-inhibitory signals can amplify T-cell responses against tumors (Peggs et al 2009). This amplification may be accomplished by blocking co-inhibitory molecules, such as cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) or programmed

cell death 1 (PD-1), from binding with their ligands, B7 or B7-H1 (programmed cell death ligand 1 [PD-L1]).

1.1.2 MEDI4736

MEDI4736 is a human monoclonal antibody (mAb) of the immunoglobulin G (IgG) 1 kappa subclass that inhibits binding of PD-L1 and is being developed by AstraZeneca/MedImmune for use in the treatment of cancer. (MedImmune is a wholly owned subsidiary of AstraZeneca; AstraZeneca/MedImmune will be referred to as AstraZeneca throughout this document.) As MEDI4736 is an engineered mAb, it does not induce antibody-dependent cellular cytotoxicity or complement-dependent cytotoxicity. The proposed mechanism of action for MEDI4736 is interference of the interaction of PD-L1.

PD-L1 is expressed in a broad range of cancers with a high frequency, up to 88% in some types of cancers. In a number of these cancers, including lung, the expression of PD-L1 is associated with reduced survival and an unfavorable prognosis. In lung cancer, only 12% of patients with tumors expressing PD-L1 survived for more than 3 years, compared with 20% of patients with tumors lacking PD-L1 (Mu et al 2011). Based on these findings, an anti-PD-L1 antibody could be used therapeutically to enhance anti-tumor immune responses in patients with cancer. Results of several non-clinical studies using mouse tumor models support this hypothesis, where antibodies directed against PD-L1 or its receptor PD-1 showed anti-tumor activity (Hirano et al 2005, Iwai et al 2002, Okudaira et al 2009, Zhang et al 2008).

MEDI4736 has been given to humans as part of ongoing studies as a single drug or in combination with other drugs. As of 14 July 2014, 509 patients have been enrolled and treated in 10 ongoing clinical studies of MEDI4736 (5 employing MEDI4736 as monotherapy and 5 as combination therapy). Details on the safety profile of MEDI4736 monotherapy are summarized in Section 1.3.2.1. Refer to the current MEDI4736 Investigator's Brochure for a complete summary of non-clinical and clinical information; see Section 6.7 for guidance on management of MEDI4736-related toxicities.

MEDI4736 monotherapy exhibits nonlinear (dose-dependent) pharmacokinetics (PK) approaching linearity with a ≥3-mg/kg dose, likely due to saturable target-mediated clearance, and has a half-life of approximately 21 days. Of the 220 patients who received MEDI4736 monotherapy for whom PK/anti-drug antibody (ADA) data were available from Study CD-ON-MEDI4736-1108 (referred to hereafter as Study 1108) as of 14 July 2014, 5 patients (1 patient each in the 0.1- and 3-mg/kg cohorts and 3 patients in the 10-mg/kg cohort) were detected ADA-positive, with an impact on PK/pharmacodynamics in 1 patient in the 3-mg/kg cohort.

1.1.3 Tremelimumab

Tremelimumab, a CTLA-4 mAb of the IgG 2 kappa isotype, is an immunomodulatory therapy (IMT) that is being developed by AstraZeneca for use in the treatment of cancer. Tremelimumab is a human IgG2 mAb directed against CTLA-4.

Binding of CTLA-4 to its target ligands (B7-1 and B7-2) provides a negative regulatory signal, which limits T-cell activation. Anti-CTLA-4 inhibitors antagonize the binding of CTLA-4 to B7 ligands and enhance human T-cell activation as demonstrated by increased cytokine (interleukin [IL]-2 and interferon [IFN] gamma) production in vitro in whole blood or peripheral blood mononuclear cell (PBMC) cultures (Tarhini and Kirkwood 2008). In addition, blockade of CTLA-4 binding to B7 by anti-CTLA-4 antibodies results in markedly enhanced T-cell activation and anti-tumor activity in animal models, including killing of established murine solid tumors and induction of protective anti-tumor immunity. (Refer to the tremelimumab IB, Edition 4.0, for more information.) Therefore, it is expected that treatment with an anti-CTLA-4 antibody, such as tremelimumab, will lead to increased activation of the human immune system, increasing anti-tumor activity in patients with solid tumors.

An extensive program of non-clinical and clinical studies has been conducted for tremelimumab both as monotherapy and combination therapy with conventional anticancer agents to support various cancer indications using different dose schedules. As of the data cut-off date of 12 November 2014 (for all studies except D4190C00006, which has a cut-off date of 4 December 2014), 973 patients have received tremelimumab monotherapy (not including 497 patients who have been treated in the blinded Phase 2b study, D4880C00003) and 208 patients have received tremelimumab in combination with other agents. Details on the safety profile of tremelimumab monotherapy are summarized in Section 1.3.2.2. Refer to the current tremelimumab IB for a complete summary of non-clinical and clinical information; see Section 6.7 for guidance on management of tremelimumab-related toxicities.

Tremelimumab exhibited a biphasic PK profile with a long terminal phase half-life of 22 days. Overall, a low incidence of ADAs (<6%) was observed for treatment with tremelimumab.

1.1.4 MEDI4736 in combination with tremelimumab

Targeting both PD-1 and CTLA-4 pathways may have additive or synergistic activity (Pardoll 2012) because the mechanisms of action of CTLA-4 and PD-1 are non-redundant; therefore, AstraZeneca is also investigating the use of MEDI4736 + tremelimumab combination therapy for the treatment of cancer.

Study D4190C00006 is a Phase Ib dose-escalation study to establish safety, PK/pharmacodynamics, and preliminary anti-tumor activity of MEDI4736 + tremelimumab combination therapy in patients with advanced NSCLC. The dosing schedule utilized is MEDI4736 every 2 or 4 weeks (q2w; q4w) up to Week 50 and 48 (12 months), combined with tremelimumab q4w up to Week 24 for 7 doses then every 12 weeks (q12w) for 2 additional doses for up to 12 months. The study is ongoing and continues to accrue.

As of 27 January 2015, a total of 74 patients have been treated in the study, including 58 patients on the q4w dosing schedule and 16 patients on the q2w dosing schedule. Patients have received between 1 and 13 doses of MEDI4736 and between 1 and 9 doses of tremelimumab. Details on the safety profile of MEDI4736 + tremelimumab combination therapy are summarized in Sections 1.2.1 and 1.3.2.3. Refer to the MEDI4736 IB

(Edition 7.0) and tremelimumab IB (Edition 4.0) for a complete summary of non-clinical and clinical information; see Section 6.7 for guidance on management of MEDI4736 + tremelimumab-related toxicities.

As of 27 January 2015 in Study D4190C00006, an approximately dose-proportional increase in PK exposure (maximum plasma concentration [C_{max}] and area under the plasma drug concentration-time curve from time zero to Day 28 post-dose [$AUC_{0\text{-}28day}$]) of both MEDI4736 and tremelimumab was observed over the dose range of 3 to 20 mg/kg MEDI4736 q4w or q2w and 1 to 10 mg/kg tremelimumab q4w. Four of 60 patients with ADA data available were ADA positive for either anti-MEDI4736 or anti-tremelimumab antibodies post-treatment; MEDI4736 PK was impacted in only 2 of these 4 patients. Complete soluble programmed cell death ligand 1 (sPD-L1) suppression (surrogate for PD-L1 targeting) was observed in almost all patients over the dose range of 3 to 20 mg/kg of MEDI4736 q4w or q2w.

1.1.5 Rationale for conducting this study

Current therapies for advanced NSCLC have poor outcomes (low 5-year survival of 17% for the US), with responses to systemic chemotherapy in the first-line setting of approximately 20% to 30%, and a median OS of approximately 10 to 12 months (Bonomi 2010, D'Addario et al 2010, Scagliotti et al 2008, Schiller et al 2002). Responses are also limited in duration. Systemic chemotherapy is associated with significant side effects, including neutropenia, nausea, vomiting and dehydration, and alopecia (Sandler et al 2006, Scagliotti et al 2008). There is still a significant unmet medical need for additional treatment options for use in this patient population as the 1-year survival rate is 30% to 40%.

As an antibody that blocks the interaction between PD-L1 and its receptors, MEDI4736 may relieve PD-L1-dependent immunosuppressive effects and, therefore, enhance the cytotoxic activity of anti-tumor T-cells. This hypothesis is supported by emerging clinical data from other mAbs targeting the PD-L1/PD-1 pathway, which provide early evidence of clinical activity and a manageable safety profile (Brahmer et al 2012, Topalian et al 2012). Responses have been observed in patients with PD-L1-positive tumors and patients with PD-L1-negative tumors. In addition, MEDI4736 monotherapy has shown durable responses in NSCLC in Study 1108 (see Section 1.3.1.1).

The rationale for combining MEDI4736 and tremelimumab is that the mechanisms of CTLA-4 and PD-1 are non-redundant, suggesting that targeting both pathways may have additive or synergistic activity (Pardoll 2012). In fact, combining immunotherapy agents has been shown to result in improved response rates relative to monotherapy. For example, the concurrent administration of nivolumab and ipilimumab to patients with advanced melanoma induced higher objective response rates (ORRs) than those obtained with single-agent therapy. Importantly, responses appeared to be deep and durable (Wolchok et al 2013). Similar results have been observed in an ongoing study of MEDI4736 + tremelimumab in NSCLC (Antonia et al 2014a), with further updated details presented in this clinical study protocol.

Based on the preliminary clinical efficacy and safety data observed in patients with NSCLC in Study D4190C00006 (with MEDI4736 + tremelimumab) and in Study 1108 (with MEDI4736 monotherapy), AstraZeneca plans to determine the activity of MEDI4736 in combination with tremelimumab and MEDI4736 monotherapy as first-line treatment in patients with NSCLC. The preliminary efficacy, safety, and tolerability data of the MEDI4736 + tremelimumab combination in Study D4190C00006 and MEDI4736 monotherapy in Study 1108 support the development of these treatments in NSCLC. The primary endpoint of this Phase III study is to determine the activity of MEDI4736 + tremelimumab combination therapy compared to Standard of Care (SoC) in patients with EGFR and ALK wild-type NSCLC when used as first-line treatment, and secondary endpoints will include comparison of the activity of the monotherapy regimen to SoC. In addition to quantifying the clinical effects of MEDI4736 in combination with tremelimumab, this will also allow comparison between the relative efficacy and tolerability of MEDI4736 + tremelimumab and MEDI4736 alone.

1.2 Rationale for study design, doses, and control groups

This study will utilize an open-label design due to the different treatment administration schedules and treatment durations.

1.2.1 MEDI4736 + tremelimumab combination therapy dose rationale

The MEDI4736 + tremelimumab doses and regimen selected for this study are based on the goal of selecting an optimal combination dose of MEDI4736 and tremelimumab that would yield sustained target suppression (sPD-L1), demonstrate promising efficacy, and have an acceptable safety profile.

As of 27 January 2015, a total of 74 patients with advanced NSCLC have been treated in Study D4190C00006. The 74 patients have received between 1 and 9 doses of tremelimumab and between 1 and 13 doses of MEDI4736. Various dose combinations were explored, with doses of tremelimumab ranging from 1 to 10 mg/kg and doses of MEDI4736 ranging from 3 to 20 mg/kg. Fifty-eight of these patients were in the q4w dosing schedule and 16 patients were in the q2w dosing schedule, with the goal of identifying the dose combination that best optimizes the risk:benefit profile in an acceptable range of PK and pharmacodynamic values.

In order to reduce the dosing frequency of MEDI4736 to align with the q4w dosing of tremelimumab, while ensuring an acceptable PK/pharmacodynamic, safety, and efficacy profile, cohorts were narrowed to 15 and 20 mg/kg MEDI4736 q4w. PK simulations from the MEDI4736 monotherapy data indicated that a similar area under the plasma drug concentration-time curve at steady state (AUC_{ss}; 4 weeks) was expected following both 10 mg/kg q2w and 20 mg/kg q4w MEDI4736. The observed MEDI4736 PK data from the D4190C00006 study were well in line with the predicted monotherapy PK data developed preclinically. This demonstrates similar exposure of MEDI4736 20 mg/kg q4w and 10 mg/kg q2w, with no alterations in PK when MEDI4736 and tremelimumab (doses ranging from 1 to 3 mg/kg) are dosed together. While the median C_{max} at steady state ($C_{max,ss}$) is expected to be higher with 20 mg/kg q4w (approximately 1.5 fold) and median trough concentration at steady state ($C_{trough,ss}$) is expected to be higher with 10 mg/kg q2w (approximately 1.25 fold), this is

not expected to impact the overall safety and efficacy profile, based on existing preclinical and clinical data.

Monotonic increases in pharmacodynamic activity were observed with increasing doses of tremelimumab relative to the activity observed in patients treated with MEDI4736 monotherapy. There was evidence of augmented pharmacodynamic activity relative to MEDI4736 monotherapy with combination doses containing 1 mg/kg tremelimumab, inclusive of both the 15 and 20 mg/kg MEDI4736 plus 1 mg/kg tremelimumab combinations.

Patients treated with doses of tremelimumab above 1 mg/kg had a higher rate of adverse events (AEs), including discontinuations due to AEs, serious AEs (SAEs), and severe AEs. Between the 10 mg/kg MEDI4736 + 1 mg/kg tremelimumab and 10 mg/kg MEDI4736 + 3 mg/kg tremelimumab cohorts treated at the q2w schedule, the number of patients reporting any AE, Grade 3 AEs, SAEs, and treatment-related AEs was higher in the 10 mg/kg MEDI4736 + 3 mg/kg tremelimumab cohort than the 10 mg/kg MEDI4736 + 1 mg/kg tremelimumab cohort. A similar pattern was noted in the q4w regimens, suggesting that, as the dose of tremelimumab increased above 1 mg/kg, a higher rate of treatment-related events may be anticipated. Further, the SAEs frequently attributed to immunotherapy, pneumonitis and colitis, were more commonly seen in cohorts using either 3 mg/kg or 10 mg/kg of tremelimumab compared to the 1-mg/kg dose cohorts. Together, these data suggest that a combination using a tremelimumab dose of 1 mg/kg appeared to minimize the rate of toxicity when combined with MEDI4736. As a result, all combination doses utilizing either the 3 or 10 mg/kg doses of tremelimumab were eliminated in the final dose selection.

In contrast, cohorts assessing higher doses of MEDI4736 with a constant dose of tremelimumab did not show an increase in the rate of AEs. The data suggested that increasing doses of MEDI4736 may not impact the safety of the combination as much as the tremelimumab dose. Further, safety data between the 10-mg/kg and 20-mg/kg cohorts were similar, with no change in safety events with increasing dose of MEDI4736.

In Study D4190C00006, of all treatment cohorts, the cohort of 11 patients treated in the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab group had the fewest AEs, Grade ≥3 AEs, SAEs, and treatment discontinuations due to AEs, but still showed strong evidence of clinical activity. This cohort had a lower number of treatment-related Grade ≥3 AEs or treatment related SAEs. No dose-limiting toxicities (DLTs) were reported.

Preliminary clinical activity of the MEDI4736 and tremelimumab combination did not appear to change with increasing doses of tremelimumab. The 15- and 20-mg/kg MEDI4736 q4w cohorts demonstrated objective responses at all doses of tremelimumab, and increasing doses of tremelimumab did not provide deeper or more rapid responses.

Efficacy data suggested that the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose cohort may demonstrate equivalent clinical activity to other dose combinations. A total of 5 of 11 patients in the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab cohort were evaluable for efficacy with at least 8 weeks of follow-up. Of these, there were 2 patients (40%) with partial response (PR), 1 patient (20%) with stable disease (SD), and 1 patient (20%) with progressive

disease (PD). (The fifth patient had only a single scan, which was conducted outside the window for these evaluations.)

Additionally, of all cohorts, the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose cohort had the fewest AEs, Grade ≥3 AEs, SAEs, and treatment discontinuations due to AEs, but still showed some evidence of clinical activity. All together, the data suggested that a 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose combination should be selected for further development.

1.2.2 Rationale for 4 cycles of combination therapy followed by MEDI4736 monotherapy

Long-term follow up on melanoma patients treated with ipilimumab, an anti-CTLA-4 targeting antibody (dosed every 3 weeks for 4 doses and then discontinued), shows that patients responding to ipilimumab derive long-term benefit, with a 3-year OS rate of approximately 22%. Furthermore, the survival curve in this population reached a plateau at 3 years and was maintained through 10 years of follow up (Schadendorf et al 2013).

Similar data have been presented for other anti-PD-1/PD-L1 targeting antibodies:

- Nivolumab (anti-PD-1) was dosed q2w for up to 96 weeks in a large Phase I dose-escalation and expansion study, and showed responses were maintained for a median of 22.94 months for melanoma (doses 0.1 mg/kg to 10 mg/kg), 17 months for NSCLC (doses 1, 3, and 10 mg/kg), and 12.9 months for renal cell carcinoma patients (doses 1 and 10 mg/kg) at the time of data analysis (Hodi et al 2014, Brahmer et al 2014, Drake et al 2013). Furthermore, responses were maintained beyond treatment discontinuation in the majority of patients who stopped nivolumab treatment (either due to protocol specified end of treatment, CR, or toxicity) for up to 56 weeks at the time of data analysis (Topalian et al 2014).

- MPDL3280a (anti-PD-L1) and the combination of nivolumab with ipilimumab, in which patients were dosed for a finite time period and responses maintained beyond treatment discontinuation have been reported (Herbst et al 2013, Wolchok et al 2013).

Similar long term results may be expected with use of other immune-mediated cancer therapeutics including anti-CTLA-4 antibodies such as tremelimumab, anti PD-L1 antibodies such as MEDI4736, or the combination of the two.

The MEDI4736 + tremelimumab combination regimen will be administered for 4 doses followed by monotherapy MEDI4736 20 mg/kg q4w.

1.2.3 MEDI4736 monotherapy dose rationale

A dose of MEDI4736 dose of 20 mg/kg q4w is supported by in-vitro data, non-clinical activity, clinical PK/pharmacodynamics, biomarkers, and activity data from Study 1108 in patients with advanced solid tumors (ongoing first-time-in-humans study).

As of 14 July 2014, 393 patients enrolled in Study 1108 have received MEDI4736, predominantly at 10 mg/kg q2w (either in the dose-escalation or dose-expansion phase of the study). Data presented at the European Society for Medical Oncology (ESMO) meeting 2014 with a later cut-off of 21 August 2014 showed that MEDI4736 was well tolerated at all doses in the NSCLC subset of patients enrolled into Study 1108, with drug-related Grade ≥3 AEs reported in 3% of patients and drug-related AEs leading to discontinuation reported in 1% of patients. No drug-related colitis or hyperglycemia of any grade, no Grade ≥3 pneumonitis reported, and no drug-related AEs leading to death were reported (Antonia et al 2014b). No DLTs were observed up to a dose of 10 mg/kg q2w or 15 mg/kg every 3 weeks (q3w).

Efficacy data on the NSCLC patients in Study 1108, presented at ESMO 2014 (cut-off date of 21 August 2014), showed a disease control rate at 12 weeks of 41% and objective response rate (ORR) of 16% among 162 evaluable patients, with activity observed in both squamous and non-squamous histologies. The ORR was higher (25%; 12 complete response [CR]/PR; n=48) in patients with PD-L1 positive tumors, defined as those with ≥25% of tumor cells with membrane staining for PD-L1, compared to patients with PD-L1 negative tumors (10%; 7 CR/PR; n=74) (Antonia et al 2014b).

Based on available PK/pharmacodynamic data from ongoing Study 1108 with doses ranging from 0.1 to 10 mg/kg q2w or 15 mg/kg q3w, MEDI4736 exhibited non-linear (dose-dependent) PK consistent with target-mediated drug disposition. The PK approached linearity at ≥3 mg/kg q2w, suggesting near complete target saturation (membrane-bound and sPD-L1), and further shows that the MEDI4736 dosing frequency can be adapted to a particular regimen given the linearity seen at higher doses than 3 mg/kg. The expected half life with doses ≥3 mg/kg q2w is approximately 21 days. A dose-dependent suppression in peripheral sPD-L1 was observed over the dose range studied, consistent with engagement of MEDI4736 with PD-L1. Dose-related changes in a variety of peripheral biomarkers have been observed over the dose range of 0.1 to 15 mg/kg. A low level of immunogenicity has been observed. No patients have experienced immune-complex disease following exposure to MEDI4736. Of the 220 patients who received MEDI4736 monotherapy and for whom PK/ADA data were available as of14 July 2014, 5 were ADA positive, with an impact on PK/pharmacodynamics reported in 1 patient at 3 mg/kg.

Data from Study 006 (Phase I trial in NSCLC patients using the combination of MEDI4736 and tremelimumab), also show an approximately dose-proportional increase in PK exposure for MEDI4736 over the dose range of 3 to 20 mg/kg MEDI4736 q4w or q2w. The observed MEDI4736 PK data from the combination study were well in line with the predicted monotherapy PK data (5^{th} median and 95^{th} percentiles) for a q4w regimen.

A population PK model was developed using the data from Study 1108 (doses=0.1 to 10 mg/kg q2w or 15 mg/kg q3w; Fairman et al 2014). Multiple simulations indicate that a similar overall exposure is expected following both 10 mg/kg q2w and 20 mg/kg q4w regimens, as represented by AUC at steady state (AUC_{ss}; 4 weeks). Median $C_{max,ss}$ is expected to be higher with 20 mg/kg q4w (~1.5 fold) and median $C_{trough,ss}$ is expected to be higher with 10 mg/kg q2w (~1.25 fold). Clinical activity with the 20 mg/kg q4w dosing regimen is anticipated to be consistent with 10 mg/kg q2w with the proposed similar dose of

20 mg/kg q4w expected to (a) achieve complete target saturation in majority of patients; (b) account for anticipated variability in PK, pharmacodynamic, and clinical activity in diverse cancer populations; (c) maintain sufficient PK exposure in case of ADA impact; and (d) achieve PK exposure that yielded maximal anti-tumor activity in animal models.

As of 8 April 2015, there is initial safety data for 16 patients receiving the 20 mg/kg q4w dosing regimen (12 patients from Study 1108 and 4 patients from the Japan Phase I trial). The toxicities observed with 20 mg/kg q4w are consistent with the 10 mg q2w regimen, and there were no DLTs observed. Of the 12 patients in Study 1108, 42% of patients have experienced any grade AE, with 2 being Grade 3 and above (17%). None of the Grade 3 and higher events was considered treatment-related. No patients on the Japan Phase I trial have experienced a Grade 3 or above AE. At present, the data do not suggest that the safety profile of MEDI4736 will be different in the 20 mg/kg q4w dosing regimen when compared to 10 mg/kg q2w regimen.

Given the similar AUC and modest differences in median peak and trough levels at steady state, the observation that both regimens maintain complete sPD-L1 suppression at trough, and the available clinical data, the 20 mg/kg q4w and 10 mg/kg q2w regimens are expected to have similar efficacy and safety profiles, supporting further development with a dose of 20 mg/kg q4w.

1.2.4 Rationale for Standard of Care as a comparator

The choice of SoC options provided in this study includes carboplatin and paclitaxel, carboplatin (or cisplatin) and gemcitabine (squamous only), carboplatin (or cisplatin) and pemetrexed (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only). Patients in the SoC group will receive treatment determined by the Investigator, from the SoC agents approved for use in NSCLC in their local market, until progression per standard practice. The SoC options provided in this study include agents that are commonly used in locally advanced or metastatic NSCLC and allow sufficient flexibility for Investigators and patients to select the agents that reflect their normal clinical practice and national guidelines (NCCN 2014 and Reck et al 2014).

1.2.5 Rationale for retreatment option

In contrast to patients treated with chemotherapy, who are unlikely to respond to rechallenge with the same agent upon progression, responses have been observed upon retreatment with IMTs. Several potential mechanisms of resistance to IMT exist, including loss of T-cell "memory" or recurrence of immune escape, which suggest retreatment for patients who initially respond or demonstrate stable disease is reasonable. Preliminary data in patients previously treated with IMTs suggest that responses are similar to those observed following initial treatment (Forde et al 2014; Hodi et al 2010). Therefore, patients in the MEDI4736 + tremelimumab combination therapy arm or the MEDI4736 monotherapy arm who achieve and maintain disease control (ie, CR, PR, or SD) through the end of the 12-month treatment period may restart treatment with their assigned treatment upon evidence of PD during follow-up.

In addition, patients in the combination therapy arm who complete the 4 dosing cycles with the combination of MEDI4736 + tremelimumab (with clinical benefit per Investigator judgement), but subsequently experience confirmed progression (according to Response Evaluation Criteria in Solid Tumors, version 1.1 [RECIST 1.1]) during treatment with MEDI4736 alone, will be given the option to restart MEDI4736 + tremelimumab therapy. All patients who restart treatment will receive a maximum of 12 months of further treatment.

1.2.6 Rationale for endpoints

The primary aim of this study is to determine the efficacy of MEDI4736 + tremelimumab combination therapy versus SoC in terms of PFS. PFS may serve as a surrogate endpoint for OS when differences between treatment groups are of sufficient magnitude and clinically important (FDA Guidance 2011, Pazdur 2008).

The secondary efficacy endpoints of ORR, DoR, proportion of patients alive and progression-free at 12 months from randomization (APF12), time to second progression (PFS2), and OS will be examined to further evaluate the anti-tumor effect and survival benefit of MEDI4736 + tremelimumab versus SoC. The secondary efficacy endpoints of PFS, ORR, PFS2, and OS will be examined to also further evaluate the anti-tumor effect and survival benefit of MEDI4736 monotherapy compared to SoC. PFS and ORR will be examined to further evaluate the anti-tumor effect of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy. ORR, DoR, and APF12 will be assessed using Investigator assessments according to RECIST 1.1.

In addition, PFS and ORR will be evaluated in a secondary analysis using Blinded Independent Central Review (BICR) assessments according to immune-related response criteria (irRECIST; Nishino et al 2013). Anti-tumor activity will be evaluated by irRECIST as a secondary endpoint in order to assess the utility of modified RECIST criteria in capturing the true clinical benefit of MEDI4736 and tremelimumab. Using irRECIST will allow taking into account the unique kinetics of responses that have been observed and well characterized with this class of agents, including in ongoing monotherapy studies.

The secondary health-related quality of life (HRQoL) assessments (the European Organisation for Research and Treatment of Cancer [EORTC] 30-item core quality of life questionnaire, version 3 [QLQ-C30 v3] and 13-item lung cancer quality of life questionnaire [QLQ-LC13]) will show the overall influence of the benefits and toxicity of the treatment from a patient's perspective and will aid in understanding of the benefit/risk evaluation. These patient-reported outcome (PRO) questionnaires are well-established instruments that have been previously included in lung cancer clinical trials.

The PK and immunogenicity of MEDI4736 and tremelimumab are being examined to assess the PK of both agents when administered in combination, and to assess their potential impact on PK, pharmacodynamics, and safety and efficacy parameters. Biological samples will be used to explore potential biomarkers in tumor, plasma, and/or serum that may influence pathogenesis, response, and clinical characteristics.

1.3 Benefit/risk and ethical assessment

The following sections include summaries of the potential benefits and risks associated with MEDI4736 monotherapy, tremelimumab monotherapy, and MEDI4736 + tremelimumab combination therapy, respectively, prior to the overall benefit:risk assessment.

1.3.1 Potential benefits

1.3.1.1 MEDI4736

The majority of the safety and efficacy data currently available for MEDI4736 are based on the first time in-human, single-agent study (Study 1108) in patients with advanced solid tumors. Data from Study 1108 were presented at the ESMO 2014 Congress. As of 21 August 2014, 162 patients with NSCLC were evaluable for response analysis. The disease control rate (DCR) at 12 weeks in patients receiving 10 mg/kg MEDI4736 q2w was 39%, and the ORR was 15% (26% [12 out of 47 patients] with known PD-L1-positive NSCLC [ie, ≥25% PD-L1 expression] and 10% [7 out of 74 patients] with known PD-L1-negative NSCLC [ie, <25% PD-L1 expression]). A total of 24% of patients receiving 10 mg/kg MEDI4736 q2w had SD for ≥12 weeks (including 21% [10 out of 47 patients] with known PD-L1-positive NSCLC and 32% [24 out of 74 patients] with known PD-L1-negative NSCLC). Responses were ongoing in 88% of patients with NSCLC, with an objective response duration ranging from 0.1 to 32.4 weeks (Antonia et al 2014b).

1.3.1.2 Tremelimumab

In a single-arm, Phase II study (Study A3671008) of tremelimumab administered at 15 mg/kg every 90 days to patients with refractory melanoma, an RR of 7% and a median OS of 10 months in the second-line setting (as compared to approximately 6 months with best supportive care reported from a retrospective analysis; Korn et al 2008) were observed (Kirkwood et al 2010). In a randomized, open-label, first-line Phase III study of tremelimumab (administered at 15 mg/kg every 90 days) versus chemotherapy (dacarbazine or temozolomide) in advanced melanoma (Study A3671009), results of the final analysis showed an RR of 11% and a median OS of 12.58 months in this first-line setting as compared to 10.71 months with standard chemotherapy; however, these results were not statistically significant (Ribas et al 2013). Additionally, a Phase II maintenance study (Study A3671015) in patients with Stage IIIB or IV NSCLC who had responded or remained stable failed to achieve statistical significance. The primary endpoint of PFS at 3 months was 22.7% in the tremelimumab arm (15 mg/kg) compared with 11.9% in the best supportive care arm (Study A3671015).

1.3.1.3 MEDI4736 + tremelimumab

The preclinical and clinical justification for this combination as noted in Section 1.2.1 also supports the synergy of this combination. Available data, such as those presented by Wolchok et al, suggest that the combination of agents targeting PD-1/PD-L1 and CTLA-4 may have profound and durable benefits in patients with melanoma (Wolchok et al 2013). Further, preliminary efficacy data from Study D4190C00006 has demonstrated that this combination is clinically active and well tolerated. As of 27 January 2015, 53 patients were

evaluable for response across various MEDI4736 + tremelimumab dose regiments. Of these, 12 patients (23%) had a best response of PR and 14 patients (26%) had a best response of SD. In the MEDI4736 20 mg/kg plus tremelimumab 1 mg/kg q4w cohort, a total of 5 of 11 patients were evaluable for efficacy with at least 8 weeks of follow-up. Of these, there were 2 patients (40%) with partial response (PR), 1 patient (20%) with stable disease (SD), and 1 patient (20%) with progressive disease (PD). (The fifth patient had only a single scan, which was conducted outside the window for these evaluations.)

Current experience with single-agent IMT studies suggests that clinical responses may be restricted to a subset of any given patient population and that it might be beneficial to enrich the patient population by selecting patients likely to respond to therapy. To date, no assay has been established or validated, and no single approach has proven accurate, for patient enrichment for IMTs. However, independent data from multiple sources using different assays and scoring methods suggests that PD-L1 expression on tumor cells and/or tumor infiltrating cells may be associated with greater clinical benefit.

Data from ongoing studies with MEDI4736 and other agents targeting the PD-1/PD-L1 pathway suggest, as shown in a number of tumor types (eg, NSCLC, renal cell carcinoma, and melanoma), that monotherapy may be more efficacious (in terms of ORR) in patients who are PD-L1-positive.

Given these findings, a number of ongoing studies are assessing the activity of agents in patients with PD-L1–positive tumors. There is also an unmet medical need in patients with PD-L1–negative tumors that needs to be addressed. Data, as of 27 January 2015 from Study 006 show that with the addition of tremelimumab to MEDI4736, the ORR can be increased to 25% in patients with PD-L1 negative NSCLC. As patients with PD-L1 positive disease can also have an increase in ORR, from 25% with MEDI4736 monotherapy, to 36% with the combination of MEDI4736 and tremelimumab, the study will enroll all patients with NSCLC, with an emphasis on those determined to be PD-L1 negative.

1.3.2 Potential risks

1.3.2.1 MEDI4736

Potential risks, based on the mechanism of action of MEDI4736 and related molecules, include immune-mediated reactions, such as enterocolitis, dermatitis, hepatitis/hepatotoxicity, endocrinopathy, pneumonitis, and neuropathy or neurologic events. Additional important potential risks include infusion-related reactions, hypersensitivity, anaphylaxis or serious allergic reactions, serious infections, and immune complex disease. Of the 393 patients with advanced solid tumors including NSCLC treated with 10 mg/kg q2w in Study 1108 (as of IB Edition 7.0 cut-off date of 14 July 2014), 331 patients (84.2%) had at least 1 adverse event (AE) regardless of causality. Treatment-related AEs were reported for 162 of 393 patients (41.2%). The most frequently reported (≥2% of patients) AEs assessed by the Investigator as treatment related (including all National Cancer Institute (NCI) Common Terminology Criteria for Adverse Event [CTCAE] grades) were fatigue (13.5%); nausea (8.4%); diarrhea, decreased appetite, and rash (5.3% each); vomiting (4.8%); pruritus (4.1%); dyspnea (3.8%); pyrexia (3.1%); hypothyroidism, increased ALT, increased AST, and cough (2.5% each);

myalgia (2.3%); and abdominal pain and dizziness (2.0% each). Two patients discontinued treatment due to AEs assessed by the Investigator as treatment related (pneumonitis and increased transaminases). Grade 3 events assessed by the Investigator as treatment related that occurred in 2 or more patients were fatigue (4 patients); increased gamma-glutamyltransferase (3 patients); and vomiting, increased ALT, increased AST, and arthralgia (2 patients each). There were 2 patients with Grade 4 events (hypercalcemia and fatigue) assessed by the Investigator as treatment related and 1 patient with a Grade 5 event assessed by the Investigator as treatment related (angiopathy). For further details on the safety profile of MEDI4736 as monotherapy or combination therapy, please refer to the current IB.

Other mAbs targeting the PD-1/PD-L1 pathway are currently in clinical development. Among the most frequent AEs noted with treatment with these antibodies are fatigue, rash, diarrhea, and pruritus. Immune-mediated AEs of Grade ≥3 reported include pneumonitis, diarrhea, increased ALT, and increased AST. Other relevant risks include those associated with biological and immunotherapeutic agents, including infusion reactions and acute immunoglobulin E-mediated allergic reactions.

1.3.2.2 Tremelimumab

Potential risks, based on the mechanism of action of tremelimumab and related molecules (ipilimumab) include potentially immune-mediated gastrointestinal events including enterocolitis, intestinal perforation, abdominal pain, dehydration, nausea and vomiting, and decreased appetite (anorexia); dermatitis including urticaria, skin exfoliation, and dry skin; endocrinopathies including hypophysitis, adrenal insufficiency, and hyperthyroidism and hypothyroidism; hepatitis including autoimmune hepatitis and increased serum ALT and AST; pancreatitis including autoimmune pancreatitis and lipase and amylase elevation; respiratory tract events including pneumonitis and interstitial lung disease; nervous system events including encephalitis, peripheral motor and sensory neuropathies, and Guillain-Barré syndrome; cytopenias including thrombocytopenia, anemia, and neutropenia; infusion-related reactions; anaphylaxis; and serious allergic reactions. The profile of AEs and the spectrum of event severity have remained stable across the tremelimumab clinical program and are consistent with the pharmacology of the target. To date, no tumor type or stage appears to be associated with unique AEs (except for vitiligo that appears to be confined to patients with melanoma). Overall, 944 of the 973 patients (97.0%) treated with tremelimumab monotherapy as of the data cutoff date of 12 November 2014 (for all studies except D4190C00006 that has a cutoff date of 04 December 2014 and not including 497 patients who have been treated in the ongoing blinded Phase IIb Study D4880C00003) experienced at least 1 AE. The events resulted in discontinuation of tremelimumab in 10.0% of patients, were serious in 36.5%, were Grade ≥3 in severity in 49.8%, were fatal in 67.7%, and were considered to be treatment related in 79.1% of patients. The frequency of any AEs and Grade ≥3 AEs was generally similar across the tremelimumab dose groups. However, a higher percentage of patients in the 10 mg/kg every 28 days and 15 mg/kg every 90 days groups compared with the All Doses <10 mg/kg group experienced treatment-related AEs, serious AEs (SAEs), AEs resulting in discontinuation of investigational product (IP), and deaths.

1.3.2.3 MEDI4736 + tremelimumab

No safety studies in animals have been performed combining tremelimumab with MEDI4736. As both CTLA-4 and PD-L1 have mechanisms of actions that enhance activation of immune cells, their potential to induce cytokine release was tested in a whole-blood assay system. MEDI4736 and tremelimumab, either alone or in combination, did not induce cytokine release in blood from any donor.

Evaluation of the safety of MEDI4736 + tremelimumab in the ongoing Study D4190C00006, in patients with NSCLC, has so far shown a manageable safety and tolerability profile of the combination therapy.

Overall, 62 (83.8%) of the 74 patients reported an AE regardless of causality. The most frequently (10 or more patients) reported AEs were fatigue (37.8%; 28 patients); diarrhea (32.4%; 24 patients); amylase increased and pruritus (16.2%; 12 patients); decreased appetite, dyspnea, nausea, and rash (14.9%; 11 patients each), and headache and pyrexia (13.5%; 10 patients). Additional safety results from this study are presented in Section 1.2.1 and the MEDI4736 IB

In the literature (Wolchok et al 2013), using the combination of the same class of drugs (eg, anti-PD-1 and anti-CTLA4 antibodies), specifically nivolumab + ipilimumab in a study involving patients with malignant melanoma, the safety profile of this combination had shown occurrences of AEs assessed by the Investigator as treatment-related in 93% of treated patients, with the most frequent events being rash (55% of patients), pruritus (47% of patients), fatigue (38% of patients), and diarrhea (34% of patients). Grade 3 or 4 AEs, regardless of causality, were noted in 72% of patients, with Grade 3 or 4 events assessed by the Investigator as treatment-related in 53%. The most frequent of these Grade 3 or 4 events assessed by the Investigator as treatment-related include increased lipase (in 13% of patients), AST (in 13%), and ALT levels (in 11%). Frequent Grade 3 or 4 selected AEs assessed by the Investigator as treatment-related in the combination therapy included hepatic events (in 15% of patients), gastrointestinal (GI) events (in 9%), and renal events (in 6%). Isolated cases of pneumonitis and uveitis were also observed.

1.3.3 Overall benefit-risk and ethical assessment

There remains a significant unmet medical need for additional treatment options for patients with EGFR and ALK wild-type, locally advanced or metastatic NSCLC who have not received prior chemotherapy or any systemic therapy for locally advanced or metastatic NSCLC.

The study design aims to minimize potential risks; intensive monitoring, including early safety assessment, is in place for those risks deemed to be most likely based on prior experience with the IPs (ie, MEDI4736 + tremelimumab, MEDI4736, and SoC).

Based upon the available non-clinical and clinical safety data, the limited survival benefit provided by the currently available treatment options to patients, the limited life expectancy due to malignant disease, the activity seen with MEDI4736 in this tumor type, and the strength

of the scientific hypotheses under evaluation, the MEDI4736 + tremelimumab combination and the MEDI4736 monotherapy treatments proposed for evaluation in this study may have the potential to provide meaningful clinical benefit with a manageable safety and tolerability profile by generating durable clinical responses, thereby improving quality of life (QoL) and potentially extending survival.

Therefore, the investigation of the potential therapeutic efficacy of the combination of MEDI4736 with tremelimumab and MEDI4736 monotherapy in patients with PD-L1-positive and -negative tumors is acceptable, and the overall benefit/risk assessment supports the proposed study design.

1.4 Study design

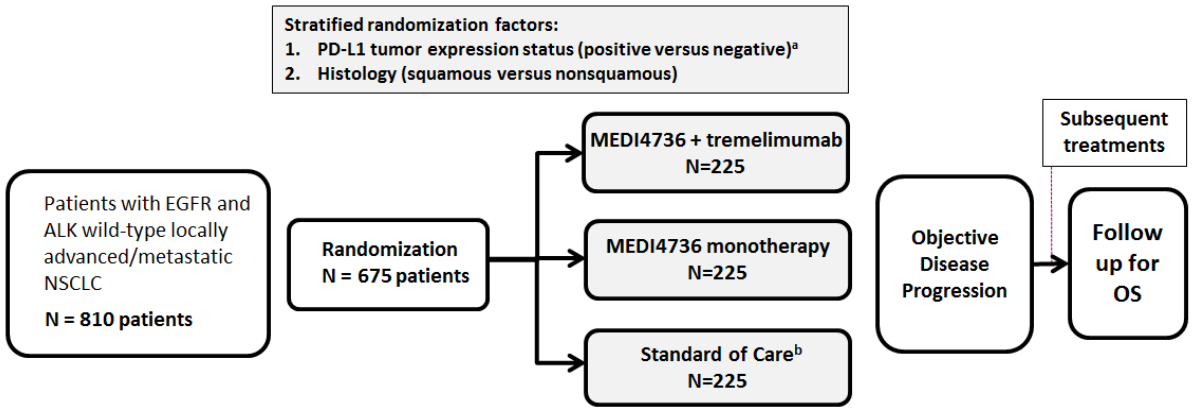
This is a randomized, open-label, multi-center, global Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with EGFR and ALK wild-type locally advanced or metastatic NSCLC. A schematic diagram of the overall study design is shown in Figure 1, and a detailed study flow chart is shown in Figure 2.

This study will enroll approximately 810 patients at sites in North America, Asia, Australia, and Europe to randomize approximately 675 patients (including at a minimum 438 patients with PD-L1–negative NSCLC) to treatment.

Patients will provide a tumor tissue sample at screening (newly acquired or archived sample <3 months old) to determine PD-L1 expression status (defined by an immunohistochemistry [IHC] assay developed by in which ≥25% PD-L1–membrane expression in tumoral tissue is considered positive and <25% is considered negative for PD-L1 expression; referred to hereafter as patients with PD-L1-positive or PD-L1-negative tumors, respectively).

Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (as described above) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy. Doses and treatment regimens are described in Section 7.2. Assessments will be conducted as indicated in Table 2, Table 3, and Table 4.

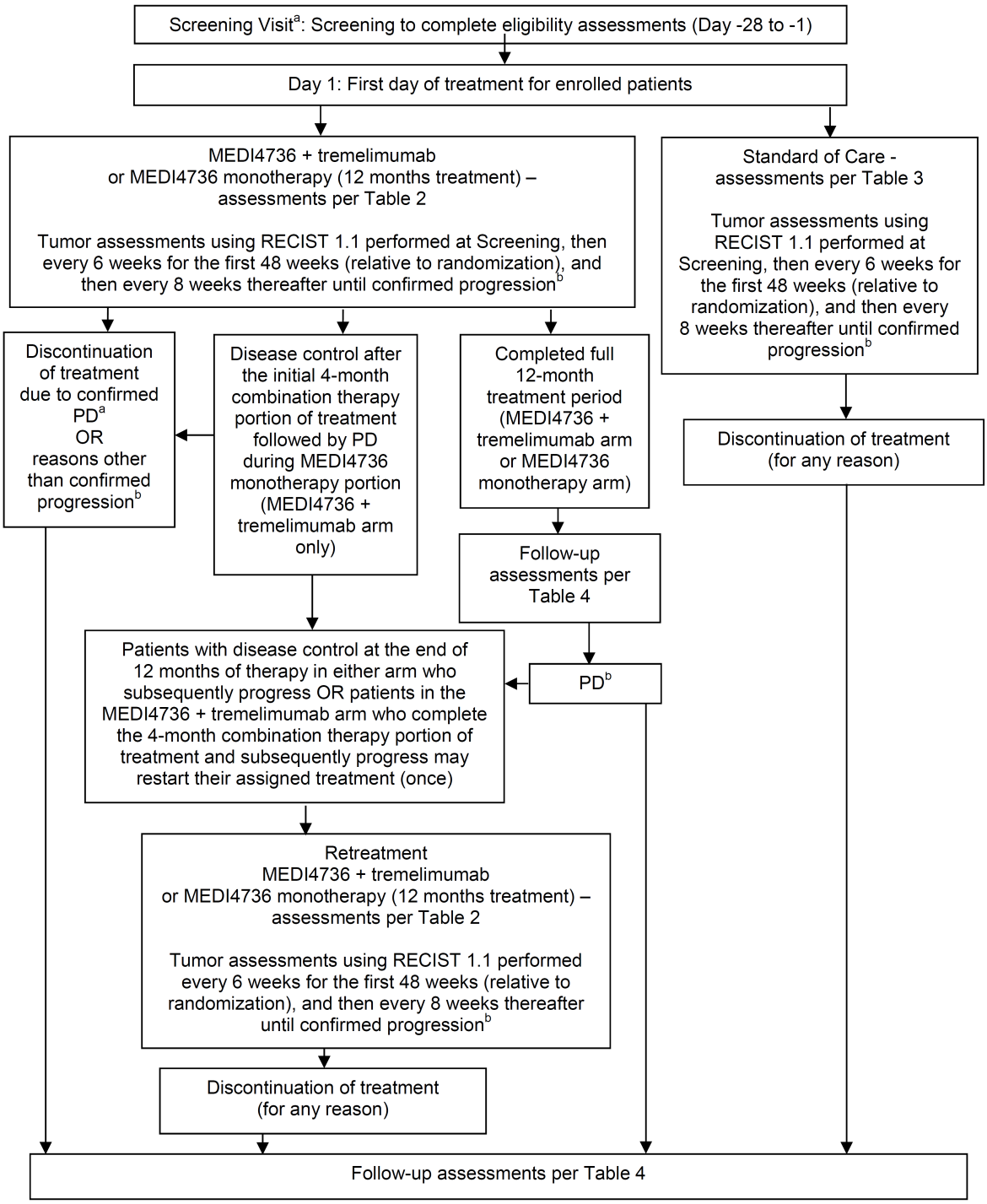
Figure 1 Overall study design



[a] Sites will be supplied with PD-L1 status upon request at disease progression.

[b] Standard of Care is an Investigator choice from the following: paclitaxel + carboplatin, gemcitabine + cisplatin (or carboplatin) (squamous only), pemetrexed + cisplatin (or carboplatin) (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only following pemetrexed/platinum induction).

Figure 2 Study flow chart



[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization.

[b] A confirmatory scan is always required following the initial demonstration of PD. (See Section 5.1 for more information.)

2. STUDY OBJECTIVES

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1-positive and/or PD-L1–negative NSCLC.

2.1 Primary objective

Primary Objective:	Outcome Measure:
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with NSCLC	PFS using Investigator assessments according to RECIST 1.1[a]

[a] Sensitivity analyses of PFS will be performed based on BICR assessment according to RECIST 1.1 (subset of patients with BICR assessment only) and Investigator assessment according to RECIST 1.1 modified for confirmation of progression. If bias cannot be excluded based upon the sample BICR assessment, then an independent evaluation of all radiographic images may be required for the assessment of the primary PFS endpoint.

2.2 Secondary objectives

Secondary Objectives:	Outcome Measures:
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, ORR, DoR, APF12, PFS2, and OS	PFS in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1 ORR, DoR, and APF12 using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, ORR, PFS2, and OS	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS in patients with PD-L1–positive NSCLC using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice OS
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS and ORR	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS and ORR in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1
To assess disease-related symptoms and HRQoL in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC QLQ-C30 v3 and the LC13 module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.

To assess the PK of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of ADAs for MEDI4736 and tremelimumab
To explore irRECIST as an assessment methodology for clinical benefit of MEDI4736 + tremelimumab compared to SoC with assessment by BICR	PFS and ORR using BICR assessment according to irRECIST

2.3 Safety objective

Safety Objective:	**Outcome Measures:**
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of locally advanced or metastatic NSCLC patients	AEs, physical examinations, laboratory findings, and vital signs

2.4 Exploratory objectives

Exploratory Objectives:	**Outcome Measures:**
To assess AEs by patient self-reporting of specific CTCAE symptoms	Collection of approximately 20 patient-reported outcomes version of CTCAE (PRO-CTCAE) symptoms via an electronic device solution
To assess patients' overall impression of the change in their health status since the start of study treatment	Patients' Global Impression of Change (PGIC) item will be collected directly from patients via an electronic device solution
To investigate the relationship between PK exposure and clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze PK exposure and the relationship with clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, as deemed appropriate
To describe and evaluate resource use associated with assigned treatments and underlying disease during assigned treatment	Health resource utilization measures including hospitalization, outpatient visits, or emergency department visits
To explore the impact of treatment and disease state on health state utility using the EQ-5D-5L during assigned treatment	EQ-5D-5L health state utility index will be used to derive health state utility based on patient-reported data

To investigate associations between pre-treatment peripheral myeloid-derived suppressor cells (MDSCs) measures and clinical activity	A graphical and/or a data modeling approach will be used to analyze the relationship between MDSC counts with clinical outcomes and/or with tumor lesion measurements
To investigate the relationship between biomarkers and clinical outcomes, efficacy, AEs, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze biomarkers (eg, IFNγ and/or PD-L1 status defined under alternative methods) and the relationship with clinical outcomes, efficacy, AEs, and/or safety parameters, as deemed appropriate

Note: Exploratory objective analyses may be reported separately from the main clinical study report.

3. PATIENT SELECTION, ENROLLMENT, RANDOMIZATION, RESTRICTIONS, DISCONTINUATION AND WITHDRAWAL

Each patient must meet all of the inclusion criteria (Section 3.1) and none of the exclusion criteria (Section 3.2) for this study. Under no circumstances will there be exceptions to this rule.

3.1 Inclusion criteria

For inclusion in the study, patients should fulfill the following criteria:

1. Age ≥18 years at the time of screening

2. Written informed consent and any locally required authorization (eg, Health Insurance Portability and Accountability Act in the US, European Union [EU] Data Privacy Directive in the EU) obtained from the patient/legal representative prior to performing any protocol-related procedures, including screening evaluations. (For patients aged <20 years and enrolling in Japan, a written informed consent should be obtained from the patient and his or her legally acceptable representative.)

3. Histologically or cytologically documented Stage IV NSCLC with locally advanced disease not amendable to curative surgery or radiation (according to version 7 of the International Association for the Study of Lung Cancer Staging Manual in Thoracic Oncology; IASLC Staging Manual in Thoracic Oncology).

4. Patients must have tumors that lack activating EGFR mutation (ie, exon 19 deletion or exon 21 L858R mutation) and ALK rearrangement. (If a patient is known to have a tumor with a KRAS mutation, then EGFR and ALK testing is not required).

5. No prior chemotherapy or any other systemic therapy for locally advanced or metastatic NSCLC. Patients who have received prior platinum-containing adjuvant, neoadjuvant, or definitive chemoradiation for locally advanced disease are eligible, provided that progression has occurred >6 months from last therapy.

6. Tumor PD-L1 status, with IHC assay confirmed by a reference laboratory, must be known prior to randomization. As such, all patients must be able to undergo a fresh tumor biopsy during screening or to provide an available tumor sample taken <3 months prior to screening. Tumor lesions used for fresh biopsies should not be target lesions, unless there are no other lesions suitable for biopsy. Fine needle aspirate specimens are not acceptable. Specimens from metastatic bone lesions are typically unacceptable unless there is a significant soft tissue component. The tumor specimen submitted to establish eligibility should be of sufficient quantity to allow for PD-L1 IHC and other exploratory biomarker analyses and is preferred in formalin-fixed paraffin embedded blocks.

7. World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 at enrolment.

8. At least 1 lesion, not previously irradiated, that can be accurately measured at baseline as ≥10 mm in the longest diameter (except lymph nodes which must have a short axis ≥15 mm) with CT or MRI and that is suitable for accurate repeated measurements as per RECIST 1.1 guidelines.

9. No prior exposure to immune-mediated therapy including, but not limited to, other anti-CTLA-4, anti-PD-1, anti-PD-L1, and anti-programmed cell death ligand 2 (anti-PD-L2) antibodies, excluding therapeutic anticancer vaccines.

10. Adequate organ and marrow function as defined below:

    - Hemoglobin ≥9.0 g/dL
    - Absolute neutrophil count $\geq 1.5 \times 10^9$ /L
    - Platelet count $\geq 100 \times 10^9$/L
    - Serum bilirubin ≤1.5 × the ULN. This will not apply to patients with confirmed Gilbert's syndrome, who will be allowed in consultation with their physician.
    - ALT and AST ≤2.5 × ULN; for patients with hepatic metastases, ALT and AST ≤5 × ULN
    - Calculated creatinine clearance >50 mL/min as determined by Cockcroft-Gault (using actual body weight)

      Males:

      $$\text{Creatinine CL (mL/min)} = \frac{\text{Weight (kg)} \times (140 - \text{Age})}{72 \times \text{serum creatinine (mg/dL)}}$$

      Females:

$$\text{Creatinine CL (mL/min)} = \frac{\text{Weight (kg)} \times (140 - \text{Age})}{72 \times \text{serum creatinine (mg/dL)}} \times 0.85$$

11. Evidence of post-menopausal status or negative urinary or serum pregnancy test for female pre-menopausal patients. Women will be considered post-menopausal if they have been amenorrheic for 12 months without an alternative medical cause. The following age-specific requirements apply:

    – Women <50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of exogenous hormonal treatments and if they have luteinizing hormone and follicle-stimulating hormone levels in the post-menopausal range for the institution or underwent surgical sterilization (bilateral oophorectomy or hysterectomy).

    – Women ≥50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of all exogenous hormonal treatments, had radiation-induced oophorectomy with last menses >1 year ago, had chemotherapy-induced menopause with >1 year interval since last menses, or underwent surgical sterilization (bilateral oophorectomy or hysterectomy).

3.2 Exclusion criteria

Patients should not enter the study if any of the following exclusion criteria are fulfilled:

1. Involvement in the planning and/or conduct of the study (applies to both AstraZeneca staff and/or staff at the study site)

2. Concurrent enrolment in another clinical study, unless it is an observational (non-interventional) clinical study or the follow-up period of an interventional study

3. Mixed small-cell lung cancer and NSCLC histology

4. Any concurrent chemotherapy, IP, biologic, or hormonal therapy for cancer treatment. Concurrent use of hormonal therapy for non-cancer-related conditions (eg, hormone replacement therapy) is acceptable.

5. Radiotherapy treatment to more than 30% of the bone marrow or with a wide field of radiation within 4 weeks of the first dose of study drug. Note: Local treatment of isolated lesions, excluding target lesions, for palliative intent is acceptable.

6. Major surgical procedure (as defined by the Investigator) within 28 days prior to the first dose of IP. Note: Local surgery of isolated lesions for palliative intent is acceptable.

7. History of allogenic organ transplantation

8. Active or prior documented autoimmune or inflammatory disorders (including inflammatory bowel disease [eg, colitis or Crohn's disease], diverticulitis with the exception of diverticulosis, celiac disease or other serious gastrointestinal chronic conditions associated with diarrhea), systemic lupus erythematosus, Sarcoidosis syndrome, or Wegener syndrome (granulomatosis with polyangiitis), Graves' disease, rheumatoid arthritis, hypophysitis, uveitis, etc within the past 3 years prior to the start of treatment. The following are exceptions to this criterion:

   - Patients with vitiligo or alopecia
   - Patients with hypothyroidism (eg, following Hashimoto syndrome) stable on hormone replacement or psoriasis not requiring systemic treatment

9. Any condition that, in the opinion of the Investigator, would interfere with the evaluation of IP or interpretation of patient safety or study results, including, but not limited to, ongoing or active infection, symptomatic congestive heart failure, uncontrolled hypertension, unstable angina pectoris, cardiac arrhythmia, interstitial lung disease, or psychiatric illness/social situations that would limit compliance with study requirement, substantially increase risk of incurring AEs from MEDI4736 or tremelimumab, or compromise the ability of the patient to give written informed consent

10. No medical contraindication to platinum (cisplatin or carboplatin)-based doublet chemotherapy

11. History of another primary malignancy except for

   - Malignancy treated with curative intent and with no known active disease ≥5 years before the first dose of study drug and of low potential risk for recurrence
   - Adequately treated non-melanoma skin cancer or lentigo maligna without evidence of disease
   - Adequately treated carcinoma in situ without evidence of disease (eg, cervical cancer in situ)

12. History of leptomeningeal carcinomatosis

13. Brain metastases or spinal cord compression unless asymptomatic or treated and stable off steroids and anti-convulsants for at least 1 month prior to study treatment. Patients with suspected brain metastases at screening should have a CT/MRI of the brain prior to study entry.

14. Mean QT interval corrected for heart rate using Fridericia's formula (QTcF) ≥470 ms

15. History of active primary immunodeficiency

16. Known history of clinical diagnosis of tuberculosis

17. Active infection including hepatitis B, hepatitis C, or human immunodeficiency virus (HIV)

18. Current or prior use of immunosuppressive medication within 14 days before the first dose of MEDI4736 or tremelimumab. The following are exceptions to this criterion:

    - Intranasal, inhaled, topical steroids, or local steroid injections (eg, intra-articular injection).
    - Systemic corticosteroids at physiologic doses not to exceed 10 mg/day of prednisone or its equivalent
    - Steroids as premedication for hypersensitivity reactions (eg, CT scan premedication)

19. Receipt of live, attenuated vaccine within 30 days prior to the first dose of IP. Note: Patients, if enrolled, should not receive live vaccine during the study and up to 30 days after the last dose of IP.

20. Female patients who are pregnant or breast-feeding or male or female patients of reproductive potential who are not willing to employ effective birth control from screening to 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy.

21. Known allergy or hypersensitivity to IP or any excipient

Procedures for withdrawal of incorrectly enrolled patients are presented in Section 3.4.

3.3 Patient enrolment and randomization

Investigators should keep a record (ie, the patient screening log) of patients who entered screening.

At screening/baseline (Days -28 to -1), the Investigators or suitably trained delegate will:

1. Obtain signed informed consent before any study specific procedures are performed. (Informed consent of study procedures may be obtained prior to the 28-day screening window in order to permit tumor biopsy sample acquisition which must be analyzed prior to randomization.)

2. Obtain a unique 7-digit enrolment number (E-code), through the IVRS/IWRS in the following format (ECCNNXXX: CC being the country code, NN being the center

number, and XXX being the patient enrolment code at the center). This number is the patient's unique identifier and is used to identify the patient on the electronic case report forms (eCRFs).

3. Obtain tumor sample and send for PD-L1 expression (Obtaining the tumor biopsy sample should be given the highest priority and, as such, the sample may be obtained and sent for PD-L1 expression status evaluation prior to the 28-day screening window in order to permit analysis prior to randomization.)

4. Determine patient eligibility (see Sections 3.1 and 3.2)

5. Obtain signed informed consent for genetic research study (optional)

At randomization, once the patient is confirmed to be eligible, the Investigator or suitably trained delegate will:

1. Define the SoC treatment (based on the most appropriate option for the patient) that the patient would receive if randomized to the SoC group prior to randomization of the patient. This must be completed for all patients. The information will be recorded in the IVRS/IWRS system.

   Note, for all patients with non-squamous tumor histology who are scheduled to receive pemetrexed if randomized to the SoC group, folic acid and vitamin B12 should commence prior to randomization for up to 7 days, in line with local practice. This is to ensure SoC treatment can begin on Day 1.

2. Obtain a unique randomization number via the IVRS/IWRS. Numbers will start at 001 and will be assigned strictly sequentially by IVRS/IWRS as patients are eligible for entry into the study. The system will randomize the eligible patient to 1 of the 3 treatment groups. (PD-L1 expression status results must be received from the central laboratory by the IVRS/IWRS prior to randomization.)

If the patient is ineligible and not randomized, the IVRS/IWRS should be contacted to terminate the patient in the system.

Patients will begin treatment on Day 1. Patients must not be treated unless all eligibility criteria have been met.

If a patient withdraws from participation in the study, then his or her enrolment/randomization code cannot be reused. Withdrawn patients will not be replaced.

3.4 Procedures for handling incorrectly enrolled patients

Patients who fail to meet the eligibility criteria should not, under any circumstances, be randomized or receive study medication. There can be no exceptions to this rule. Patients who are enrolled but found to not meet all the eligibility criteria must not be randomized, and must not be initiated on treatment and must be withdrawn from the study as a screen failure.

When a patient does not meet all the eligibility criteria but is randomized in error or incorrectly started on treatment, the Investigator should inform the Study Physician immediately, and the Study Physician and the Investigator should discuss whether to continue or discontinue the patient from treatment. The Study Physician must ensure that all decisions are appropriately documented.

3.5 Methods for assigning treatment groups

The actual treatment given to patients will be determined by the randomization scheme in the IVRS/IWRS. The randomization scheme will be produced by a computer software program that incorporates a standard procedure for generating randomization numbers.
One randomization list will be produced for each of the randomization strata. A blocked randomization will be generated, and all centers will use the same list in order to minimize any imbalance in the number of patients assigned to each treatment group.

Patients will be identified to the IVRS/IWRS per country regulations. Randomization codes will be assigned strictly sequentially, within each stratum, as patients become eligible for randomization. The IVRS/IWRS will provide the kit identification number to be allocated to the patient at the randomization visit.

3.6 Methods for ensuring blinding

Not applicable; this study is not blinded.

3.7 Methods for unblinding

Not applicable; this study is not blinded.

3.8 Restrictions

The following restrictions apply while the patient is receiving study treatment and for the specified times before and after:

1. Females of childbearing potential who are sexually active with a non-sterilized male partner must use 2 methods of effective contraception (Table 1) from the time of screening and must agree to continue using such precautions for 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy; cessation of birth control after this point should be discussed with a responsible physician. Periodic abstinence, the rhythm method, and the withdrawal method are not acceptable methods of birth control.

   – Females of childbearing potential are defined as those who are not surgically sterile (ie, bilateral tubal ligation, bilateral oophorectomy, or complete hysterectomy) or post-menopausal (defined 12 months with no menses without an alternative medical cause).

2. Non-sterilized males who are sexually active with a female partner of childbearing potential must use 2 acceptable methods of effective contraception (see Table 1)

from screening through 180 days after receipt of the final dose of MEDI4736 + tremelimumab combination therapy or 90 days after receipt of the final dose of MEDI4736 monotherapy.

3. **Patients in the SoC group:** Follow the local prescribing information relating to contraception, the time limits for such precautions, and any additional restrictions for agents in the SoC group.

4. **All patients:** Patients should not donate blood while participating in this study and for 3 months following the last dose of study treatment.

Restrictions relating to concomitant medications are described in Section 7.7.

Table 1 Effective methods of contraception (2 methods must be used)

Barrier methods	Intrauterine device methods	Hormonal methods
Male condom plus spermicide	Copper T	Implants
Cap plus spermicide	Progesterone T[a]	Hormonal shot or injection
Diaphragm plus spermicide	Levonorgestrel-releasing intrauterine system (eg, Mirena®)[a]	Combined pill Minipill Patch

[a] This is also considered to be a hormonal method.

3.9 Discontinuation of investigational product

An individual patient will not receive any further IP (MEDI4736 + tremelimumab combination, MEDI4736 monotherapy, or SoC) if any of the following occur in the patient in question:

- Withdrawal of consent from further treatment with IP. The patient is, at any time, free to discontinue treatment, without prejudice to further treatment. A patient who discontinues treatment is normally expected to continue to participate in the study unless they specifically withdraw their consent to further participation in any study procedures and assessments (see Section 3.10.2).

- An AE that, in the opinion of the Investigator or AstraZeneca, contraindicates further dosing

- Any AE that meets criteria for discontinuation as defined in Section 6.1.

- Pregnancy or intent to become pregnant

- Non-compliance with the study protocol that, in the opinion of the Investigator or AstraZeneca, warrants withdrawal from study treatment (eg, refusal to adhere to scheduled visits)

- Initiation of alternative anticancer therapy including another investigational agent
- Confirmed PD and Investigator determination that the patient is no longer benefiting from treatment with IP

3.9.1 Procedures for discontinuation of a patient from investigational product

At any time, patients are free to discontinue IP without prejudice to further treatment. A patient who decides to discontinue IP will always be asked about the reason(s) for discontinuation and the presence of any AE. If possible, they will be seen and assessed by an Investigator. AEs will be followed up (see Section 6). The Study Physician should be notified of any ongoing AE that may delay treatment or necessitate permanent discontinuation of treatment.

Patients who are permanently discontinued from further receipt of IP, regardless of the reason, will be identified as having permanently discontinued treatment. Patients who are permanently discontinued will enter follow-up (see Table 4). All patients will be followed for survival until the end of the study. Patients who decline to return to the site for evaluations should be contacted by telephone as indicated in Table 4 as an alternative.

Any patient who discontinues study treatment for reasons other than objective disease progression should have tumor assessments performed as scheduled in Table 4 until objective disease progression is documented or death occurs, unless consent is withdrawn.

Patients who have permanently discontinued from further receipt of IP will need to be discontinued from the IVRS/IWRS.

3.10 Criteria for withdrawal of the patient from the study

3.10.1 Screen failures

Screen failures are patients who do not fulfill the eligibility criteria for the study, and therefore must not be randomized. These patients should have the reason for study withdrawal recorded as “eligibility criteria not fulfilled” (ie, patient does not meet the required inclusion/exclusion criteria). This reason for study withdrawal is only valid for screen failures (ie, not randomized patients). Patients can be rescreened a single time, but they cannot be re-randomized.

3.10.2 Withdrawal of the informed consent

Patients are free to withdraw from the study at any time (IP and assessments) without prejudice to further treatment.

Patients who withdraw consent for further participation in the study will not receive any further IP or further study observation, with the exception of follow-up for survival, which will continue until the end of the study unless the patient has expressly withdrawn their consent to survival follow-up. Note that the patient may be offered additional tests or tapering of treatment to withdraw safely.

A patient who withdraws consent will always be asked about the reason(s) for withdrawal and the presence of any AE. The Investigator will follow up AEs outside of the clinical study. The patient will return electronic PRO devices, if applicable.

Patients will be considered lost to follow-up only if no contact has been established by the time the study is completed (see Section 9.3), such that there is insufficient information to determine the patient's status at that time. Patients who refuse to continue participation in the study, including telephone contact, should be documented as "withdrawal of consent" rather than "lost to follow-up." Investigators should document attempts to re-establish contact with missing patients throughout the study period. If contact with a missing patient is re-established, the patient should not be considered lost to follow-up and evaluations should resume according to the protocol.

3.10.2.1 Survival status for withdrawn consent and lost to follow up patients

At the time of PFS analyses, all enrolled patients' survival status should be re-checked, this includes those patients who withdrew consent or are classified as "lost to follow up."

- Lost to Follow up – site personnel should check hospital records, the patients' current physician, and a publicly available death registry (if available) to obtain a current survival status in the 7 days following data cutoff. (The SURVIVE module will be updated.)

In the event that the patient has actively withdrawn consent to the processing of their personal data, the survival 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.

3.11 Discontinuation of the study

The study may be stopped if, in the judgment of AstraZeneca, study patients are placed at undue risk because of clinically significant findings that meet any of the following criteria:

- Meet individual stopping criteria or are otherwise considered significant
- Are assessed as causally related to study drug
- Are not considered to be consistent with continuation of the study

In addition, the study may be stopped based on the findings of the interim safety analysis conducted by the Independent Data Monitoring Committee (IDMC) (see Section 6.8).

Regardless of the reason for termination, all data available for the patients at the time of discontinuation of follow-up must be recorded in the eCRFs. All reasons for discontinuation of treatment must be documented.

In terminating the study, AstraZeneca will ensure that adequate consideration is given to the protection of the patients' interests. If this study is discontinued, all other studies involving MEDI4736 or tremelimumab will remain open to enrolment and screening, if deemed appropriate by AstraZeneca.

4. STUDY PLAN AND TIMING OF PROCEDURES

The procedures for the screening and the 12-month treatment periods in this study are presented for the MEDI4736 + tremelimumab combination therapy group and MEDI4736 monotherapy group in Table 2 and for the SoC therapy group in Table 3. The procedures for the follow-up period are presented in Table 4.

Table 2 Schedule of assessments for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy treatment and retreatment periods (12 months)

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Informed consent											
Informed consent: study procedures	X[a]										4.1, 10.4
Consent: genetic sample and analysis (optional)	X										3.3
Study procedures											
Physical exam (full)	X										5.2.2
Targeted physical exam (based on symptoms)		X	X	X	X	X	X	X	X	X	5.2.2
Vital signs[b]	X	X	X	X	X	X	X	X	X	X	5.2.4
ECG[c]	X	As clinically indicated									5.2.3
Concomitant medications	<----------									---------->	7.7
Demography, including baseline characteristics and tobacco use	X										4.1
Eligibility criteria	X										3.1, 3.2
Laboratory assessments											
Clinical chemistry	X	X[d]	X	X	X	X	X	X	X	X	Table 5
Liver function tests		To be completed prior to each infusion									Table 5
Hematology	X	X[d]	X	X	X	X	X	X	X	X	Table 6
TSH, free T_3, and free T_4[e]	X	X	X	X	X	X	X	X	X	X	5.2.1
Urinalysis	X	X	X	X	X	X	X	X	X	X	Table 7
Hepatitis B and C and HIV	X										5.2.1
Pregnancy test[f]	X	As clinically indicated									5.2.1

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Pharmacokinetics											
MEDI4736 PK sample (serum)		X[g]			X[g]			X[g]			5.4.1
Tremelimumab PK sample (serum; combination therapy arm only)		X[g]			X[g]						5.4.1
Monitoring											
WHO/ECOG performance status	X	X	X	X	X	X	X	X	X	X	5.3.3
AE/SAE assessment	<----------									---------->	6.2.1
Drug accountability		X	X	X	X	X	X	X	X	X	7.6
Pre-randomization medication											
Folic acid[h]	X	Discontinue as randomized to combination therapy									3.3
IM Vitamin B12[h]	X	Discontinue as randomized to combination therapy									3.3
IP administration											
Combination therapy arm											
MEDI4736 (combination therapy)[i,j]		X	X	X	X	X	X	X	X	X	7.2.1
Tremelimumab[i,j]		X	X	X	X						7.2.1
Monotherapy arm											
MEDI4736 (monotherapy)[j]		X	X	X	X	X	X	X	X	X	7.2.1
PRO assessments[k]											
EORTC QLQ-C30, EQ-5D-5L		Every 4 weeks for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)									5.3.1.1, 5.3.1.5
EORTC QLQ-LC13, PRO-CTCAE[l]		Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)									5.3.1.2, 5.3.1.3
PGIC			X	X	X			X		X	5.3.1.4

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Other laboratory assessments and assays											
Immunogenicity assessment (ADA sampling to identify ADA responses in patient circulation)		X			X			X[m]			5.4.2
sPD-L1 (serum)		X			X			X			5.5.2
Circulating soluble factors (plasma)		X	X		X						5.5.2
Tumor biopsy (newly acquired or archived <3 months old)	X[a]										5.5.1
Whole blood for SNP genotyping (pre-dose)		X									5.5.2
Whole blood for gene expression (PaxGene-RNA tubes)		X	X								5.5.2
Myeloid-derived suppressor cells (Cyto-Chex tube)		X									5.5.2
PBMCs		X	X								5.5.2
Tumor evaluation (CT or MRI) (RECIST 1.1)[n,o]	X	Every 6 weeks for the first 48 weeks relative to the date of randomization, and then every 8 weeks thereafter									5.1
PGx sample (optional DNA element for long-term storage/future use)	X										5.6
Health economics measurements											
Hospital resource use module (HOSPAD)[p]		To be completed at each hospitalization and unscheduled visit by site staff									8.5.8

[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization. For patients undergoing retreatment, if clinically feasible, a new biopsy should be obtained.
[b] Body weight is recorded along with vital signs.
[c] Any clinically significant abnormalities detected require a confirmatory ECG.
[d] If screening laboratory assessments are performed within 3 days prior to Day 1 (first infusion day), they do not need to be repeated at Day 1.
[e] Free T_3 and free T_4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
[f] For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
[g] Up to 60 minutes pre-dose and within 10 minutes of the end of infusion.
[h] To be administered in line with local practice for patients with non-squamous tumors who will receive pemetrexed if randomized to SoC arm.

[i] During the combination portion of treatment, tremelimumab will be administered first; the MEDI4736 infusion will start approximately 1 hour (maximum 2 hours) after the end of the tremelimumab infusion. If there are no clinically significant infusion reactions with the first cycle, then for all other cycles, the MEDI4736 can be given immediately after the tremelimumab infusion has finished.

[j] Results for urea and electrolytes, full blood count, and liver function tests must be available before commencing an infusion (within 3 days).

[k] Patients will complete PROs using handheld devices at home.

[l] PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.

[m] For monotherapy arm only.

[n] RECIST assessments will be performed using CT/MRI assessments of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients at baseline and follow-up. Baseline assessments should be performed no more than 28 days before the date of randomization and, ideally, should be performed as close as possible to the start of study treatment. The confirmatory scans should be performed no less than 4 weeks after the initial assessment of PD (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their next scheduled visit.

[o] Patients with confirmed PD who continue to receive MEDI4736 + tremelimumab or MEDI4736 at the discretion of the Investigator (following consultation with AstraZeneca) can receive treatment for a maximum of 12 months. Patients will have scans done every 6 weeks for the first 48 weeks, and then every 8 weeks thereafter (relative to the date of randomization) until confirmed objective disease progression.

[p] HOSPAD module should be completed by site staff whenever the patient has attended or been admitted in to the hospital and at unscheduled visits. A reminder will be provided at each clinic visit.

Note: All assessments on treatment days are to be performed prior to infusion, unless otherwise indicated.

C Cycle; IM intramuscular; PGIC Patients' Global Impression of Change; SNP Single nucleotide polymorphism; TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine.

Table 3 Schedule of assessments for Standard of Care therapy treatment period

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Informed consent													
Informed consent: study procedures	X[a]												4.1, 10.4
Consent: genetic sample and analysis (optional)	X												3.3
Study procedures													
Physical exam (full)	X												5.2.2
Targeted physical exam (based on symptoms)		X	X	X	X	X	X	X	X	X	X	X	5.2.2
Vital signs[b]	X	X	X	X	X	X	X	X	X	X	X	X	5.2.4
ECG[c]	X	As clinically indicated											5.2.3
Concomitant medications	<----------											---------->	7.7
Demography, including baseline characteristics and tobacco use	X												4.1
Eligibility criteria	X												3.1, 3.2
Laboratory assessments													
Clinical chemistry[d,e]	X	X	X	X	X	X	X	X	X	X	X	X	Table 5
Hematology[d,e]	X	X	X	X	X	X	X	X	X	X	X	X	Table 6
TSH, free T_3, and free T_4[f]	X					X				X			5.2.1
Urinalysis	X												Table 7
Hepatitis B and C and HIV	X												5.2.1

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Pregnancy test[g]	X	As clinically indicated											5.2.1
Monitoring													
WHO/ECOG performance status	X	X	X	X	X	X	X	X	X	X	X	X	5.3.3
AE/SAE assessment	<----------											---------->	6.2.1
Pre-randomization medication													
Folic acid[h]	X	Continue in line with local practice											3.3
IM Vitamin B12[h]	X	Continue in line with local practice											
SoC administration													
Platinum-based chemotherapy		X	Cycle every 3 weeks										7.2.1
PRO assessments[i]													
EORTC QLQ-C30, EQ-5D-5L		Every 4 weeks relative for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)											5.3.1.1, 5.3.1.5
EORTC QLQ-LC13, PRO-CTCAE[j]		Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)											5.3.1.2, 5.3.1.3
PGIC		Weeks 4, 8, 12, 24, and 48											5.3.1.4
Other laboratory assessments and assays													
sPD-L1 (serum)		X				X				X			5.5.2
Circulating soluble factors		X		X		X							5.5.2
Tumor biopsy (newly acquired or archived <3 months old)	X[a]												5.5.1
Whole blood for SNP genotyping		X											5.5.2

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**±1**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Whole blood gene expression (PaxGene-RNA tubes)		X		X									5.5.2
Myeloid-derived suppressor cells (Cyto-Chex tube)		X											5.5.2
PBMCs		X		X									5.5.2
Tumor evaluation (CT or MRI) (RECIST 1.1)[k]	X	Every 6 weeks for the first 48 weeks relative to the date of randomization, and then every 8 weeks thereafter											5.1
PGx sample (optional DNA element for long-term storage/future use)	X												5.6
Health economics measurements													
Hospital resource use module (HOSPAD)[l]		To be completed at each hospitalization and unscheduled visit by site staff											8.5.8

[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization.
[b] Pre-dose and as clinically indicated before every infusion or administration.
[c] Any clinically significant abnormalities detected require a confirmatory ECG.
[d] To be collected every 3 weeks prior to the start of infusion and as clinically indicated.
[e] If screening laboratory assessments are performed within 3 days prior to Day 1 (first infusion day), they do not need to be repeated at Day 1.
[f] Free T_3 and free T_4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
[g] For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
[h] To be administered in line with local practice for patients with non-squamous tumors who will receive pemetrexed if randomized to SoC arm.
[i] Patients will complete PROs using handheld devices at home.
[j] PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.
[k] RECIST assessments will be performed using CT/MRI assessments of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients at baseline and follow-up. Baseline assessments should be performed no more than 28 days before the date of randomization and, ideally, should be performed as close as possible to the start of study treatment. The confirmatory scans should be performed no less than 4 weeks after the initial assessment of PD (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their next scheduled visit.

[l] HOSPAD module should be completed by site staff whenever the patient has attended or been admitted in to the hospital and at unscheduled visits. A reminder will be provided at each clinic visit.

Note: All assessments on treatment days are to be performed prior to infusion or administration, unless otherwise indicated.

C Cycle; IM intramuscular; PGIC Patients' Global Impression of Change; SNP Single nucleotide polymorphism; TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine.

Table 4 Schedule of assessments for patients who have completed/discontinued treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or Standard of Care therapy

	Time since last dose of IP							
	Day (±3)	**Months (±1 week)**						**12 months and every 6 months (±2 weeks)**
Evaluation	**30**	**2**	**3**	**4**	**6**	**8**	**10**	
Physical examination (full)[a]	X							
Vital signs (temperature, respiratory rate, blood pressure, and pulse)	X							
Weight	X							
Pregnancy test[b]	X	As clinically indicated						
AE/SAE assessment	X	X	X					
Concomitant medications	X	X	X					
WHO/ECOG performance status	At timepoints consistent with tumor assessments; at 30, 60, and 90 days following confirmation of progression; and then at initiation of subsequent anticancer therapy[c]							
Subsequent anticancer therapy[d]; and second progression assessment[e]	<---------->							
Survival status[f]		X	X	X	X	X	X	X (every 2 months)
Hematology	X	X	X					
Clinical chemistry	X	X	X					
TSH, free T_3, and free T_4[g]	X							
Pharmacokinetic assessment[h]			X					
Immunogenicity assessment (ADA sampling) to identify ADA responses in patient circulation[h]			X		X			
sPD-L1 concentration (to assess target engagement)[h]			X					
EORTC QLQ-C30[i], EQ-5D-5L[i]	Every 4 weeks relative for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)							

	Time since last dose of IP							
	Day (±3)	Months (±1 week)						12 months and every 6 months (±2 weeks)
Evaluation	**30**	**2**	**3**	**4**	**6**	**8**	**10**	
EORTC QLQ-LC13[i], PRO-CTCAE[i,j]	Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)							
Hospital resource use module (HOSPAD)	X							
Tumor assessment (CT or MRI)[k]	Every 6 weeks for the first 48 weeks (relative to the date of randomization), and then every 8 weeks thereafter until confirmed objective disease progression per standard practice post-progression/death (whichever comes first)							

[a] Physical exams are described in Section 5.2.2.
[b] For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
[c] WHO/ECOG performance status should also be collected at other site visits that the patient attends, if appropriate site staff are available to collect such information. In addition, WHO performance status should be provided when information on subsequent anticancer therapy is provided, where possible.
[d] Details of any treatment for NSCLC (including surgery) post the last dose of study treatment must be recorded in the eCRF.
[e] Second disease progression (PFS2) assessment will be performed by the Investigator and defined according to local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression, or death.
[f] Patients may be contacted in the week following data cutoffs to confirm survival status. Details of any treatment for NSCLC (including surgery) post the last dose of study treatment must be recorded in the eCRF.
[g] Free T3 and free T4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
[h] For patients in the MEDI4736 + tremelimumab or MEDI4736 monotherapy groups only.
[i] Patients will complete PROs using handheld devices at home.
[j] PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.
[k] Only for patients yet to progress, RECIST 1.1 assessments will be performed using CT/MRI assessments of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients. The confirmatory scans should preferably be performed no less than 4 weeks after the initial assessment of PD (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits (relative to the date of randomization). All confirmatory scans should be recorded on the database.

TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine.

4.1 Enrollment/screening period

All screening and enrolment procedures will be performed according to the assessment schedule in Table 2 and Table 3. Demographic data and other characteristics will be recorded including date of birth or age, gender, smoking history, and race/ethnicity, according to local regulations. A standard medical and surgical history will be obtained.

Written informed consent and any locally required privacy act document authorization must be obtained prior to performing any protocol-specific procedures, including screening/baseline evaluations. All patients will be required to provide consent to supply a sample of their tumor (archived or newly acquired biopsy) for entry into this study. This consent is included in the main patient informed consent form (ICF).

Screening/baseline evaluations may be performed over more than 1 visit.

The timing of vital sign assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 2 and Table 3.

4.2 Treatment period

All procedures to be conducted during the treatment period will be performed according to the assessment schedule (see Table 2 and Table 3).

Whenever vital signs and blood draws are scheduled for the same nominal time, the assessments should occur in the following order: vital signs and then blood draws. The timing of the vital signs assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 2 and Table 3.

4.3 Follow-up period

All procedures to be conducted during the follow-up period will be performed according to the assessment schedule (see Table 4).

Whenever vital signs, ECGs, and blood draws are scheduled for the same nominal time, the assessments should occur in the following order: ECG, vital signs, and then blood draws. The timing of the first 2 assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 4.

5. STUDY ASSESSMENTS

A Web-Based Data Capture (WBDC) system will be used for data collection and query handling. The Investigator will ensure that data are recorded on the eCRFs as specified in this study protocol and in accordance with the instructions provided.

The Investigator ensures the accuracy, completeness, and timeliness of the data recorded and the provision of answers to data queries according to the clinical study agreement (CSA). The Investigator will sign the completed eCRFs. A copy of the completed eCRFs will be archived at the study site.

5.1 Efficacy assessments

RECIST 1.1 criteria will be used to assess patient response to treatment by determining PFS, ORR, DoR, and APF12 using investigator assessments (primary), and a BICR will be performed on a sample of patients (for sensitivity analysis). The RECIST 1.1 guidelines for measurable, non-measurable, target, and non-target lesions and the objective tumor response criteria (CR, PR, SD, or PD) are presented in Appendix F. PFS2 defined by local standard clinical practice, and OS will also be evaluated.

The methods of assessment of tumor burden used at baseline are CT and MRI scans of the chest and abdomen (including liver and adrenal glands). Any other areas of disease involvement should be additionally imaged based on the signs and symptoms of individual patients.

The baseline assessment should be performed no more than 28 days before the start of IP treatment and ideally as close as possible to the start of the assigned IP. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; Table 2, Table 3, and Table 4) then q8w thereafter, until confirmed objective disease progression per RECIST 1.1 (irrespective of the reason for stopping treatment or subsequent therapy). 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.

For patients who discontinue IP (including SoC) due to toxicity in the absence of confirmed objective progression, objective tumor assessments per the scheduled assessments should be continued every 6 weeks for 48 weeks (relative to randomization) and then every 8 weeks until confirmed objective disease progression.

A confirmatory scan is required for all patients following the initial demonstration of PD. The confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD in the absence of clinically significant deterioration. Treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC may continue between the initial assessment of progression and confirmation of progression. Progression would be considered confirmed per RECIST 1.1 criteria available in Appendix F using Investigator assessments.

If a patient discontinues treatment (and/or receives a subsequent anticancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

Categorization of objective tumor response assessment will be based on the RECIST 1.1 criteria of response: CR, PR, SD, and PD. Target lesion progression will be calculated in comparison to when the tumor burden was at a minimum (ie, smallest sum of diameters previously recorded on study). In the absence of progression, tumor response (CR or PR) and SD will be calculated in comparison to the baseline tumor measurements obtained before starting treatment.

Following confirmed progression, patients should continue to be followed up for survival every 1 to 2 months as outlined in the study plan (Table 4). Exceptions are patients with confirmed PD who continue to receive IP at the discretion of the Investigator (after consultation with AstraZeneca); these patients can receive treatment for a maximum of 12 months and will have scans for RECIST 1.1 assessments every 6 weeks (relative to the date of randomization per Table 2 and Table 3) for the first 48 weeks of treatment and then q8w until disease progression. Subsequent anticancer therapy information will be collected at the timepoints indicated in Table 4.

Patients in the MEDI4736 + tremelimumab or MEDI4736 monotherapy groups who will receive retreatment must have a baseline tumor assessment within 28 days of restarting treatment and additional scans every 6 weeks for the first 48 weeks relative to the date of randomization, and then q8w thereafter until disease progression. All assessments in Table 2 will be followed for patients who receive retreatment.

It is important to follow the assessment schedule as closely as possible. Please refer to the study plans (Table 2 and Table 3 [screening and the treatment period] and Table 4 [for follow-up of patients who have completed or discontinued IP treatment]) and Appendix F.

5.1.1 Central reading of scans

A BICR of radiological scans will be performed on a random sample of patients to confirm the robustness of the PFS endpoint. This will be prespecified in the statistical analysis plan/independent review charter. If bias cannot be excluded based upon the sample BICR assessment, then an independent evaluation of all radiographic images may be required for the assessment of the primary PFS endpoint.

All images will be collected centrally. Guidelines for imaging collection and storage will be provided in a separate document. The management of patients will be based solely upon the results of the RECIST assessment conducted by the Investigator.

5.1.2 Survival assessments

Assessments for survival must be made every 2 months following treatment discontinuation. Survival information may be obtained via telephone contact with the patient, patient's family,

or by contact with the patient's current physician. The details of first and subsequent therapies for cancer, after discontinuation of treatment, will be collected.

In addition, patients on treatment or in survival follow-up will be contacted following the data cut-off for the primary analysis and all subsequent survival analyses to provide complete survival data. These contacts should generally occur within 7 days of the data cut off.

5.2 Safety assessments

5.2.1 Laboratory safety assessments

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 (see Table 2, Table 3, and Table 4).

Clinical laboratory safety tests, including serum pregnancy tests, will be performed in a licensed clinical laboratory according to local standard procedures. Sample tubes and sample sizes may vary depending on the laboratory method used and routine practice at the site. Urine pregnancy tests may be performed at the site using a licensed test (urine or serum pregnancy test). Abnormal clinically significant laboratory results should be repeated as soon as possible (preferably within 24 to 48 hours).

Additional safety samples may be collected if clinically indicated at the discretion of the Investigator. The date, time of collection, and results (values, units, and reference ranges) will be recorded on the appropriate eCRF.

The laboratory variables to be measured are presented in Table 5 (clinical chemistry), Table 6 (hematology), and Table 7 (urinalysis).

Pregnancy tests on either urine (human chorionic gonadotropin [hCG]) or blood (serum β-hCG) samples will be performed for pre-menopausal women of childbearing potential at Screening as specified in the assessment schedule (see Table 2, Table 3, and Table 4). Tests will be performed by the hospital's local laboratory. If results are positive, the patient must not start or continue treatment. In the event of a suspected pregnancy during the study, the test should be repeated.

Other safety tests to be performed at Screening include assessment for hepatitis B surface antigen, hepatitis C antibodies, HIV antibodies, thyroid-stimulating hormone (TSH), free triiodothyronine (T_3), and free thyroxine (T_4).

Table 5 Clinical chemistry (serum or plasma)

Albumin	Glucose
Alkaline phosphatase	Lactate dehydrogenase
Alanine aminotransferase	Lipase
Amylase	Magnesium[c]
Aspartate aminotransferase	Potassium
Bicarbonate	Sodium
Calcium	Total bilirubin[a]
Chloride	Total protein
Creatinine	Urea or blood urea nitrogen, depending on local practice
Gamma glutamyltransferase[b]	Uric acid

[a] If total bilirubin is ≥2 × ULN (and evidence of Gilbert's syndrome) then fractionate into direct and indirect bilirubin.
[b] At baseline and as clinically indicated.

Table 6 Hematology

Basophils	Mean corpuscular volume
Eosinophils	Monocytes
Hematocrit	Neutrophils
Hemoglobin	Platelet count
Lymphocytes	Red blood cell count
Mean corpuscular hemoglobin	Total white cell count
Mean corpuscular hemoglobin concentration	

Note: For coagulation parameters, activated partial thromboplastin time and international normalized ratio are to be assessed at baseline and as clinically indicated.

Table 7 Urinalysis

Bilirubin	Ketones
Blood	pH
Color and appearance	Protein
Glucose	Specific gravity

Note: Microscopy should be used as appropriate to investigate white blood cells and use the high power field for red blood cells.

If a patient shows an AST or ALT ≥3 × ULN together with total bilirubin ≥2 × ULN, refer to Appendix E for further instructions. These cases should be reported as SAEs if, after evaluation, they meet the criteria for a Hy's law case or if any of the individual liver test parameters fulfill any of the SAE criteria. All patients with an elevated AST, ALT, or

bilirubin value (the latter at ≥1.5 × ULN) at the time of the last dose of study treatment should have a further liver chemistry profile (AST, ALT, bilirubin, and alkaline phosphatase) performed 30 days (±3 days) after permanent discontinuation of study treatment.

Any clinically significant abnormal laboratory values should be repeated as clinically indicated and recorded on the eCRF. Situations in which laboratory safety results should be reported as AEs are described in Section 6.3.4.

All patients with Grade 3 or 4 laboratory values at the time of completion or discontinuation from study treatment must have further tests performed until the laboratory values have returned to Grade 1 or 2, unless these values are not likely to improve because of the underlying disease.

5.2.2 Physical examination

Physical examinations will be performed according to the assessment schedules (see Table 2, Table 3, and Table 4). Full physical examinations will include assessments of the head, eyes, ears, nose, and throat and the respiratory, cardiovascular, GI, urogenital, musculoskeletal, neurological, dermatological, hematologic/lymphatic, and endocrine systems. Height will be measured at Screening only. Targeted physical examinations are to be utilized by the Investigator on the basis of clinical observations and symptomatology. Situations in which physical examination results should be reported as AEs are described in Section 6.3.4.

5.2.3 Electrocardiograms

Resting 12-lead ECGs will be recorded at screening and as clinically indicated throughout the study (see Table 2 and Table 3). ECGs should be obtained after the patient has been in a supine position for 5 minutes and recorded while the patient remains in that position.

At Screening, a single ECG will be obtained on which QTcF must be <470 ms.

In case of clinically significant ECG abnormalities, including a QTcF value >470 ms, 2 additional 12-lead ECGs should be obtained over a brief period (eg, 30 minutes) to confirm the finding.

Situations in which ECG results should be reported as AEs are described in Section 6.3.4.

5.2.4 Vital signs

Vital signs (blood pressure [BP], pulse, temperature, and respiration rate) will be evaluated according to the assessment schedules (see Table 2, Table 3, and Table 4).

On infusion days, patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups will be monitored during and after infusion of IP as presented in the bulleted list below. Patients in the SoC group will be monitored pre-dose and as clinically indicated before every infusion or administration.

Supine BP will be measured using a semi-automatic BP recording device with an appropriate cuff size, after the patient has rested for at least 5 minutes. BP and pulse will be collected from patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups before, during, and after each infusion at the following times (based on a 60-minute infusion):

- Prior to the beginning of the infusion (measured once from approximately 30 minutes before up to 0 minutes [ie, the beginning of the infusion])
- Approximately 30 minutes during the infusion (**halfway** through infusion)
- At the end of the infusion (approximately 60 minutes ±5 minutes)
- A 1-hour observation period is recommended after the first infusion of MEDI4736 and tremelimumab. If no clinically significant infusion reactions are observed during or after the first cycle, subsequent infusion observation periods can be at the Investigator's discretion (suggested 30 minutes after each MEDI4736 and tremelimumab infusion).

If the infusion takes longer than 60 minutes, then BP and pulse measurements should follow the principles as described above or be taken more frequently if clinically indicated. The date and time of collection and measurement will be recorded on the appropriate eCRF. Additional monitoring with assessment of vital signs is at the discretion of the Investigator per standard clinical practice or as clinically indicated.

Body weight is also recorded along with vital signs.

Situations in which vital signs results should be reported as AEs are described in Section 6.3.4.

5.3 Other assessments

5.3.1 Patient-reported outcomes

"PRO" is an umbrella term referring to all outcomes and symptoms that are directly reported by the patient. PROs have become a significant endpoint when evaluating effectiveness of treatments in clinical trials. The following PROs will be administered in this study: EORTC QLQ-C30 (core questionnaire), QLQ-LC13 (lung cancer module), PRO-CTCAE, PGIC, and the EQ-5D-5L (see Appendix G).

The PRO instruments will be completed by the patients using a handheld ePRO device. All assessments should be completed without assistance from anyone according to the assessment schedules (see Table 2, Table 3, and Table 4). It takes approximately 30 to 45 minutes for patients to complete the questionnaires; therefore, the burden to the patient is moderate.

5.3.1.1 EORTC QLQ-C30

The EORTC QLQ-C30 v3 questionnaire is included for the purpose of assessing HRQoL and is a well-established measure of HRQoL/health status, and commonly used as an endpoint in cancer clinical trials. It assesses HRQoL/health status through 9 multi-item scales: 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea and vomiting), and a global health and QoL scale. Six single-item symptom measures are also included: dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties (see Appendix G). For each of the 15 domains, final scores are transformed such that they range from 0 to 100, where higher scores indicate greater functioning, greater HRQoL, or greater level of symptoms (Aaronson et al 1993).

5.3.1.2 EORTC QLQ-LC13

For patients with NSCLC, a disease-specific 13-item self-administered questionnaire for lung cancer was developed (QLQ-LC13; Appendix G) to be used in conjunction with the EORTC QLQ-C30 (Bergman et al 1994). It comprises both multi-item and single-item measures of lung cancer-associated symptoms (ie, coughing, hemoptysis, dyspnea, and pain) and side effects from conventional chemotherapy and radiotherapy (ie, hair loss, neuropathy, sore mouth, and dysphagia). Similar to the EORTC QLQ-C30, all questions except 1 have a 4-point scale: "Not at all," "A little," "Quite a bit," and "Very much." One question (#43 "Did you take any medicine for pain?") has a response option of "Yes" or "No." The scoring approach for the QLQ-LC13 is similar to the EORTC QLQ-C30.

5.3.1.3 PRO-CTCAE

The PRO-CTCAE is included to address tolerability from the patients' perspective. It was developed by the NCI. The PRO-CTCAE will only be administered in those countries where a linguistically validated version exists. It was developed in recognition that collecting symptom data directly from patients using PRO tools can improve the accuracy and efficiency of symptomatic AE data collection. This was based on findings from multiple studies demonstrating that physicians and nurses underestimate symptom onset, frequency, and severity in comparison with patient ratings (Antonia et al 2014b, Litwin et al 1998, Sprangers and Aaronson 1992). These symptoms have been converted to patient terms (eg, the CTCAE term "myalgia" has been converted to "aching muscles"). For several symptoms, like fatigue and pain, additional questions are asked about symptom frequency, severity, and interference with usual activities. The items included in the PRO-CTCAE have undergone extensive qualitative review among experts and patients. These items have been extensively evaluated by cancer patients to be clear, comprehendible, and measure the symptom of interest. In this study, only items that are considered relevant for the trial, site of cancer, and cancer treatment are selected (see Appendix G).

5.3.1.4 Patients' Global Impression of Change

The PGIC item is included to assess how a patient perceives his/her overall change in health status since the start of study treatment. Patients will choose from response options from "Very Much Improved" to "Very Much Worse."

5.3.1.5 EQ-5D-5L

The EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal (EuroQol Group 1990). Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care as well as in population health surveys. The questionnaire assesses 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 response options (no problems, slight problems, moderate problems, severe problems, and extreme problems) that reflect increasing levels of difficulty (EuroQol Group 2013).

Since 2009, the EuroQol group has been developing a more sensitive version of the EQ-5D (the EQ-5D-5L) which expands the range of responses to each dimension from 3 to 5 levels of increasing severity (Herdman et al 2011). Preliminary studies indicate that the 5L version improves upon the properties of the 3L measure in terms of reduced ceiling effect, increased reliability and an improved ability to differentiate between different levels of health (Janssen et al 2008a; Janssen et al 2008b; Pickard et al 2007).

The patient will be asked to indicate his/her current health state by selecting the most appropriate level in each of the 5 dimensions. The questionnaire also includes a visual analogue scale, where the patient will be asked to rate current health status on a scale of 0 to 100, with 0 being the worst imaginable health state (see Appendix G).

5.3.2 Administration of the patient-reported outcome questionnaires

Patients will complete the PRO assessments by using a handheld electronic device (ePRO).

Each center must allocate the responsibility for the administration of the PRO devices to a specific individual (eg, a research nurse or study coordinator) and, if possible, assign a back-up person to cover if that individual is absent. The PRO questionnaires must be completed per the schedule of assessments (see Table 2, Table 3, and Table 4). Patients will be instructed to bring their handheld devices to every clinic visit. In instances where the PRO assessments coincide with a clinic visit, the research nurse or study coordinator should verify that the questionnaire was completed before the visit. If not, the patient should complete the questionnaire prior to any other study procedures and before discussion of disease progression to avoid biasing the patient's responses to the questions.

The following best practice guidelines should be followed when collecting PRO data via an electronic device:

- The research nurse or appointed site staff must explain the value and relevance of participation to patients and inform them that these questions are being asked in order to find out from them directly how they feel. The research nurse or appointed site staff should also stress that the information is confidential. Therefore, if the patient has any medical problems, he/she should discuss them with the doctor or research nurse separately from the ePRO assessment.

- The research nurse or appointed site staff must train the patient on how to use the ePRO device using the materials and training provided by the ePRO vendor, and also provide guidance on whom to call if there are problems with the device.

- The research nurse or appointed site staff must remind patients that there are no right or wrong answers and avoid introducing bias by not clarifying items. The patient should not receive help from relatives, friends, or clinic staff to answer the PRO questionnaires.

- The research nurse or appointed site staff must monitor compliance; minimizing missing data is a key aspect of study success. Compliance must be checked at each study visit and should be checked more frequently to identify problems early. If compliance drops below 85%, a check-in call from the site to ask the patient if he/she has any difficulties is highly recommended.

5.3.3 WHO/ECOG performance status

WHO/ECOG performance status will be assessed at the times specified in the assessment schedules (see Table 2, Table 3, and Table 4) based on the following:

0. Fully active; able to carry out all usual activities without restrictions

1. Restricted in strenuous activity, but ambulatory and able to carry out light work or work of a sedentary nature (eg, light housework or office work)

2. Ambulatory and capable of self-care, but unable to carry out any work activities; up and about more than 50% of waking hours.

3. Capable of only limited self-care; confined to bed or chair more than 50% of waking hours

4. Completely disabled; unable to carry out any self-care and totally confined to bed or chair

Any significant changes from baseline or screening must be reported as an AE.

5.4 Pharmacokinetics

5.4.1 Collection of samples and determination of drug concentration

Blood samples for determination of MEDI4736 and tremelimumab concentration in serum will be obtained according to the assessment schedules (see Table 2 and Table 4).

Samples for determination of MEDI4736 and tremelimumab concentration in serum will be analyzed by a designated third party on behalf of AstraZeneca. Samples will be collected, labeled, stored, and shipped as detailed in the Laboratory Manual. Full details of the analytical method used will be described in a separate Bioanalytical Validation Report.

5.4.2 Collection of samples to measure for the presence of ADAs

The presence of ADA will be assessed in serum samples taken according to the assessment schedules (see Table 2 and Table 4).

Samples will be measured for the presence of ADAs and ADA-neutralizing antibodies for both IPs (MEDI4736 and tremelimumab) using validated assays. Tiered analysis will be performed to include screening, confirmatory, and titer assay components, and positive-negative cut points previously statistically determined from drug-naïve validation samples will be employed.

5.4.3 Storage and destruction of pharmacokinetic/ADA samples

PK and ADA samples will be disposed of a maximum of 10 years after the IPs are approved for marketing.

PK and ADA samples may be disposed of or destroyed and anonymized by pooling. Additional analyses may be conducted on the anonymized, pooled PK samples to further evaluate and validate the analytical method. Results from such analyses may be reported separately from the clinical study report (CSR).

Incurred sample reproducibility analysis, if any, will be performed alongside the bioanalysis of the test samples. The results from the evaluation will not be reported in the CSR but separately in a Bioanalytical Validation Report.

Any residual back-up PK samples may be used for future exploratory biomarker research (in this case, residual back-up PK samples will be shipped to AstraZeneca Biobank; see details in the Laboratory Manual).

5.5 Biomarker analysis

The patient's consent to the use of donated biological samples is mandatory. Tissue samples will be obtained from all screened patients.

Pre-treatment tumor PD-L1 expression will be evaluated in all randomized patients. Data will be compared between arms to determine if baseline PD-L1 status is prognostic and/or predictive of outcomes associated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus SoC. Baseline tumor requirements are briefly described in Section 5.5.1.

Based on availability of tissue, additional exploratory biomarkers may be evaluated as described in Section 5.5.3. Also, descriptions of exploratory, peripheral measures are described in this section. Samples will be obtained according to the assessment schedules provided in Table 2, Table 3, and Table 4.

Details for collection, volumes, storage, and shipment of biologic samples are presented in a separate Laboratory Manual.

All samples collected for biomarker analyses will be stored at the study site, a reference laboratory, or at AstraZeneca facilities and may be used for subsequent research relevant to evaluating biological and/or clinical response to immunotherapy.

The results may be pooled with biomarker data from other MEDI4736 and tremelimumab studies to evaluate biological responses across indications and to compare results in monotherapy versus combination settings.

5.5.1 Collection of patient samples for stratification by PD-L1

At screening, there are 2 mandatory options for provisions of tissue to be used for determination of eligibility. There is 1 subsequent mandatory provision of tissue at progression if retreatment is planned :

- MANDATORY: Provision of a recent tumor biopsy formalin fixed and embedded in paraffin. A freshly collected tumor biopsy is strongly preferred; however, if not clinically feasible, an archival sample taken less than 3 months prior to screening may be submitted.

  Samples should be collected via an image-guided core needle (18 gauge or larger) or be collected as an excisional or incisional tumor biopsy sample

  When tissue is newly obtained for the purpose of entry into this study, 2 cores should be placed in formalin and processed to a single paraffin embedded block, as described in the Laboratory Manual.

  The tumor specimen submitted to establish eligibility should be of sufficient quantity to allow for PD-L1 immunohistochemistry analyses (see the Laboratory Manual). Newly acquired or archived specimens with limited tumor content and fine needle aspirates are inadequate for defining tumor PD-L1 status .

  Tumor lesions used for fresh biopsies should not be the same lesions used as RECIST 1.1 target lesions, unless there are no other lesions suitable for biopsy. If a RECIST 1.1 target lesion is used for biopsy, the lesion must be ≥2 cm in the longest diameter and must be biopsied outside of the screening period.

- MANDATORY: The collection of additional archived tumor tissue block greater than 3 months old (formalin-fixed paraffin-embedded) is mandated, where such samples exist in a quantity sufficient to allow for analysis. Tumor tissue block is preferred. If a tissue block is unavailable, unstained sections from the tissue block may be submitted. Please consult the laboratory manual for specific instructions and guidelines regarding sections.

- MANDATORY: The collection of tumor biopsies at the time of progression prior to retreatment is mandated. The Investigator must consult with the Study Physician if such sampling is not feasible.

- OPTIONAL: The collection of additional biopsies upon progression of patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy arms is strongly encouraged.

  Additional tumor biopsies collected as part of clinical care (eg, for mixed responses or upon PD) can be submitted for further analysis.

See the Laboratory Manual for further details of requirements including sample QC and shipping.

A brief description of exploratory tumor markers likely to be explored by IHC or RNA analysis is provided in Section 5.5.2.

The PD-L1 IHC assay will be used to determine PD-L1 IHC status in this study for the purposes of stratification and for the analysis of the original diagnostic sample.

To meet the requirement of FDA approval of a companion diagnostic, sections of the tumor will be retained at for potential additional studies, as requested by the FDA, to support potential test approval.

5.5.2 Exploratory biomarkers

Blood and tumor samples for exploratory biomarker analyses will be obtained according to the schedules presented in Table 2, Table 3, and Table 4. Details for collection, volumes, storage, and shipment of biologic samples are presented in a separate Laboratory Manual.

Pharmacodynamic changes in biomarker measures will be monitored, when applicable. Baseline measures (and early, on-treatment changes) will be correlated with outcomes. Note that samples will be obtained from patients randomized to each treatment arm. Comparisons will be made between baseline measures to determine if biomarkers (or combination of markers) are prognostic or predictive of outcomes associated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus SoC.

Additional sample collections and analyses may be completed at select study sites by site-specific amendment. All samples collected for such exploratory analyses will be stored at site, a reference laboratory, or at AstraZeneca's facilities and may be used for subsequent research relevant to evaluating response to immunotherapy.

The exploratory biomarker plan is described by sample type below.

Whole blood for DNA/SNP genotyping

Genomic DNA will be extracted from whole blood obtained pre-treatment from all patients. Genotyping of immunomodulatory genes such as, but not limited to, PD-1, PD-L1, CTLA-4, and human leukocyte antigen loci may be completed to determine if natural variation within such genes is associated with likelihood of clinical benefit (and/or with likelihood of drug-related AEs). Genes associated with NSCLC development, progression, or likelihood of

response to chemotherapy may likewise be investigated. Genotyping will occur retrospectively, data will not be shared with patients, and results will not impact treatment decisions.

Genotypes also may be correlated with biomarker measures (eg, gene and/or protein expression) obtained from other sample types described in this exploratory biomarker section. A primary hypothesis is that different genotypes will be associated with different expression levels of factors within the PD-1 and CTLA-4 signaling pathways. Such variations in expression may affect the ability of an individual to mount an appropriate immune reaction to tumor and/or affect the likelihood of response to therapeutics targeting these pathways. Therefore, genotyping may provide easy-to-measure, baseline information regarding a patient's immune system, and a goal of this research is to understand how such genetic information may be used to predict pharmacodynamic responses to therapy.

Whole blood gene expression (PaxGene-RNA)

Whole blood samples will be obtained pre- or post-treatment from all patients as described in Table 2 or Table 3. Total RNA will be prepared for quantification of RNA and/or miRNA expression using reverse transcription quantitative polymerase chain reaction (RT-QPCR), microarray, sequencing, or similar technology.

Focus is likely to be given to the expression of immunomodulatory genes previously found to be up-regulated in response to MEDI4736 and/or tremelimumab (data not shown). This battery of genes is likely to be similarly up-regulated in response to MEDI4736 + tremelimumab combination therapy (but not in response to SoC). Moreover, pre-treatment expression of such genes may indicate active immune responses that may be augmented by checkpoint inhibitor immunotherapies; correlations with outcome data will be completed on select candidate, predictive markers with the aim of identifying useful expression thresholds for identifying patients likely to receive benefit.

Myeloid-derived suppressor cells

Recent, collective findings suggest that a baseline measure of circulating MDSCs may be used as a prognostic tool in different disease settings and may specifically predict the likelihood of response to ipilimumab (anti-CTLA-4 therapy) (Kitano et al 2014, Meyer et al 2014). Flow cytometry will be completed on all patients to quantify pre-treatment, circulating MDSC subtypes. Different MDSC count or percentage thresholds will be analyzed for their ability to predict clinical benefit from MEDI4736 + tremelimumab combination, MEDI4736 monotherapy, or SoC therapies.

Peripheral blood mononuclear cells

Whole blood samples will be collected for preparation of PBMCs and storage for potential downstream analyses. A variety of assays may be pursued, including but not limited to: immune cell composition/activation status analyses by flow cytometry, T cell functional assays (eg, Enzyme-Linked ImmunoSpot, receptor occupancy analyses to measure target engagement, tetramer analyses to monitor antigen-specific T cells, RNA/miRNA expression,

and/or the assessment of the diversity and clonality of T cell receptor gene rearrangements using DNA.

Soluble factors - plasma

Plasma will be obtained pre- or post-treatment from all patients as described in Table 2, Table 3, and Table 4. The concentrations of a panel of cytokines and chemokines will be assessed. Focus is likely to be given to factors involved in T helper 1 cell-driven immune responses, including but not limited to IFN gamma, IL-18, chemokine (C-X-C motif) ligand (CXCL) 9, and CXCL10. Pharmacodynamic effects will be monitored to determine the specificity of response to immunotherapy (versus SoC). High pre-treatment expressions (concentrations) of such factors may indicate active immune responses, which may be augmented by checkpoint inhibitor immunotherapies; correlations with outcome data will be completed on select candidate, predictive markers with an aim of identifying useful expression thresholds for identifying patients likely to receive benefit, or alternatively, for identifying patients likely to suffer drug-related AEs.

Similarly, the concentrations of a battery of immune cell ligands or receptors may be assessed. Proteins of special interest include, but are not limited to, CTLA-4, PD-1, PD-L1, B7-1, B7-2, and IL-6R.

Focus may be given to circulating miRNAs currently thought to be putative, non-invasive prognostic biomarkers for cancer (Schwarzenbach et al 2014) and to subsets within which may be particularly useful for predicting responses to immunotherapies (Wang et al 2013). Additional, candidate miRNAs of interest may include those which regulate effectors within the PD-1 and CTLA-4 signaling pathways (eg, mIR-513, capable of PD-L1 down-regulation (Gong et al 2009).

Lastly, plasma may also be used for the detection/quantification of autoantibodies (against tumor-associated antigens). Seroconversion following treatment will be used as an indicator of overcoming tolerance. Pre-treatment seropositivity against specific antigens may provide predictive value, particularly when combined with data regarding the presence of antigen-specific T cells (Yuan et al 2011). Therefore, select, candidate autoantibody measures may be evaluated for associations with clinical benefit and for directing PBMC-based research described in Section 5.5.1.

Tumor markers

Tissue obtained as part of screening procedures and for establishing PD-L1 expression status will be analyzed for additional markers by immunohistochemistry. A primary goal is to measure CD8 and CD4/FoxP3 protein expression in an effort to enumerate cytotoxic versus regulatory T cells. Based on availability of tissue, a panel of additional, immune-relevant markers expressed on tumor-infiltrating lymphocytes or on tumor cells may be assessed. Markers of special interest include, but are not limited to, Ox40, GITR, PD-L2, Tim-3, CD137, and Lag 3.

Other tissue-based approaches may be pursued including RT-QPCR and in situ hybridization (eg, for detection of IFN gamma signaling genes such as CXCL9, CXCL10, and IFN gamma itself), and/or somatic mutation detection methodologies.

Management of biomarker data

The biomarker data will have unknown clinical significance. AstraZeneca will not provide biomarker research results to patients, their family members, any insurance company, an employer, clinical study investigator, general physician, or any other third party, unless required to do so by law. The patient's samples will not be used for any purpose other than those described in the study protocol.

Individual patients will not be identified in any report or publication resulting from this work. The data and results of this research may be reviewed with collaborators and published, but neither the patient's name nor any other personal identifiers will appear in any publication or report.

5.5.3 Storage, re-use, and destruction of biological samples

Samples will be stored for a maximum of 15 years from the end of study, after which they will be destroyed. 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. The results of this biomarker research may be pooled with biomarker data from other studies involving MEDI4736 or tremelimumab to generate hypotheses to be tested in future research.

5.5.4 Labeling and shipment of biological samples

The Principal Investigator will ensure that samples are labeled and shipped in accordance with the Laboratory Manual and the Biological Substance, Category B, Regulations (materials containing or suspected to contain infectious substances that do not meet Category A criteria); see Appendix C "IATA 6.2 Guidance Document."

Any samples identified as Infectious Category A materials will not be shipped, and no further samples will be taken from the involved patients unless agreed upon with AstraZeneca and appropriate labeling, shipment, and containment provisions are approved.

5.5.5 Chain of custody of biological samples

A full chain of custody will be maintained for all samples throughout their life cycle.

The Principal Investigator at each center will keep full traceability of collected biological samples from the patients while in storage at the center until shipment or disposal (where appropriate) and will keep documentation of receipt of arrival.

The sample receiver will keep full traceability of the samples while in storage and during use until used or disposed of or until further shipment and will keep documentation of receipt of arrival.

AstraZeneca will keep oversight of the entire life cycle through internal procedures, monitoring of study sites, and auditing of external laboratory providers.

Samples retained for further use will be registered in the AstraZeneca Biobank during the entire life cycle.

5.5.6 Withdrawal of informed consent for donated biological samples

If a patient withdraws consent to the use of donated biological samples, the samples will be disposed of or destroyed and the action documented. If samples have already been analyzed, AstraZeneca is not obliged to destroy the results of this research.

The Principal Investigator will:

- Ensure that AstraZeneca is immediately notified of the patients' withdrawal of informed consent to the use of donated samples
- Ensure that biological samples from that patient, if stored at the study site, are immediately identified, disposed of or destroyed and the action documented
- Ensure that the laboratory(ies) holding the samples is/are immediately informed about the withdrawn consent and that samples are disposed of or destroyed, the action is documented, and the signed document is returned to the study site
- Ensure that the patient and AstraZeneca are informed about the sample disposal

5.6 Pharmacogenetics

Refer to Appendix D for details of the genetic research (optional DNA component).

6. SAFETY REPORTING AND MEDICAL MANAGEMENT

The Principal Investigator is responsible for ensuring that all staff involved in the study are familiar with the content of this section.

6.1 Definitions of serious adverse event

An SAE is an AE occurring during any study phase (ie, run-in, treatment, washout, or follow-up) that fulfills one or more of the following criteria:

- Results in death
- Is immediately life-threatening
- Requires in-patient hospitalization or prolongation of existing hospitalization

- Results in persistent or significant disability or incapacity or substantial disruption of the ability to conduct normal life functions
- Is a congenital abnormality or birth defect
- Is an important medical event that may jeopardize the patient or may require medical intervention to prevent one of the outcomes listed above

For further guidance on the definition of an SAE, see Appendix B.

6.2 Recording of adverse events

6.2.1 Time period for collection of adverse events

AEs and SAEs will be collected from the time of signature of informed consent throughout the treatment period and including the follow-up period (90 days after the last dose of IP). AEs and SAEs collected prior to randomization will be reported as pre-randomization AEs and SAEs.

6.2.2 Follow-up of unresolved adverse events

During the course of the study, all AEs and SAEs should be proactively followed up for each patient. Every effort should be made to obtain a resolution for all events, even if the events continue after discontinuation or study completion.

Any AEs that are unresolved at the patient's last visit in the study are followed up by the Investigator for as long as medically indicated, but without further recording in the eCRF. AstraZeneca retains the right to request additional information for any patient with ongoing AE(s)/SAE(s) at the end of the study, if judged necessary.

6.2.3 Variables

The following variables will be collected for each AE:

- AE (verbatim)
- The date and time when the AE started and stopped
- The maximum CTCAE grade reported
- Whether the AE is serious or not
- Investigator causality rating against the IPs (yes or no)
- Action taken with regard to IPs
- Administration of treatment for the AE
- Whether the AE caused the patient's withdrawal from the study (yes or no)

- Outcome

In addition, the following variables will be collected for SAEs:

- Date the AE met criteria for SAE
- Date the Investigator became aware of the SAE
- Seriousness criteria fulfilled
- Date of hospitalization
- Date of discharge
- Probable cause of death
- Date of death
- Whether an autopsy was performed
- Causality assessment in relation to study procedure(s)
- Description of the AE

The grading scales found in the revised NCI CTCAE version 4.03 will be utilized for all events with an assigned CTCAE grading. For those events without assigned CTCAE grades, the recommendation in the CTCAE criteria that converts mild, moderate, and severe events into CTCAE grades should be used. A copy of the CTCAE version 4.03 can be downloaded from the Cancer Therapy Evaluation Program website (http://ctep.cancer.gov).

It is important to distinguish between serious and severe AEs. Severity is a measure of intensity, whereas seriousness is defined by the criteria in Section 6.1. An AE of severe intensity need not necessarily be considered serious. For example, nausea that persists for several hours may be considered severe nausea, but it is not an SAE unless it meets the criteria shown in Section 6.1. On the other hand, a stroke that results in only a limited degree of disability may be considered a mild stroke, but it would be an SAE if it satisfies the criteria shown in Section 6.1.

6.3 Definition of adverse events

An AE is the development of an undesirable medical condition or the deterioration of a pre-existing medical condition during or following exposure to a pharmaceutical product, whether or not the condition is considered to be causally related to the product. An undesirable medical condition can be a symptom (eg, nausea or chest pain), sign (eg, tachycardia or enlarged liver), or the abnormal result of an investigation (eg, laboratory findings or ECG). In clinical studies, an AE can include an undesirable medical condition

occurring at any time, including run-in or washout periods, even if no study treatment has been administered.

The term 'AE' is used to include both serious and non-serious AEs.

6.3.1 Causality collection

The Investigator will assess the causal relationship between the IPs and each AE and answer "yes" or "no" to the question "Do you consider that there is a reasonable possibility that the event may have been caused by the investigational product?"

For SAEs, causal relationship will also be assessed for other medications and study procedures. Note that, for SAEs that could be associated with any study procedure, the causal relationship is implied as "yes."

A guide to the interpretation of the causality question is found in Appendix B.

6.3.2 Relationship to protocol procedures

The Investigator is also required to provide an assessment of the relationship of SAEs to protocol procedures on the SAE report form. This includes both non-treatment–emergent (ie, SAEs that occur prior to the administration of IP) and treatment-emergent SAEs. A protocol-related SAE may occur as a result of a procedure or intervention required during the study (eg, blood collection). The following guidelines should be used by Investigators to assess the relationship of SAEs to the protocol:

- Protocol related: The event occurred due to a procedure or intervention that was described in the protocol for which there is no alternative etiology present in the patient's medical record.

- Not protocol related: The event is related to an etiology other than the procedure or intervention that was described in the protocol. The alternative etiology must be documented in the study patient's medical record.

6.3.3 Adverse events based on signs and symptoms

All AEs spontaneously reported by the patient or reported in response to the open question from the study personnel: "Have you had any health problems since the previous visit/you were last asked?" or revealed by observation will be collected and recorded in the eCRF. When collecting AEs, the recording of diagnoses is preferred, when possible, to recording a list of signs and symptoms. However, if a diagnosis is known and there are other signs or symptoms that are not generally part of the diagnosis, the diagnosis and each sign or symptom will be recorded separately.

6.3.4 Adverse events based on examinations and tests

The results from protocol-mandated laboratory tests and vital signs measurements will be summarized in the CSR. Deterioration as compared to baseline in protocol-mandated

laboratory values and vital signs should therefore only be reported as AEs if they fulfill any of the SAE criteria or are the reason for discontinuation of treatment with the IPs.

If deterioration in a laboratory value or vital sign is associated with clinical signs and symptoms, the sign or symptom will be reported as an AE and the associated laboratory result or vital sign will be considered as additional information. Whenever possible, the reporting Investigator should use the clinical rather than the laboratory term (eg, anemia versus low hemoglobin value). In the absence of clinical signs or symptoms, clinically relevant deteriorations in non-mandated parameters should be reported as AEs.

Deterioration of a laboratory value that is unequivocally due to disease progression should not be reported as an AE/SAE.

Any new or aggravated clinically relevant abnormal medical finding at a physical examination as compared with the baseline assessment will be reported as an AE.

6.3.5 Hy's law

Cases where a patient shows elevations in liver biochemistry may require further evaluation and occurrences of AST or ALT ≥3 × ULN together with total bilirubin ≥2 × ULN may need to be reported as SAEs. Please refer to Appendix E for further instruction on cases of increases in liver biochemistry and evaluation of Hy's law.

6.3.6 Disease progression

Disease progression can be considered as a worsening of a patient's condition attributable to the disease for which the IP is being studied. It may be an increase in the severity of the disease under study and/or increases in the symptoms of the disease. The development of new or progression of existing metastasis to the primary cancer under study should be considered as disease progression and not an AE. Events that are unequivocally due to disease progression should not be reported as an AE during the study.

6.3.7 New cancers

The development of a new cancer should be regarded as an SAE. New primary cancers are those that are not the primary reason for the administration of the study treatment and have been identified after the patient's inclusion in this study.

6.3.8 Deaths

All deaths that occur during the study, or within the protocol-defined follow-up period after the administration of the last dose of study treatment, must be reported as follows:

- Death clearly resulting from disease progression should be reported to the Study Physician at the next monitoring visit and should be documented in the eCRF. It should not be reported as an SAE.

- Where death is not due (or not clearly due) to progression of the disease under study, the AE causing the death must be reported to the Study Physician as an SAE

within 24 hours. The report should contain a comment regarding the co-involvement of PD, if appropriate, and should assign main and contributory causes of death.

- Deaths with an unknown cause should always be reported as an SAE. A post-mortem may be helpful in the assessment of the cause of death, and if performed, a copy of the post-mortem results should be forwarded to AstraZeneca Drug Safety or its representative within the usual timeframes.

6.4 Reporting of serious adverse events

All SAEs have to be reported, whether or not they are considered causally related to the IPs or to any study procedure. All SAEs will be recorded in the eCRF.

If any SAE occurs in the course of the study, then the Investigator or other site personnel should inform the appropriate AstraZeneca representatives within 1 day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all the necessary information is provided to the AstraZeneca Patient Safety data entry site **within 1 calendar day** of initial receipt for fatal and life-threatening events **and within 5 calendar days** of initial receipt for all other SAEs.

For fatal or life-threatening AEs in which important or relevant information is missing, active follow-up is undertaken immediately. The Investigator or other site personnel will inform AstraZeneca representatives of any follow-up information on a previously reported SAE within 1 calendar day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

Once the Investigator or other site personnel indicates that an AE is serious in the WBDC system, an automated e-mail alert is sent to the designated AstraZeneca representative.

If the WBDC system is not available, then the Investigator or other study site personnel will report an SAE to the appropriate AstraZeneca representative by telephone.

The AstraZeneca representative will advise the Investigator or study site personnel how to proceed.

The reference documents for the definition of expectedness or listedness are the IBs for MEDI4736 and tremelimumab.

The Principal Investigator is responsible for ensuring that procedures and expertise are available to handle medical emergencies during the study. **A medical emergency usually constitutes an SAE and is to be reported as such.**

6.5 Overdose

Use of IP in doses in excess of that specified in the protocol is considered to be an overdose. There is currently no specific treatment in the event of overdose of IP, and possible symptoms of overdose are not established.

- An overdose with associated AEs will be recorded as the AE diagnosis or symptoms in the relevant AE modules of the eCRF and in the Overdose eCRF module.

- An overdose without associated symptoms will only be reported in the Overdose eCRF module.

If an overdose of an AstraZeneca study drug occurs in the course of the study, then the Investigator or other site personnel will inform appropriate AstraZeneca representatives immediately, or **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all relevant information is provided to the AstraZeneca Patient Safety data entry site.

For overdoses associated with an SAE, the standard reporting timelines apply, see Section 6.4. For other overdoses, reporting must occur within 30 days.

6.6 Pregnancy

All pregnancies and outcomes of pregnancy should be reported to AstraZeneca.

6.6.1 Maternal exposure

If a patient becomes pregnant during the course of the study, the IPs should be discontinued immediately.

Pregnancy itself is not regarded as an AE unless there is a suspicion that the IP under study may have interfered with the effectiveness of a contraceptive medication. Congenital abnormalities or birth defects and spontaneous miscarriages should be reported and handled as SAEs. Elective abortions without complications should not be handled as AEs. The outcome of all pregnancies (spontaneous miscarriage, elective termination, ectopic pregnancy, normal birth, or congenital abnormality) should be followed up and documented even if the patient was discontinued from the study.

If any pregnancy occurs in the course of the study, then the Investigator or other site personnel should inform the appropriate AstraZeneca representatives within 1 day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all relevant information is provided to the AstraZeneca Patient Safety data entry site within 1 to 5 calendar days for SAEs (see Section 6.4) and within 30 days for all other pregnancies.

The same timelines apply when outcome information is available.

6.6.2 Paternal exposure

Male patients should refrain from fathering a child or donating sperm during the study and for 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy. Please follow the local prescribing information relating to contraception and the time limit for such precautions for SoC agents.

Pregnancy of the patient's partner is not considered to be an AE. However, the outcome of all pregnancies (spontaneous miscarriage, elective termination, ectopic pregnancy, normal birth, or congenital abnormality) occurring from the date of the first dose until 90 days after the last dose should, if possible, be followed up and documented.

Where a report of pregnancy is received, prior to obtaining information about the pregnancy, the Investigator must obtain the consent of the patient's partner. Therefore, the local study team should adopt the generic ICF template in line with local procedures and submit it to the relevant Ethics Committees (EC)/Internal Review Boards (IRB) prior to use.

6.7 Management of investigational product-related toxicities

The following general guidance should be followed for management of toxicities.

- Treat each of the toxicities with maximum supportive care (including holding the agent suspected of causing the toxicity if required).
- If the symptoms promptly resolve with supportive care, consideration should be given to continuing the same dose of the assigned IP along with appropriate continuing supportive care. If medically appropriate, dose modifications are permitted (see Sections 6.7.1 and 6.7.2). In addition, guidelines on dose modifications are provided in Table 8.
- All dose modifications should be documented with clear reasoning and documentation of the approach taken.

In addition, guidelines on dose modification and toxicity management for immune-mediated reactions, infusion-related reactions, and non-immune-mediated reactions for MEDI4736 and tremelimumab will be provided to investigators. There are certain circumstances in which MEDI4736 and tremelimumab should be permanently discontinued. Following the first dose of IP, subsequent administration of MEDI4736 and tremelimumab can be modified. All toxicities will be graded according to CTCAE version 4.03. Dose reductions are not permitted. Dose modifications of MEDI4736 and tremelimumab may be required in the event of treatment-related toxicity. All toxicities will be graded according to NCI CTCAE v4.03. In case of doubt, the Investigator should consult with the Study Physician.

Acetaminophen and/or an antihistamine (eg, diphenhydramine) or equivalent medications per institutional standard may be administered prior to infusion at the discretion of the

Investigator for primary prophylaxis against infusion-related reactions. In the event of Grade ≤2 infusion-related reaction, the infusion rate of IP may be decreased by 50% or interrupted until resolution of the event (up to 4 hours) and re-initiated at 50% of the initial rate until completion of the infusion. In patients experiencing Grade ≤2 infusion-related reaction, subsequent infusions may be administered at 50% of the initial rate. If a patient experiences an infusion-related reaction, acetaminophen and/or an antihistamine (eg, diphenhydramine) and/or corticosteroid or equivalent medications per institutional standard may be administered prior to subsequent infusions at the discretion of the Investigator for secondary prophylaxis of infusion-related reactions. If the infusion-related reaction is Grade 3 or higher in severity, treatment with IP will be discontinued.

As with any antibody, allergic reactions to dose administration are possible. Appropriate drugs and medical equipment to treat acute anaphylactic reactions must be immediately available, and study personnel must be trained to recognize and treat anaphylaxis. The study site must have immediate access to emergency resuscitation teams and equipment in addition to the ability to admit patients to an intensive care unit if necessary.

6.7.1 MEDI4736 + tremelimumab and MEDI4736 monotherapy adverse events of special interest

Adverse events of special interest (AESIs) are events of scientific and medical interest specific to the further understanding of the MEDI4736 and tremelimumab safety profile and require close monitoring and rapid communication by the Investigator to AstraZeneca. MEDI4736 and tremelimumab AESIs may be serious or non-serious. The rapid reporting of these AESIs allows ongoing analysis of these events in order to characterize and understand them in association with the use of these IPs.

MEDI4736, an anti-PD-L1 antibody, binds with high affinity and specificity to PD-L1 and blocks its binding to PD-1 (CD279) and B7-1 (CD80) molecules, thus promoting anti-tumor immunity and tumor cell killing. Tremelimumab, a CTLA-4 antibody, blocks the inhibitory signal resulting from CTLA-4 binding to B7 ligands on antigen-presenting cells, thus maintaining T-cell homeostasis. Potential risks based on these mechanisms of action include immune-mediated reactions such as enterocolitis, dermatitis, hepatotoxicity or hepatitis, endocrinopathy, neuropathy or neurologic events, pancreatitis, and pneumonitis.

The class including anti-PD-L1 drugs and other immune checkpoint antibodies, such as anti-PD-1 or anti-CTLA-4, have a wide spectrum of immune-mediated reactions that have been considered inflammatory in nature and can affect any organ of the body.

For MEDI4736 and tremelimumab, AESIs will comprise the following:

Pneumonitis

AEs of pneumonitis are also of interest for AstraZeneca, as pneumonitis has been observed with use of anti-PD-1 mAbs (but not with anti-PD-L1 mAbs). Initial work-up should include a high-resolution CT scan, ruling out infection, and pulse oximetry. Pulmonary consultation

is highly recommended. Guidelines for the management of patients with immune-related adverse events (irAEs) including pneumonitis are provided in Table 8.

Infusion reactions

AEs of infusion reactions (also termed infusion-related reactions) are of special interest to AstraZeneca and are defined, for the purpose of this protocol, as all AEs occurring from the start of IP infusion up to 48 hours after the infusion start time. For all infusion reactions, the eCRF should be completed as instructed in Section 6.2, and all SAEs should be reported to AstraZeneca Patient safety as described in Section 6.4.

Hypersensitivity reactions

Hypersensitivity reactions as well as infusion-related reactions have been reported with anti-PD-L1 and anti-PD-1 therapy (Brahmer et al 2012). As with the administration of any foreign protein and/or other biologic agents, reactions following the infusion of mAbs can be caused by various mechanisms, including acute anaphylactic (immunoglobulin E-mediated) and anaphylactoid reactions against the mAbs and serum sickness. Acute allergic reactions may occur, may be severe, and may result in death. Acute allergic reactions may include hypotension, dyspnea, cyanosis, respiratory failure, urticaria, pruritis, angioedema, hypotonia, arthralgia, bronchospasm, wheeze, cough, dizziness, fatigue, headache, hypertension, myalgia, vomiting, and unresponsiveness. Guidelines for the management of patients with hypersensitivity (including anaphylactic reaction) and infusion-related reactions are provided in Table 8.

Hepatic function abnormalities (hepatotoxicity)

Hepatic function abnormality is defined as any increase in ALT or AST to greater than 3 × ULN and concurrent increase in total bilirubin to be greater than 2 × ULN. Concurrent findings are those that derive from a single blood draw or from separate blood draws taken within 8 days of each other. Follow-up investigations and inquiries will be initiated promptly by the investigational site to determine whether the findings are reproducible and/or whether there is objective evidence that clearly supports causation by a disease (eg, cholelithiasis and bile duct obstruction with distended gallbladder) or an agent other than the IP. Guidelines for management of patients with hepatic function abnormality are provided in Table 8.

Gastrointestinal disorders

Diarrhea/colitis is the most commonly observed treatment emergent SAE when tremelimumab is used as monotherapy. In rare cases, colon perforation may occur that requires surgery (colectomy) or can lead to a fatal outcome if not properly managed. Guidelines on management of diarrhea and colitis in patients receiving tremelimumab are provided in Table 8.

Endocrine disorders

Immune-mediated endocrinopathies include hypophysitis, adrenal insufficiency, and hyper- and hypothyroidism. Guidelines for the management of patients with immune-mediated endocrine events are provided in Table 8.

Pancreatic disorders

Immune-mediated pancreatitis includes autoimmune pancreatitis, and lipase and amylase elevation. Guidelines for the management of patients with immune-mediated pancreatic disorders are provided in Table 8.

Neurotoxicity

Immune-mediated nervous system events include encephalitis, peripheral motor and sensory neuropathies, Guillain-Barré, and myasthenia gravis. Guidelines for the management of patients with immune-mediated neurotoxic events are provided in Table 8.

6.7.2 Immune-related adverse events

Based on the mechanism of action of MEDI4736 and tremelimumab leading to T-cell activation and proliferation, there is a possibility of observing irAEs during the conduct of this study. Potential irAEs may be similar to those seen with the use of ipilimumab, BMS-936558 (anti-PD-1 mAb), and BMS-936559 (anti-PD-L1 mAb) and may include immune-mediated enterocolitis, dermatitis, hepatitis (hepatotoxicity), pneumonitis, and endocrinopathies (Hodi et al 2010, Brahmer et al 2012, Topalian et al 2012). These AEs are inflammatory in nature and can affect any organ. With anti-PD-L1 and anti-CTLA-4 combination therapy, the occurrence of overlapping or increasing cumulative toxicities that include irAEs could potentially occur at higher frequencies than with either MEDI4736 or tremelimumab monotherapy. Patients should be monitored for signs and symptoms of irAEs. In the absence of an alternate etiology (eg, infection or PD), an immune-related etiology should be considered for signs or symptoms of enterocolitis, dermatitis, pneumonitis, hepatitis, and endocrinopathy. In addition to the dose modification guidelines provided in Table 8, it is recommended that irAEs are managed according to the general treatment guidelines outlined for ipilimumab (Weber et al 2012). These guidelines recommend the following:

1. Patients should be evaluated to identify any alternative etiology.

2. In the absence of a clear alternative etiology, all events of an inflammatory nature should be considered immune related.

3. Symptomatic and topical therapy should be considered for low-grade events.

4. Systemic corticosteroids should be considered for a persistent low-grade event or for a severe event.

5. More potent immunosuppressives should be considered for events not responding to systemic steroids (eg, infliximab or mycophenolate).

If the Investigator has any questions in regards to an AE being an irAE, the Investigator should immediately contact the Study Physician.

Table 8 Dosing modification and toxicity management guidelines for immune-mediated, infusion-related, and non immune-mediated reactions (MEDI4736 monotherapy or combination therapy with tremelimumab)

Immune-mediated reactions		
	Dose modifications	**Toxicity management**
irAEs (overall management)	Drug administration modifications of study drug/study regimen will be made to manage potential irAEs based on severity of treatment-emergent toxicities graded per NCI CTCAE, v4.03. Grade 1: No dose modification Grade 2: Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1 – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date Grade 3: Depending on the individual toxicity, may permanently discontinue study drug/study regimen. Please refer to guidelines below. Grade 4: Permanently discontinue study drug/study regimen	It is recommended that management of irAEs follow the guidelines presented in this table. – Subjects should be thoroughly evaluated to rule out any alternative etiology (eg, disease progression, concomitant medications, or infections) – In the absence of a clear alternative etiology, all events should be considered potentially immune-related. – Symptomatic and topical therapy should be considered for low-grade (Grade 1 or 2, unless otherwise specified) events – Systemic corticosteroids (eg, prednisone or IV equivalent) should be considered for persistent low-grade or severe (Grade ≥3) events – If symptoms recur or worsen during corticosteroid tapering, increase the corticosteroid dose until stabilization or improvement of symptoms, then resume corticosteroid tapering at a slower rate – More potent immunosuppressives – TNF antagonist class (eg, infliximab) or mycophenolate should be considered for events not responding to systemic steroids after discussion with Study Physician – Discontinuation of study drug is not mandated for Grade 3 or 4 inflammatory reactions attributed to local tumor response (eg, inflammatory reaction at sites of metastatic disease or lymph nodes)
Pneumonitis/ILD	Any Grade	– Monitor subjects for signs and symptoms of pneumonitis or ILD (new onset or worsening shortness of breath or cough). Subjects should be evaluated with imaging and pulmonary function tests, including other diagnostic procedures as described below – Initial work-up may include clinical evaluation, monitoring of oxygenation via pulse oximetry (resting and exertion), laboratory work-up, and high-resolution CT scan.
	Grade 1: No dose modification required. However, consider holding study drug/study regimen dosing as clinically appropriate and during diagnostic work-up for other etiologies.	For Grade 1 (radiographic changes only): – Monitor and closely follow up in 2 to 4 days for clinical symptoms, pulse oximetry (resting and exertion), and laboratory work-up and then as clinically indicated – Consider pulmonary and infectious disease consult

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Pneumonitis/ILD (continued)	Grade 2	Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1 – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date	For Grade 2 (mild to moderate new symptoms): – Monitor symptoms daily and consider hospitalization – Discuss with study physician and consider systemic steroids (eg, prednisone 1 to 2 mg/kg/day or IV equivalent – Reimaging as clinically indicated – If no improvement within 3 to 5 days, additional workup and treatment with IV methylprednisolone 2 to 4 mg/kg/day should be considered – If no improvement within 3 to 5 days, further immunosuppressive therapy (eg, infliximab) should be considered. – Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics – Consider pulmonary and infectious disease consult
	Grade 3 or 4	Permanently discontinue study drug/study regimen	For Grade 3 or 4 (severe or new symptoms, new/worsening hypoxia, life threatening): – Discuss with Study Physician – Pulmonary and infectious disease consult – Hospitalize the subject – Supportive care (oxygen, etc) – Initiate empiric IV corticosteroids (eg, methylprednisolone or equivalent) at 1 to 4 mg/kg/day – If no improvement within 3 to 5 days, additional workup and treatment with additional immunosuppressive therapy (eg, infliximab) should be considered – Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics

Immune-mediated reactions		
	Dose modifications	**Toxicity management**
Diarrhea/Enterocolitis	Any Grade	– Monitor for symptoms that may be related to diarrhea/enterocolitis (abdominal pain, cramping, or changes in bowel habits) – Subjects should be thoroughly evaluated to rule out any alternative etiology (eg, disease progression or infections) – Steroids should be considered if an alternative etiology is not determined, even for low-grade events, in order to prevent potential progression to higher grade event – Use analgesics carefully; they can mask symptoms of perforation and peritonitis
	Grade 1 No dose modification	For Grade 1: – Close monitoring for worsening symptoms – Consider symptomatic treatment, including hydration, electrolyte replacement, dietary changes (eg, American Dietetic Association colitis diet), and loperamide
	Grade 2 Hold study drug/study regimen until resolution to Grade ≤1 – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date	For Grade 2: – Consider symptomatic treatment, including hydration, electrolyte replacement, dietary changes (eg, American Dietetic Association colitis diet), and loperamide and/or budesonide – If event is persistent (>3 to 5 days) or worsens, consider prednisone 0.5 to 1 mg/kg/day or IV equivalent – If not responsive within 3 to 5 days, consider IV corticosteroids (eg, methylprednisolone IV or equivalent) at 1 to 2 mg/kg/day – If event is not responsive within 3 to 5 days or worsens, additional workup and treatment with IV methylprednisolone 2 to 4 mg/kg/day should be considered – If no improvement within 3 to 5 days, further immunosuppressives (eg, infliximab) should be considered – Consult Study Physician if no resolution to Grade ≤1 in 3 to 4 days – Once improving, gradually taper steroids over ≥4 weeks

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Diarrhea/Enterocolitis (continued)	Grade 3 or 4	Permanently discontinue study drug/study regimen	For Grade 3 or 4: – Discuss with Study Physician – Monitor stool frequency and volume and maintain hydration – Urgent GI consult and imaging as appropriate – Initiate empiric IV corticosteroids (eg, methylprednisolone IV or equivalent) at 1 to 4 mg/kg/day – If no improvement within 3 to 5 days, consider further immunosuppressives (eg, infliximab). – Caution: ensure GI consult to rule out bowel perforation and refer to label before using infliximab. Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics
Hepatitis (elevated LFTs)	Any Grade		– Monitor and evaluate liver function test: AST, ALT, ALP, and total bilirubin – Evaluate for alternative etiologies (eg, viral hepatitis, disease progression, or concomitant medications)
	Grade 1	No dose modification If it worsens, treat as Grade 2 event	– Continue LFT monitoring per protocol
	Grade 2	Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1 – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date	For Grade 2 : – Discuss with Study Physician if no resolution to Grade ≤1 in 1 to 2 days – Recheck LFTs in 1 to 2 days. If event is persistent (>3 to 5 days) or worsens, consider prednisone 0.5 to 1 mg/kg/day or IV equivalent. – If no improvement within 3 to 5 days, consider additional workup and treatment with IV methylprednisolone 2 to 4 mg/kg/day o If no improvement within 3 to 5 days, consider further immunosuppressives (eg, mycophenolate mofetil) o Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Hepatitis (elevated LFTs) (continued)	Grade 3	For elevations in transaminases ≤8 × ULN or elevations in bilirubin ≤5 × ULN – Hold study drug/study regimen dose until resolution to Grade ≤1 or baseline – Resume study drug/study regimen administration at the next scheduled dose if elevations downgrade Grade ≤1 or baseline within 14 days Permanently discontinue study drug/study regimen if the elevations do not downgrade to Grade ≤1 or baseline within 14 days For elevations in transaminases >8 × ULN or elevations in bilirubin >5 × ULN, discontinue study drug/study regimen Permanently discontinue study drug/study regimen for any case meeting Hy's law criteria	For Grade 3 or 4: – Discuss with the Study Physician – Initiate empiric IV corticosteroids (eg, methylprednisolone IV or equivalent) at 1 to 4 mg/kg/day o If no improvement within 3 to 5 days, consider further immunosuppressive therapy (eg, mycophenolate mofetil) o If still no further improvement within 3 to 5 days consider other immunosuppressive therapy per local guidelines – Hepatology consult, abdominal workup, and imaging as appropriate. o Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics
	Grade 4	Permanently discontinue study drug/study regimen	
Rash (excluding bullous skin formations)	Any Grade		Monitor for signs and symptoms of dermatitis (rash and pruritus) **IF THERE IS ANY BULLOUS FORMATION, THE STUDY PHYSICIAN SHOULD BE CONTACTED**
	Grade 1	No dose modification	For Grade 1: – Consider symptomatic treatment including oral antipruritics (eg, diphenhydramine or hydroxyzine) and topical therapy (eg, urea cream)

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Rash (excluding bullous skin formations) (continued)	Grade 2	For persistent (>1 to 2 weeks) Grade 2 events, hold scheduled study drug/study regimen until resolution to Grade ≤1 or baseline – If toxicity worsens, then treat as Grade 3 – If toxicity improves, then resume administration at next scheduled dose	For Grade 2 : – Consider symptomatic treatment including oral antipruritics (eg, diphenhydramine or hydroxyzine) and topical therapy (eg, urea cream) – Consider moderate-strength topical steroid – If no improvement of rash/skin lesions occurs within 3 to 5 days or is worsening, discuss with Study Physician and consider systemic steroids prednisone 0.5 to 1 mg/kg/day or IV equivalent – Consider dermatology consult – Consider skin biopsy if persistent for >1 to 2 weeks or recurs
	Grade 3	Hold study drug/study regimen until resolution to Grade ≤1 or baseline If temporarily holding the study drug/study regimen does not provide improvement of the Grade 3 skin rash to Grade ≤1 or baseline within 30 days, then permanently discontinue study drug/study regimen	For Grade 3 or 4: – Discuss with Study Physician – Consider hospitalization – Monitor extent of rash (Rule of Nines) – Consult dermatology – Consider skin biopsy (preferably more than 1) as clinically feasible. – Initiate empiric IV corticosteroids (eg, methylprednisolone IV or equivalent) at 1 to 4 mg/kg/day – Once improving, gradually taper steroids over ≥4 weeks and consider prophylactic antibiotics
	Grade 4	Permanently discontinue study drug/study regimen	
Endocrinopathy (eg, hyperthyroidism, hypothyroidism, hypopituitarism, or adrenal insufficiency)	Any Grade		– Monitor subjects for signs and symptoms of endocrinopathies. Nonspecific symptoms include headache, fatigue, behavior changes, changed mental status, vertigo, abdominal pain, unusual bowel habits, hypotension, and weakness. – Subjects should be thoroughly evaluated to rule out any alternative etiology (eg, disease progression, including brain metastases or infections) – Monitor and evaluate thyroid function tests: TSH, free T_3 and free T_4 and other relevant endocrine labs depending on suspected endocrinopathy. – If a subject experiences an AE that is thought to be possibly of autoimmune nature (eg, thyroiditis, pancreatitis, hypophysitis, or diabetes insipidus), the Investigator should send a blood sample for appropriate autoimmune antibody testing

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Endocrinopathy (eg, hyperthyroidism, hypothyroidism, hypopituitarism, or adrenal insufficiency) (continued)	Grade 1	No dose modification	For Grade 1 (including those with asymptomatic TSH elevation): – Monitor subject with appropriate endocrine function tests – If TSH <0.5 × LLN, or TSH >2 × ULN or consistently out of range in 2 subsequent measurements, include free T_4 at subsequent cycles as clinically indicated and consider endocrinology consult.
	Grade 2	For Grade 2 endocrinopathy other than hypothyroidism, hold study drug/study regimen dose until subject is clinically stable – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date	For Grade 2 (including those with symptomatic endocrinopathy): – Isolated hypothyroidism may be treated with replacement therapy without treatment interruption and without corticosteroids – Discuss with Study Physician – Initiate hormone replacement as needed for management – Evaluate endocrine function, and as clinically indicated, consider pituitary scan – For subjects with abnormal endocrine work up, except for those with isolated hypothyroidism, consider short-term, high-dose corticosteroids (eg, methylprednisolone or IV equivalent) with relevant hormone replacement (eg, levothyroxine, hydrocortisone, or sex hormones). For subjects with normal endocrine work up (lab or MRI scans), repeat labs/MRI as clinically indicated.
	Grade 3 or 4	For Grade 3 or 4 endocrinopathy other than hypothyroidism, hold study drug/study regimen dose until endocrinopathy symptom(s) are controlled Resume study drug/study regimen administration if controlled at the next scheduled dose	For Grade 3 or 4: – Isolated hypothyroidism may be treated with replacement therapy without treatment interruption and without corticosteroids – Discuss with Study Physician – Initiate empiric IV corticosteroids (eg, methylprednisolone IV or equivalent) at 1 to 2 mg/kg/day – Administer hormone replacement therapy as necessary – For adrenal crisis, severe dehydration, hypotension, or shock: immediately initiate IV corticosteroids with mineralocorticoid activity – Consult endocrinologist – Once improving, gradually taper immunosuppressive steroids over ≥4 weeks

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Immune mediated neurotoxicity (except myasthenia gravis and Guillain-Barré)	Any Grade		– Subjects should be evaluated to rule out any alternative etiology (eg, disease progression, infections, metabolic syndromes, or medications) – Monitor subject for general symptoms (headache, nausea, vertigo, behavior change, or weakness) – Consider appropriate diagnostic testing (eg, electromyogram and nerve conduction investigations) – Symptomatic treatment with neurological consult as appropriate
	Grade 1	No dose modifications	
	Grade 2	For acute motor neuropathies or neurotoxicity, hold study drug/study regimen dose until resolution to Grade ≤1 For sensory neuropathy/neuropathic pain, consider holding study drug/study regimen dose until resolution to Grade ≤1. – If toxicity worsens, then treat as Grade 3 or 4 – If toxicity improves to baseline, then treat at next scheduled treatment date	– Discuss with the Study Physician – Consider neurology consult – Sensory neuropathy/neuropathic pain may be managed by appropriate medications (eg, gabapentin or duloxetine) – Consider systemic steroids prednisone 1 to 2 g/kg/day or IV equivalent at 0.5 to 1 mg/kg/day – If no improvement within 3 to 5 days, consider additional workup and treatment with additional immunosuppressive therapy (eg, IV IgG)
	Grade 3	– Hold study drug/study regimen dose until resolution to Grade ≤1 – Permanently discontinue study drug/study regimen if Grade 3 irAE does not resolve to Grade ≤1 within 30 days.	– Discuss with Study Physician – Consult neurology consult – Consider hospitalization – Consider empiric IV corticosteroids (eg, methylprednisolone or IV equivalent) at 1 to 2 mg/kg/day – If no improvement within 3 to 5 days, consider additional workup and treatment with additional immunosuppressants (eg, IV IgG) – Once stable, gradually taper steroids over ≥4 weeks
	Grade 4	– Permanently discontinue study drug/study regimen	

Immune-mediated reactions		
	Dose modifications	**Toxicity management**
Immune-mediated peripheral neuromotor syndromes, such as Guillain-Barré and myasthenia gravis	Any Grade	– The prompt diagnosis of immune-mediated peripheral neuromotor syndromes is important, since certain subjects may unpredictably experience acute decompensations that can result in substantial morbidity or, in the worst case, death. Special care should be taken for certain sentinel symptoms which may predict a more severe outcome, such as prominent dysphagia, rapidly progressive weakness, and signs of respiratory insufficiency or autonomic instability – Subjects should be evaluated to rule out any alternative etiology (eg, disease progression, infections, metabolic syndromes, or medications). It should be noted that the diagnosis of immune-mediated peripheral neuromotor syndromes can be particularly challenging in subjects with underlying cancer, due to the multiple potential confounding effects of cancer (and its treatments) throughout the neuraxis. Given the importance of prompt and accurate diagnosis, it is essential to have a low threshold to obtain a neurological consult – Neurophysiologic diagnostic testing (eg, electromyogram and nerve conduction investigations, and "repetitive stimulation" if myasthenia is suspected) are routinely indicated upon suspicion of such conditions and may be best facilitated by means of a neurology consultation – Important to consider that the use of steroids as the primary treatment of Guillain-Barré is not typically considered effective. Subjects requiring treatment should be considered for plasmapheresis (or IV IgG, as an alternative)
	Grade 1 No dose modification	– Discuss with the Study Physician – Care should be taken to monitor subjects for sentinel symptoms of a potential decompensation, as described above – Consider a neurology consult unless the symptoms are very minor and stable

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Immune-mediated peripheral neuromotor syndromes, such as Guillain-Barré and myasthenia gravis (continued)	Grade 2	Hold study drug/study regimen dose until resolution to Grade ≤1 Permanently discontinue study drug/study regimen if it does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency or autonomic instability	Grade 2: Moderate – Discuss with the Study Physician – Care should be taken to monitor subjects for sentinel symptoms of a potential decompensation as described above – Obtain a neurology consult – Sensory neuropathy/neuropathic pain may be managed by appropriate medications (eg, gabapentin or duloxetine) – *MYASTHENIA GRAVIS* o Steroids may be successfully used to treat myasthenia gravis. Important to consider that steroid therapy (especially with high doses) may result in transient worsening of myasthenia and should typically be administered in a monitored setting under supervision of a consulting neurologist. o Subjects unable to tolerate steroids may be candidates for treatment with plasmapheresis or IV IgG. Such decisions are best made in consultation with a neurologist, taking into account the unique needs of each subject. o If myasthenia gravis-like neurotoxicity present, consider starting AChE-inhibitor therapy in addition to steroids. Such therapy, if successful, can also serve to reinforce the diagnosis. – *GUILLAIN-BARRÉ:* o Important to consider here that the use of steroids as the primary treatment of Guillain-Barré is not typically considered effective. Subjects requiring treatment should be considered for plasmapheresis (or IV IgG, as an alternative).
	Grade 3	Hold study drug/study regimen dose until resolution to Grade ≤1 Permanently discontinue study drug/study regimen if Grade 3 irAE does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency or autonomic instability	– For severe or life threatening (Grade 3 or 4) events: – Discuss with Study Physician – Recommend hospitalization – Monitor symptoms and obtain neurological consult *MYASTHENIA GRAVIS* o Steroids may be successfully used to treat myasthenia gravis. It

Immune-mediated reactions			
	Dose modifications		**Toxicity management**
Immune-mediated peripheral neuromotor syndromes, such as Guillain-Barré and myasthenia gravis (continued)	Grade 4	Permanently discontinue study drug/study regimen	should typically be administered in a monitored setting under supervision of a consulting neurologist. o Subjects unable to tolerate steroids may be candidates for treatment with plasmapheresis or IV IgG. o If myasthenia gravis-like neurotoxicity present, consider starting AChE-inhibitor therapy in addition to steroids. Such therapy, if successful, can also serve to reinforce the diagnosis. *GUILLAIN-BARRÉ*: – Important to consider here that the use of steroids as the primary treatment of Guillain-Barré is not typically considered effective. Subjects requiring treatment should be considered for plasmapheresis (or IV IgG, as an alternative).

<table>
<tr><th colspan="3">Infusion-related reactions</th></tr>
<tr><th>Infusion-related reactions</th><th>Dose modifications</th><th>Toxicity management</th></tr>
<tr><td rowspan="2"></td><td>Any Grade</td><td>Management per institutional standard at the discretion of Investigator
Monitor subjects for signs and symptoms of infusion-related reactions (eg, fever and/or shaking chills, flushing and/or itching, alterations in heart rate and blood pressure, dyspnea or chest discomfort, or skin rashes) and anaphylaxis (eg, generalized urticaria, angioedema, wheezing, hypotension, or tachycardia)</td></tr>
<tr><td>Grade 1
The infusion rate of study drug/study regimen may be decreased by 50% or temporarily interrupted until resolution of the event</td><td rowspan="2">For Grade 1 or Grade 2:
– Acetaminophen and/or antihistamines may be administered per institutional standard at the discretion of the Investigator
– Consider premedication per institutional standard prior to subsequent doses</td></tr>
<tr><td rowspan="2"></td><td>Grade 2
The infusion rate of study drug/study regimen may be decreased 50% or temporarily interrupted until resolution of the event (up to 4 hours).

Subsequent infusions may be given at 50% of the initial infusion rate</td></tr>
<tr><td>Grade 3 or 4
Permanently discontinue study drug/study regimen</td><td>For Grade 3 or 4:
– Manage severe infusion-related reactions per institutional standards (eg, IM epinephrine, followed by IV diphenhydramine and ranitidine, and IV glucocorticoid)</td></tr>
</table>

Non-immune mediated reactions		
CTC Grade/Severity	**Dose modification**	**Toxicity management**
Any Grade	Note: dose modifications are not required for AEs not deemed to be related to study treatment (ie, events due to underlying disease) or for laboratory abnormalities not deemed to be clinically significant.	Treat accordingly as per institutional standard
1	No dose adjustment	Treat accordingly as per institutional standard
2	Hold study drug/study regimen until resolution to Grade ≤1 or baseline	Treat accordingly as per institutional standard
3	Hold study drug/study regimen until resolution to Grade ≤1 or baseline For AEs that downgrade to Grade ≤2 within 7 days or resolve to Grade ≤1 or baseline within 14 days, resume study drug/study regimen administration at next scheduled dose. Otherwise, discontinue study drug/study regimen.	Treat accordingly as per institutional standard
4	Discontinue study drug/study regimen (note, for Grade 4 labs, decision to discontinue would be based on accompanying clinical signs/symptoms and as per Investigator's clinical judgment and in consultation with the Sponsor)	Treat accordingly as per institutional standard

AChE Acetylcholine esterase; ADA American Dietetic Association; ALP Alkaline phosphatase; GI Gastrointestinal; IDS Infectious Disease Service; ILD Interstitial lung disease; IM Intramuscular; PO By mouth; TNF Tumor necrosis factor; TSH Thyroid stimulating hormone; ULN Upper limit of normal.

6.7.3 Standard of Care agents

IP-related toxicity management, including dose delays, reductions, and adjustments for patients in the SoC group should be performed as indicated in the local prescribing information for the relevant agent.

6.8 Study governance and oversight

The safety of all AstraZeneca clinical studies is closely monitored on an ongoing basis by AstraZeneca representatives in consultation with Patient Safety. Issues identified will be addressed; for instance, this could involve amendments to the study protocol and letters to Investigators.

An IDMC comprised of independent experts will be convened and will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review safety assessments and make recommendations to continue, amend, or stop the study based on safety findings. The committee will meet at least every 6 months thereafter.

Full details of the IDMC procedures, processes, and interim analyses can be found in the IDMC Charter.

7. INVESTIGATIONAL PRODUCT AND OTHER TREATMENTS

7.1 Identity of investigational product(s)

AstraZeneca will supply MEDI4736 and tremelimumab, while the SoC treatments (paclitaxel + carboplatin, gemcitabine + cisplatin, gemcitabine + carboplatin, pemetrexed + cisplatin, pemetrexed + carboplatin, and pemetrexed maintenance) will be supplied locally (Table 9).

Table 9 List of investigational products for this study

Investigational product	Dosage form and strength	Manufacturer
MEDI4736	50 mg/mL, solution, IV	MedImmune
Tremelimumab	20 mg/mL, solution, IV	MedImmune
Standards of Care		
Paclitaxel[a]	IV (as sourced locally)	Sourced locally
Carboplatin[a]	IV (as sourced locally)	Sourced locally
Gemcitabine[a]	IV (as sourced locally)	Sourced locally
Cisplatin[a]	IV (as sourced locally)	Sourced locally
Pemetrexed[a]	IV (as sourced locally)	Sourced locally

[a] Under certain circumstances when local sourcing is not feasible, a Standard of Care treatment may be supplied centrally through AstraZeneca.

7.1.1 MEDI4736

MEDI4736 will be supplied by AstraZeneca as a 500-mg vial concentrate for solution for infusion. The solution contains 50 mg/mL MEDI4736, 26 mM histidine/histidine-hydrochloride, 275 mM trehalose dihydrate, and 0.02% (weight/volume) polysorbate 80; it has a pH of 6.0. Total in-use storage time from needle puncture of MEDI4736 vial to the start of administration should not exceed 4 hours at room temperature or 24 hours at 2°C to 8°C (36°F to 46°F). If the in-use storage time exceeds these limits, a new dose must be prepared from new vials. MEDI4736 does not contain preservatives, and any unused portion must be discarded.

Preparation of MEDI4736 doses for administration with an IV bag

Doses of 20 mg/kg will be administered using a 250-mL IV bag containing 0.9% (weight/volume) saline and delivered through an IV administration set with a 0.2-µm in-line filter.

Patient weight at baseline should be used for dosing calculations unless there is a ≥10% change in weight. Dosing day weight can be used for dosing calculations instead of baseline weight per institutional standard. A volume of 0.9% (weight/volume) saline equal to the calculated volume of MEDI4736 to be added to the IV bag must be removed from the bag prior to addition of MEDI4736. The calculated volume of MEDI4736 is then added to the IV bag, and the bag is mixed by gentle inversion to ensure homogeneity of the dose in the bag.

No incompatibilities between MEDI4736 and polyethylene, polypropylene, polyvinylchloride, or polyolefin copolymers have been observed.

Dose calculation

The volume of MEDI4736 (in mL) to add to the IV bag is calculated as follows:

In combination with tremelimumab: 20 mg/kg × patient weight (kg) ÷ MEDI4736 concentration (nominal: 50 mg/mL)

As monotherapy: 20 mg/kg × patient weight (kg) ÷ MEDI4736 concentration (nominal: 50 mg/mL)

Example: For a patient weighing 80 kg, dosed at 10 mg/kg, 16 mL [10 mg/kg × 80 kg divided by 50 mg/mL] of MEDI4736 is to be diluted in a 250 mL IV bag containing 0.9% (weight/volume) saline. First, 16 mL of saline is removed from the IV bag, and then 16 mL of MEDI4736 is added to the bag. The bag is mixed by gentle inversion to ensure homogeneity of the dose in the bag and the diluted MEDI4736 is administered as described above.

7.1.2 Tremelimumab

Tremelimumab is supplied as a sterile solution for IV infusion, filled in 20 mL clear glass vials with a rubber stopper and aluminum seal. Each vial contains 20 mg/mL (with a nominal fill of 20 mL, accounting to 400 mg/vial) of tremelimumab, in an isotonic solution at pH 5.5.

Product preparation and reconstitution of tremelimumab

The dose of tremelimumab for administration must be prepared by the Investigator's or site's designated IP manager using aseptic technique. Total in-use storage time from needle puncture of the product vial to start of administration should not exceed 4 hours at room temperature or 24 hours at 2°C to 8°C (36°F to 46°F). If in-use storage time exceeds these limits, a new dose must be prepared from new vials. Tremelimumab does not contain preservatives and any unused portion must be discarded.

P**reparation of tremelimumab doses for administration with an IV bag**

All details can be found in the Drug Handling Instructions.

Dose calculation of tremelimumab

The dose will be calculated using the following formula:

1 mg/kg × Patient Weight (kg) ÷ tremelimumab concentration (nominal 20 mg/mL)

The corresponding volume of MEDI4736 and tremelimumab should be rounded to the nearest tenth of an mL (0.1 mL). Dose adjustments for each cycle are only needed for greater than 10% change in weight.

7.1.3 Standard of Care treatment

Each SoC agent will be sourced as commercially available material/locally sourced, prescribed according to local regulations, and will be administered according to prescribing information or treatment guidance in general use by the Investigating site. Under certain circumstances when local sourcing is not feasible, an SoC will be supplied centrally by AstraZeneca. This will be labeled with local language translated text in accordance with regulatory guidelines.

7.2 Dose and treatment regimens

Patients will be randomized in a 1:1:1 ratio to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC.

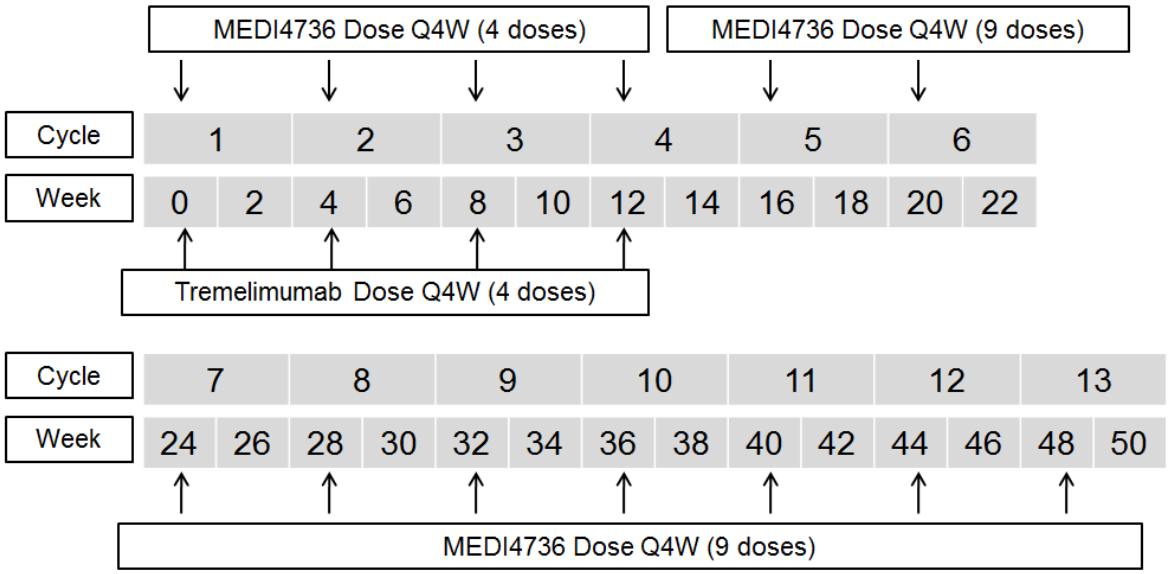
7.2.1 Treatment regimens

MEDI4736 + tremelimumab combination therapy

Patients in the MEDI4736 + tremelimumab combination therapy group will receive 20 mg/kg MEDI4736 via IV infusion q4w for up to 4 doses/cycles and 1 mg/kg tremelimumab via IV infusion q4w for up to 4 doses/cycles, and then continue 20 mg/kg MEDI4736 q4w starting on Week 16 for up to 8 months (9 doses) (see Figure 3). Dosing outside the window should be discussed with the Study Physician. Tremelimumab will be administered first. MEDI4736 infusion will start approximately 1 hour after the end of tremelimumab infusion. The duration will be approximately 1 hour for each infusion. A 1-hour observation period is recommended

after the first infusion of MEDI4736 and tremelimumab. If no clinically significant infusion reactions are observed during or after the first cycle, subsequent infusion observation periods can be at the Investigator's discretion (suggested 30 minutes after each MEDI4736 and tremelimumab infusion).

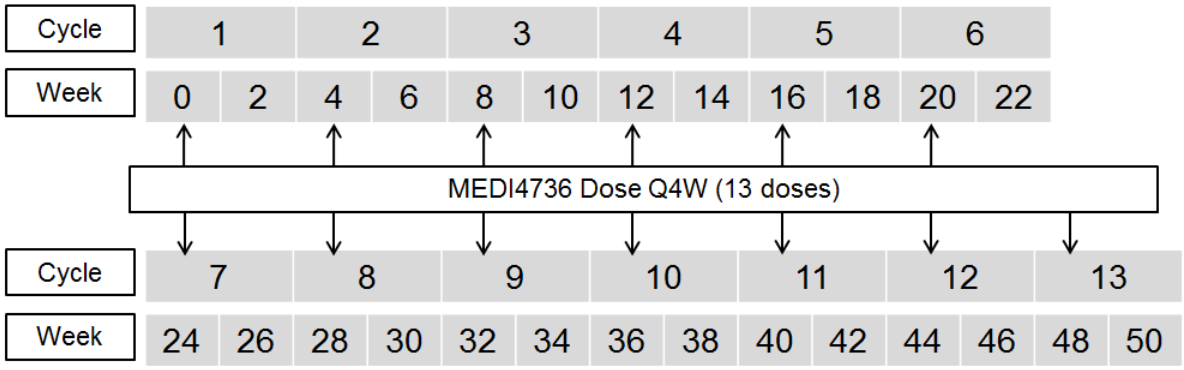
Figure 3 MEDI4736 + tremelimumab combination therapy dosing scheme



MEDI4736 monotherapy

Patients in the MEDI4736 monotherapy treatment group will receive 20 mg/kg MEDI4736 via IV infusion q4w for up to 12 months (up to 13 doses; see Figure 4).

Figure 4 MEDI4736 monotherapy dosing scheme



Standard of Care treatment

Patients in the SoC group will receive 1 of the following treatments until documented PD (unless the Investigator considers the patient would continue to receive benefit from treatment), initiation of alternative anticancer therapy, unacceptable toxicity, withdrawal of consent to continued treatment, or other reasons to discontinue treatment criterion occur:

- Paclitaxel + carboplatin: Paclitaxel 200 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + cisplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + cisplatin 75 or 80 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + carboplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Pemetrexed + cisplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and cisplatin 75 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.
- Pemetrexed + carboplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.

For all SoC therapies, a particular treatment (paclitaxel, gemcitabine, cisplatin, carboplatin, or pemetrexed) will not be used in patients who have previously received that treatment for locally advanced or metastatic disease or who have experienced recurrence or progression of disease within 6 months of prior multimodal therapy using that particular treatment.

A confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

7.2.2 Duration of treatment and criteria for retreatment

Until specific treatment discontinuation criteria are met, treatment will continue for a 12-month period for the MEDI4736 + tremelimumab and MEDI4736 monotherapy groups and with no maximum treatment duration for the SoC group.

For patients randomized to SoC, crossover to MEDI4736 + tremelimumab or MEDI4736 monotherapy following disease progression will not be permitted.

For patients randomized to MEDI4736 + tremelimumab or MEDI4736 monotherapy, retreatment with their assigned treatment is allowed (once only) for patients meeting the retreatment criteria below. The same treatment guidelines followed during the initial 12-month treatment period will be followed during the re-treatment period, including the same dose and frequency of treatments and the same schedule of assessments.

Patients randomized to MEDI4736 monotherapy may undergo retreatment in 1 clinical scenario, described below:

1. Patients who achieve or maintain disease control (ie, CR, PR, or SD) through to the end of the 12-month treatment period may restart their assigned treatment (MEDI4736 monotherapy) upon evidence of PD, with or without confirmation according to RECIST 1.1, during follow-up.

Patients randomized to the combination of MEDI4736 and tremelimumab may undergo retreatment in 2 clinical scenarios, described below:

1. Patients who achieve and maintain disease control (ie, CR, PR, or SD) through to the end of the 12-month treatment period may restart treatment with the combination upon evidence of PD, with or without confirmation according to RECIST 1.1, during follow-up.

2. Patients who complete the 4 dosing cycles of the combination of MEDI4736 and tremelimumab portion of the regimen (with clinical benefit per Investigator judgment), but subsequently have evidence of PD during the MEDI4736 monotherapy portion of the combination regimen, with or without confirmation according to RECIST 1.1, may restart treatment with the combination.

For either MEDI4736 + tremelimumab or MEDI4736 monotherapy treatment group, before restarting their assigned treatment, the investigator should ensure that the patient:

1. Does not have any significant, unacceptable, or irreversible toxicities that indicate continuing treatment will not further benefit the patient

2. Still fulfils the eligibility criteria for this study, including reconsenting to restart MEDI4736 or MEDI4736 and tremelimumab

3. Has not have received an intervening systemic anticancer therapy after their assigned treatment discontinuation

4. Has had a baseline tumor assessment within 28 days of restarting their assigned treatment; all further scans should occur with the same frequency as during the initial 12 months of treatment (relative to the date of randomization) until study treatment is stopped (maximum of 12 months of further treatment).

During the retreatment period, patients in the MEDI4736 + tremelimumab group will resume MEDI4736 dosing at 20 mg/kg q4w with 1 mg/kg of tremelimumab q4w for 4 doses each. Patients will then continue with MEDI4736 monotherapy at 20 mg/kg q4w, beginning at Week 16, 4 weeks after the last dose of combination therapy (a total of 9 additional doses). Patients in the MEDI4736 monotherapy treatment group will resume MEDI4736 dosing at 20 mg/kg q4w for up to 12 months (up to 13 doses).

For all treatment groups, treatment through progression is at the Investigator's discretion, and the Investigator should ensure that patients do not have any significant, unacceptable, or irreversible toxicities that indicate that continuing treatment will not further benefit the patient. A patient with confirmed progression randomized to the MEDI4736 monotherapy arm cannot continue therapy or obtain retreatment if dosing is ongoing and the progression occurs in a target lesion that has previously shown a confirmed response. A patient with a confirmed progression randomized to the MEDI4736 + tremelimumab arm cannot continue therapy or obtain retreatment if dosing is ongoing in the combination portion of therapy (q4w dosing) and progression occurs in a target lesion that has previously shown a confirmed response.

Patients who AstraZeneca and/or the Investigator determine may not continue treatment will enter follow-up.

7.3 Labeling

Labels will be prepared in accordance with Good Manufacturing Practice (GMP) and local regulatory guidelines. The labels will fulfill GMP Annex 13 requirements for labelling. Label text will be translated into the local language.

Labels will be provided as either a single panel label or as multi-language booklet labels.

7.4 Storage

All study drugs should be kept in a secure place under appropriate storage conditions. The IP label on the pack/bottle/carton specifies the appropriate storage. Storage is also described in the IB.

7.5 Compliance

The administration of all study drugs (including IP) should be recorded in the appropriate sections of the eCRF.

Treatment compliance will be assured by reconciliation of site drug accountability logs.

7.6 Accountability

The study drug provided for this study will be used only as directed in the study protocol. The study personnel will account for all study drugs

Drug accountability should be performed until the patient stops study treatment completely. Study site personnel will account for all study drugs received at the site, for all unused study drugs, and for appropriate destruction of study drugs. Certificates of delivery, destruction, and return should be signed.

7.7 Concomitant and other treatments

The Investigator must be informed as soon as possible about any medication taken from the time of screening until the end of the clinical phase of the study (final study visit). Any concomitant medication(s), including herbal preparations, taken during the study will be recorded in the eCRF.

Restricted, prohibited, and permitted concomitant medications are described in the following tables. Refer to Section 6.7 for guidance on management of IP-related toxicities.

Prohibited medication/class of drug:	**Usage:**
Additional investigational anticancer therapy concurrent with those under investigation in this study	Should not be given whilst the patient is on IP treatment (including SoC)
mAbs against CTLA-4, PD-1, or PD-L1	Should not be given whilst the patient is on IP treatment (including SoC) through 90 days after the last dose of IP.
Any concurrent chemotherapy, local therapy (except palliative radiotherapy for non-target lesions, eg, radiotherapy, surgery, radiofrequency ablation), biologic therapy, or hormonal therapy for cancer treatment	Should not be given whilst the patient is on IP treatment (including SoC). (Concurrent use of hormones for non-cancer-related conditions [eg, insulin for diabetes and hormone replacement therapy] is acceptable.)
Immunosuppressive medications, including, but not limited to: systemic corticosteroids at doses exceeding 10 mg/day of prednisone or its equivalent, methotrexate, azathioprine, and TNF-α blockers	Should not be given whilst the patient is on IP treatment (including SoC). (Use of immunosuppressive medications for the management of IP-related AEs or in patients with contrast allergies is acceptable. In addition, use of inhaled, topical, and intranasal corticosteroids is permitted.
Live attenuated vaccines	Should not be given through 30 days after the last dose of IP (including SoC) during the study

Rescue/supportive medication/class of drug:	**Usage:**
Concomitant medications or treatments (eg, acetaminophen or diphenhydramine) deemed necessary by the Investigator to provide adequate prophylactic or supportive care, except for those medications identified as "prohibited" as listed above	To be administered as prescribed by the Investigator

Rescue/supportive medication/class of drug:	Usage:
Best supportive care (including antibiotics, nutritional support, growth factor support, correction of metabolic disorders, optimal symptom control, and pain management [including palliative radiotherapy, etc])	Should be used when necessary for all patients

7.7.1 Other concomitant treatment

Medications other than those described in Section 7.7 that are considered necessary for the patient's safety and well-being may be given at the discretion of the Investigator and should be recorded in the appropriate sections of the eCRF.

7.8 Post study access to study treatment

After the final analysis, AstraZeneca will continue to supply open-label drug to patients receiving MEDI4736 + tremelimumab or MEDI4736 monotherapy up to completion of a patient's current 12-month treatment period (initial or repeat) (see Section 7.2.2).

8. STATISTICAL ANALYSES BY ASTRAZENECA

8.1 Statistical considerations

All statistical analyses will be performed by AstraZeneca or its representatives.

A comprehensive statistical analysis plan (SAP) will be prepared and finalized within 3 months of the first randomized patient and any subsequent amendments will be documented, with final amendments completed prior to reporting of the data. The primary aim of the study is to compare the efficacy and safety of MEDI4736 in combination with tremelimumab to SoC.

8.2 Sample size estimate

The study will plan to enroll approximately 810 patients in order to randomize 675 eligible patients 1:1:1 to MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC. The 675 patients will comprise at a minimum 438 patients who have PD-L1-negative tumors.

The study is sized to characterize the PFS benefit of MEDI4736 in combination with tremelimumab versus SoC in patients with EGFR and ALK wild-type locally advanced or metastatic NSCLC (ie, regardless of PD-L1 tumor expression status) and patients with PD-L1–negative tumors. The sizing assumes a 3-month delay in separation of the PFS curves between each arm, hence the use of average hazard ratios (HRs).

The primary analysis of PFS will be performed when

- (approximately) 338 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups (75% maturity) AND

- (approximately) 223 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 negative tumors (76% maturity)

MEDI4736 + tremelimumab versus SoC (primary endpoint)

If PFS at 12 months was 33.5% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average PFS HR is 0.59, the trial will have >90% power to demonstrate statistical significance at the 5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.81. With a 15-month recruitment period and a minimum follow up period of 6.5 months assumed, it is anticipated that this analysis will be performed 21.5 months after the first patient has been recruited.

***MEDI4736 + tremelimumab versus SoC (PD-L1–negative population*)**

If the boundary is crossed for the primary PFS analysis, then PFS in the PD-L1–negative population of MEDI4736 + tremelimumab versus SoC will be tested at the 5% alpha level (in line with the hierarchical testing strategy). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and SoC treatment groups and a true average PFS HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 90% power to demonstrate statistical significance at the 5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.77.

MEDI4736 monotherapy versus SoC (ITT and PD-L1–positive populations)

If the boundary is crossed for the primary PFS analysis and for PFS of MEDI4736 + tremelimumab versus SoC in the PD-L1–negative population, then a statistical analysis of MEDI4736 monotherapy versus SoC will be performed in all randomized patients as well as in patients with PD-L1-positive tumors (in line with the hierarchical testing strategy):

- **ITT population:** With approximately 450 patients randomized across the MEDI4736 monotherapy and SoC treatment groups and a true average PFS HR of 0.59, an estimated 338 progression/death events (75% maturity) are expected to have occurred at 21.5 months from "first patient in," providing >90% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.78.

- **PD-L1-positive population:** With approximately 158 patients with PD-L1-positive tumors randomized across the MEDI4736 monotherapy and SoC treatment groups and an true average PFS HR of 0.51, an estimated 114 progression/death events (72% maturity) are expected to have occurred at 21.5 months from "first patient in," providing >90% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.65.

MEDI4736 + tremelimumab versus MEDI4736 monotherapy (PD-L1–negative population)

If the boundary is crossed for the primary PFS analysis and for PFS of MEDI4736 + tremelimumab versus SoC in the PD-L1–negative population, then a statistical analysis of MEDI4736 + tremelimumab versus MEDI4736 monotherapy will be performed in the PD-L1-negative population (in line with the hierarchical testing strategy). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups, if PFS at 12 months was 30% with MEDI4736 + tremelimumab and 13% with MEDI4736 monotherapy (~5.8-month median PFS) and assuming an true average PFS HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 86% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that would be statistically significant being an average HR of 0.74.

MEDI4736 + tremelimumab versus SoC (overall survival)

If OS at 18 months was 49% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9-month median OS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average OS HR is 0.70, the trial will have approximately 84% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.78. With a 15-month additional follow up after the final PFS primary analysis, it is anticipated that this analysis will be performed 36.5 months after the first patient has been recruited.

8.3 Definitions of analysis sets

Definitions of the analysis sets for each outcome variable are provided in Table 10.

Table 10 Summary of outcome variables and analysis populations

Outcome variable	Population
Efficacy data	
PFS	Full analysis set (ITT population)
ORR, DoR, APF12, PFS2, OS, PROs, and symptom endpoints	Full analysis set (ITT population)

Outcome variable	Population
PFS, ORR	PD-L1-negative analysis set
PFS	PD-L1-positive analysis set
Demography	Full analysis set (ITT population)
PK data	PK analysis Set
Safety Data	
Exposure	Safety analysis Set
AEs	Safety analysis Set
Laboratory measurements	Safety analysis Set
Vital signs	Safety analysis Set
ECGs	Safety analysis Set

8.3.1 Full analysis set

The full analysis set (FAS) will include all randomized patients. The full analysis set will be used for all efficacy analyses (including PROs). Treatment groups will be compared on the basis of randomized study treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the analysis in the treatment group to which they were randomized.

8.3.2 PD-L1-negative analysis set

The PD-L1-negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 negative as defined by an IHC assay developed by (ie, <25% PD-L1–membrane expression in tumoral tissue).

8.3.3 PD-L1-positive analysis set

The PD-L1-positive analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive as defined by an IHC assay developed by (ie, ≥25% PD-L1–membrane expression in tumoral tissue).

8.3.4 Safety analysis set

The safety analysis set (SAS) will consist of all patients who received at least 1 dose of study treatment. Safety data will not be formally analyzed but summarized using the safety analysis set, according to the treatment received, that is, erroneously treated patients (e.g., those randomized to treatment A but actually given treatment B) will be summarized according to the treatment they actually received.

8.3.5 Pharmacokinetic analysis set

All patients who received at least 1 dose of IP per the protocol for whom any post-dose data are available and who do not violate or deviate from the protocol 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.

8.4 Outcome measures for analyses

8.4.1 Calculation or derivation of efficacy variables

The analysis of the primary endpoint, PFS, and the analyses of the secondary endpoints, ORR, DoR, and APF12, will be based on Investigator tumor assessments according to RECIST 1.1. In addition, time to secondary progression (PFS2) will be defined by local clinical practice and OS will also be evaluated.

A BICR will be performed on a sample of patients in order to perform a sensitivity analysis of PFS. In addition, PFS and ORR by irRECIST criteria using BICR assessments will also be performed for exploratory purposes on this patient subset.

8.4.1.1 RECIST 1.1-based endpoints

Investigator RECIST 1.1-based assessments

All RECIST assessments, 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.

At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD, or PD depending on the status of their disease compared with baseline and previous assessments. Baseline will be assessed within the 28 days prior to randomization. If a patient has had a tumor assessment that cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE; unless there is evidence of progression in which case the response will be assigned as PD).

Please refer to Appendix F for the definitions of CR, PR, SD, and PD.

Blinded Independent Central Review of RECIST 1.1-based assessments

A BICR of radiological scans will be performed on a random sample of patients to confirm the robustness of the PFS endpoint. This will be prespecified in the statistical analysis plan/independent review charter, which will outline the percentage of patients to perform the sample BICR on, the method used to identify the subset of patients, the method for comparing the PFS results obtained by local review with the PFS results of the sample BICR, and the criteria for determining whether BICR should be performed on all patients. If bias cannot be excluded based upon the sample BICR assessment, then an independent evaluation of all radiographic images may be required for assessment of the primary PFS endpoint.

All images will be collected centrally. Prior radiotherapy will also be provided to the BICR to allow the selection of appropriate target lesions. The imaging scans will be reviewed by 2 independent radiologists using RECIST 1.1 and will be adjudicated, if required. For each patient, the BICR will define the overall visit response data (CR, PR, SD, PD, or NE) and the relevant scan dates for each time point (ie, for visits where response or progression is/is not identified). 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). Endpoints (of PFS) will be derived from the overall visit response date and the scan dates. If there are significant discrepancies between the site and BICR evaluations, a BICR of all patients will be performed.

PFS and ORR by irRECIST criteria using BICR assessments will also be performed for exploratory purposes.

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

Primary endpoint - progression-free survival

PFS (per RECIST 1.1 using Investigator assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression. 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. If the patient has no evaluable visits or does not have baseline data, they will be censored at 0 days unless they die within 2 visits of baseline.

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 investigational assessments, the date of progression will be determined based on the earliest of the RECIST assessment/scan dates of the component that indicates 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 target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the target lesions, and similarly for non-target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the non-target lesions.

A sensitivity analysis of PFS will be performed using BICR assessments according to RECIST 1.1.

PFS based on RECIST 1.1 modified for confirmation of progression will be performed for exploratory purposes using the algorithm described above for the RECIST 1.1 Investigator assessments, but following a modification whereby any objective disease progression must be confirmed by the next scheduled scan. The confirmatory scan must be no sooner than 4 weeks after the initial suspected progression. If disease progression is confirmed (or disease progression occurs and no further scans are recorded) then the date of progression will be when it was originally observed. Patients with a single disease progression and no further tumor assessment scans will be treated as PD in the analysis. In the absence of significant clinical deterioration, the investigational site is advised to continue the patient on their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy until progression has been confirmed. If progression is not confirmed, the patient should continue their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy treatment and on-treatment assessments. Treatment through PD in the SoC group is at the Investigator's discretion; however, a confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

PFS by irRECIST criteria using BICR assessments will also be reported.

8.4.1.2 Secondary endpoints

Objective response rate

ORR (per RECIST 1.1 using Investigator assessments) is defined as the number (%) of patients with at least 1 visit response of CR or PR. 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 by irRECIST criteria using BICR assessments will also be reported.

Duration of response

DoR (per RECIST 1.1 using Investigator assessments) will be defined as the time from the date of first documented response until the first date of documented progression or death in the absence of disease progression. 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 CR or PR. 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.

Time from randomization to second progression (PFS2)

PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death. The date of second progression will be recorded by the Investigator in the eCRF and defined according to local standard clinical practice and may involve any of the following: objective radiological

imaging, symptomatic progression, or death. Second progression status will be reviewed (every 6 weeks for the first 48 weeks relative to the date of randomization and then every 8 weeks thereafter) following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression, that is, censored at the latest of the PFS or PFS2 assessment date if the patient has not had a second progression or death.

Proportion of patients alive and progression free at 12 months

The APF12 will be defined as the Kaplan-Meier estimate of PFS (per RECIST 1.1 as assessed using Investigator assessments) at 12 months.

Overall survival

OS is defined as the time from the date of randomization until death due to any cause. 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.

Note: Survival calls will be made following the date of data cut-off for the analysis (these contacts should generally occur within 7 days of the data cut off). If patients are confirmed to be alive or if the death date is post the data cut-off date, these patients will be censored at the date of data cut-off. Death dates may be found by checking publicly available death registries.

Best objective response

BoR is calculated based on the overall visit responses from each RECIST assessment, described in Appendix F. It is the best response a patient has had during their time in the study up until RECIST progression or the last evaluable assessment in the absence of RECIST progression.

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

BoR will be determined programmatically based on RECIST 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 assessments prior to death.

For patients who die with no evaluable RECIST assessments, if the death occurs ≤90 days (ie, 2*(6 weeks ±3 days)) after randomization, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST assessments, if the death occurs >90 days (ie, 2*(6 weeks ±3 days)) after the date of randomization then BoR will be assigned to the NE category.

Progression events that have been censored due to them being >90 days after the last evaluable assessment will not contribute to the BoR derivation.

8.4.2 Calculation or derivation of safety variables

8.4.2.1 Adverse events

Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of Medical Dictionary for Regulatory Activities [MedDRA] preferred terms and CTCAE grade) will be listed individually by patient.

Any AE occurring before treatment with IP will be included in the data listings but will not be included in the summary tables of AEs. Any AE occurring within 90 days of discontinuation of IP (ie, the last dose of MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC) may be included in the AE summaries, but the majority of those summaries will omit those AEs observed after a patient has received further therapy for cancer. Further details will be provided in the SAP. Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of IP) will be flagged in the data listings.

A separate data listing of AEs occurring more than 90 days after discontinuation of IP will be produced. These events will not be included in AE summaries.

8.4.2.2 Other significant adverse events

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs and AEs leading to discontinuation. Based on the expert's judgment, significant AEs of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered other significant adverse events (OAEs) and reported as such in the CSR. A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs. Examples of these are marked hematological and other laboratory abnormalities, and certain events that lead to intervention (other than those already classified as serious) or significant additional treatment.

8.4.2.3 Safety assessments

For the change from baseline summaries for vital signs, laboratory data, ECGs, and physical examination, the baseline value will be the latest result obtained prior to the start of study treatment.

The QTcF will be derived during creation of the reporting database using the reported ECG values (RR and QT).

QTcF = QT/RR^ (1/3) where RR is in seconds

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

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

The denominator used in laboratory summaries will only include evaluable patients, in other words those who had sufficient data to have the possibility of an abnormality.

For example:

- If a 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 CTCAE criterion does not consider changes from baseline, to be evaluable the patient need only have 1 post dose-value recorded.

The denominator in vital signs data should include only those patients with recorded data.

8.4.3 Calculation or derivation of patient-reported outcome variables

PRO questionnaires will be assessed using the EORTC QLQ-C30 with the QLQ-LC13 module (HRQoL and lung cancer specific symptoms), PRO-CTCAE, PGIC, and EQ-5D-5L. All items/questionnaires will be scored according to published scoring guidelines or the developer's guidelines, if published guidelines are not available. All PRO analyses will be based on the Full Analysis Set (FAS; ITT population), unless stated.

8.4.3.1 EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea/vomiting), 5 individual items (dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a global measure of health status. The EORTC QLQ-C30 will be scored according to the EORTC QLQ-C30 scoring manual (Fayers et al 2001). An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, each of the functional scales, and the global health status scale in the EORTC QLQ-C30 according to the EORTC QLQ-C30 Scoring Manual. Higher scores on the global health status and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity.

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥10 for scales/items from the EORTC QLQ-C30 (Osoba et al 1998). For example, a clinically meaningful improvement in physical function (as assessed by EORTC QLQ-C30) is defined as an increase in the score from baseline of ≥10, whereas a clinically meaningful deterioration is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, the change in symptoms/functioning from baseline will be categorized as improvement, no change or deterioration as shown in Table 11.

Table 11 Mean change and visit response in health-related quality of life

Score	Change from baseline	Visit response
EORTC QLQ-C30 Global quality of life score	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change
EORTC QLQ-C30 symptom scales/items	≥+10	Deterioration
	≤-10	Improvement
	Otherwise	No change
EORTC QLQ-C30 functional scales	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change

For each subscale, if <50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded. If there is evidence that the missing data are systematic, missing values will be handled to ensure that any possible bias is minimized.

Time to HRQoL/function deterioration

For the following HRQoL items of the EORTC QLQ-C30, time to deterioration will be analyzed:

The Global Health Status/ QoL scale consisting of items 29 and 30 of the EORTC QLQ C30. Item 29: How would you rate your overall health during the past week? Item 30: How would you rate your overall quality of life during the past week? Patients are asked to rate their overall health and overall quality of life on a scale from 1 (very poor) to 7 (excellent).

Time to deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful deterioration (a decrease in the function scales or the global health status/HRQoL from baseline of ≥10) that is confirmed at a subsequent visit or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to HRQoL/function deterioration. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the HRQoL/function change could be evaluated.

Patients whose HRQoL (as measured by EORTC QLQ-C30) has not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the HRQoL/function could be evaluated. Also, if HRQoL deteriorates after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated. If a patient has no evaluable visits or does not have baseline data they will be censored at 0 days.

Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥10 for EORTC QLQ-C30 symptom scales) in that symptom from baseline.

HRQoL/function improvement rate

The HRQoL/function improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (an increase from baseline score ≥10 for EORTC QLQ-C30 functional scales and global health status/HRQoL) in that scale from baseline.

8.4.3.2 Lung cancer module (EORTC QLQ-LC13)

The QLQ-LC13 is a lung cancer specific module from the EORTC for lung cancer comprising 13 questions to assess lung cancer symptoms (cough, hemoptysis, dyspnea, and site-specific pain), treatment-related side-effects (sore mouth, dysphagia, peripheral neuropathy, and alopecia), and pain medication.

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥5 for scales/items from the QLQ-LC13 (Osoba et al 1998). For example, a clinically meaningful deterioration or worsening in chest pain (as assessed by the QLQ-LC13) is defined as an increase in the score from baseline of ≥5, whereas a clinically meaningful improvement is defined as a decrease in the score from baseline of ≥5. At each post-baseline assessment, the change in symptoms from baseline will be categorized as an improvement, no change or deterioration as shown in Table 12.

Table 12 Visit response for health-related quality of life and disease-related symptoms

Score	Change from baseline	Visit response
QLQ-LC13 symptom scales/items	≥+5	Deterioration
	≤-5	Improvement
	Otherwise	No change

Time to symptom deterioration

For each of the following key symptom scales/items in the QLQ-LC13, time to deterioration will be analyzed:

- Dyspnea (multi-item scale based on 3 questions: were you short of breath when you rested; walked; climbed stairs)
- Cough: 1 item (how much did you cough?)
- Hemoptysis: 1 item (did you cough up blood?)
- Pain (3 individual items): a) Have you had pain in your chest; b) your arm or shoulder; c) other parts of your body?)

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales/single items of the EORTC QLQ-C30.

Time to symptom deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful symptom deterioration (an increase in the score from baseline of ≥5) that is confirmed at a subsequent visit or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to symptom deterioration. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the symptom change could be evaluated.

Patients whose symptoms (as measured by QLQ-LC13) have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the symptom could be evaluated. Also, if symptoms deteriorate after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where the symptom could be evaluated. If a patient has no evaluable visits or does not have baseline data they will be censored at 0 days.

Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥5 for QLQ-LC13 symptom scales/items) in that symptom from baseline.

8.4.3.3 Calculation or derivation of healthy state utility (EQ-5D-5L)

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied (Oemar and Janssen 2013). In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

8.4.4 Calculation or derivation of pharmacokinetic variables

8.4.4.1 Population pharmacokinetics and exposure-response/safety analysis

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. The results of such an analysis will be reported in a separate report. The PK, pharmacodynamic, 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.

8.4.4.2 Pharmacokinetic non-compartmental analysis

The PK analyses will be performed at AstraZeneca. The actual sampling times will be used in the PK calculations. PK concentration data and summary statistics will be tabulated. Individual and mean blood concentration-time profiles will be generated. PK parameters will be determined using standard non-compartmental methods. The following PK parameters will be determined after the first and steady-state doses: peak and trough concentration (as data allow). Samples below the lower limit of quantification will be treated as missing in the analyses.

8.4.4.3 Immunogenicity analysis

Immunogenicity results will be analyzed descriptively by summarizing the number and percentage of patients who develop detectable ADAs against MEDI4736 and tremelimumab. The immunogenicity titer and presence of neutralizing ADAs will be reported for samples confirmed positive for the presence of ADAs. The effect of immunogenicity on PK, PDx, efficacy, and safety will be evaluated, if the data allow.

8.4.5 Calculation or derivation of biomarker variables

Biomarker status, as defined in the secondary objectives, will be assessed according to pre-specified criteria that will be detailed in the SAP.

8.5 Methods for statistical analyses

The formal statistical analysis will be performed to test the main hypotheses:

- H0: No difference between MEDI4736 + tremelimumab and SoC
- H1: Difference between MEDI4736 + tremelimumab and SoC

The primary endpoint is PFS in all patients regardless of PD-LI status using Investigator assessments per RECIST 1.1. The study has been sized to characterize the PFS benefit of MEDI4736 + tremelimumab versus SoC. The analysis will be performed when:

- (approximately) 338 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups (75% maturity) AND
- (approximately) 223 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 negative tumors (76% maturity) AND

Descriptive statistics will be used for all variables, as appropriate, and will be presented by treatment group. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by frequency counts and percentages for each category. Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment arm.

Baseline will be the last assessment of the variable under consideration prior to the intake of the first dose of IP, except for efficacy variables. For efficacy variables, baseline is defined as the last visit prior to randomization.

All data collected will be listed. Efficacy and PRO data will be summarized and analyzed based on the FAS. PK data will be summarized and analyzed based on the PK Analysis Set. Safety data will be summarized on the Safety Analysis Set.

All outputs will be summarized by treatment arm for all randomized patients (ITT) and where required, for all randomized patients within the PD-L1-negative and PD-L1-positive subgroups.

Results of all statistical analysis will be presented using a 95% confidence interval (CI) and 2-sided p-value, unless otherwise stated.

The following table (Table 13) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is

regarded as primary for that endpoint. Note, all endpoints compare MEDI4736 + tremelimumab versus SoC in all randomized patients (ITT population), unless otherwise indicated.

Table 13 Pre-planned statistical and sensitivity analyses to be conducted

Endpoints analyzed	Notes
Progression-free survival	Stratified log-rank tests for: Primary analysis using Investigator RECIST 1.1 assessments Secondary analysis using Investigator assessments (RECIST 1.1): – MEDI4736 + tremelimumab versus SoC for PD-L1-negative population (stratified only for histology) – MEDI4736 monotherapy versus SoC (ITT population) – MEDI4736 monotherapy versus SoC for PD-L1-positive population (stratified only for histology) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-negative population (stratified only for histology) Sensitivity analyses using BICR assessments (RECIST 1.1) for a subset of patients sampled from the ITT population for BICR Sensitivity analysis based on RECIST 1.1 modified for confirmation of progression using Investigator assessments Exploratory analysis using BICR data for irRECIST for a subset of patients sampled from the ITT population for BICR
Objective response rate	Logistic regression for: Secondary analysis for the ITT population and PD-L1-negative population using Investigator RECIST 1.1 assessments Exploratory analysis using BICR data for irRECIST for a subset of patients sampled from the ITT population for BICR
Duration of response	Analysis methods as described by Ellis et al 2008 for: Secondary analysis using Investigator assessments (RECIST 1.1)
Proportion of patients alive and progression free at 12 months	Hazard ratio using the Kaplan Meier estimates of progression free survival at 12 months (following method described by Klein et al 2007)
Time from randomization to second progression	Stratified log-rank test
Overall survival	Stratified log-rank test
Time to deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Stratified log-rank test

Multiple testing strategy

The multiple testing procedure (as shown in Figure 5) will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoint of PFS and the key secondary endpoints intended for label claims.

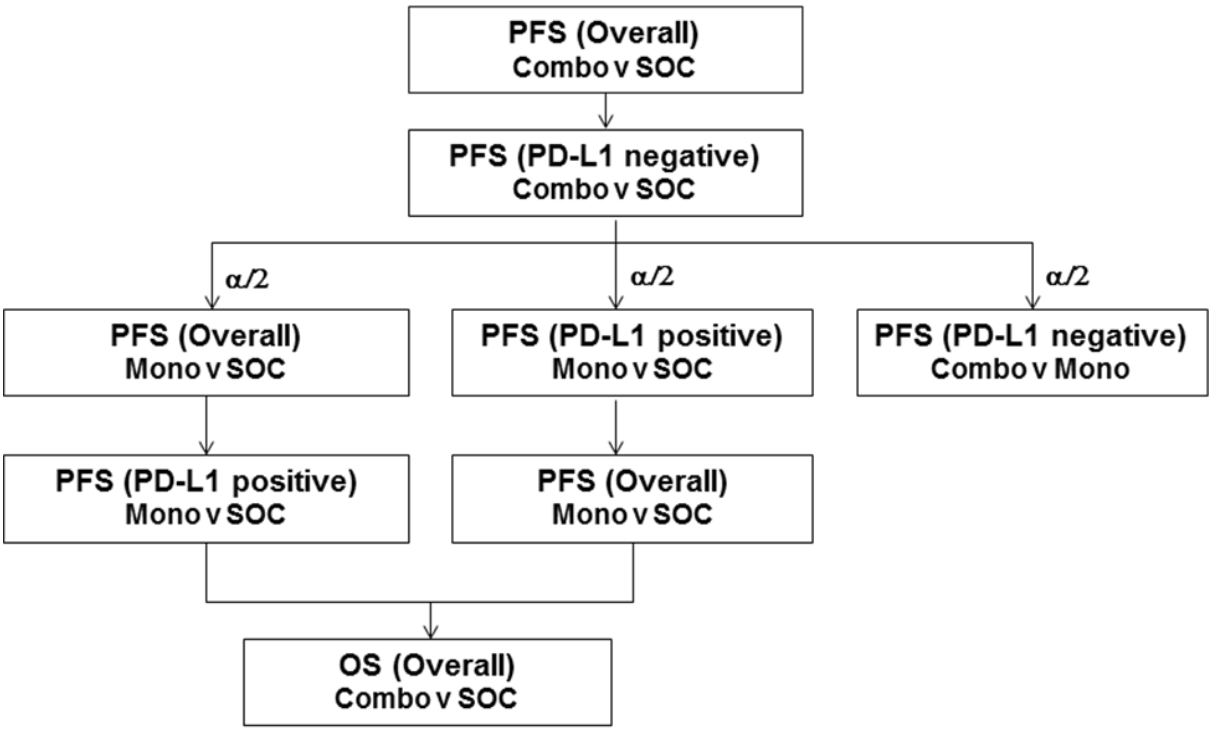
Hypotheses will be tested using a multiple testing procedure with an alpha-exhaustive recycling strategy (Burman et al 2009). With this approach, hypotheses will be tested in a pre-defined order, where PFS for MEDI4736 + tremelimumab versus SoC (ITT population) is tested first, followed by test of PFS for MEDI4736 + tremelimumab versus SoC (PD-L1 negative population). The other hypotheses will then be tested in the multiple testing procedure using alpha (test mass) splitting and alpha recycling, where the test mass that becomes available after each rejected hypothesis is recycled to secondary hypotheses not yet rejected. This testing procedure stops when the entire test mass is allocated to non-rejected hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses. Figure 5 shows the multiple testing framework.

If the tests for MEDI4736 + tremelimumab versus SoC is statistically significant at the 0.05 level in both the PD-L1 all comers and the PD-L1-negative populations, then the following 3 comparisons of (i) MEDI4736 monotherapy versus SoC in the all comer, (ii) MEDI4736 monotherapy versus SoC in the PD-L1 positive, and (iii) the MEDI4736 + tremelimumab versus MEDI4736 monotherapy in the PD-L1 negatives will be conducted, each at the 2.5% (2-sided) level.[1] The details on the alpha-exhaustive recycling procedure will be provided in the SAP.

[1] If either the null hypothesis for MEDI4736 + tremelimumab combination therapy versus SoC in the all comer (H1) or the PD-L1 negative population (H2) or both are true, then strong control of familywise error rate is maintained, since these tests will be conducted sequentially at the 0.05 level and will serve as gate keepers.

If H1 and H2 are both not true, ie, PFS in MEDI4736 + tremelimumab combination therapy is different from SoC in the all comer and the PD-L1-negative population, then the only opportunity where the overall type I error will not be preserved at the 0.05 level, by comparing the 3 hypotheses in the third level of the hierarchy at the 0.025 level, is if the null hypothesis is true in all the 3 cases; that is, there is no difference between the 2 arms in each comparison. However, given that H1 and H2 are both not true, it will not be possible for the remaining 3 hypotheses to be simultaneously true. Therefore, the overall false positive error from making these 3 comparisons at the 0.025 level should not be greater than 0.05. Hence the family-wise error rate will be maintained at the 0.05 level.

Figure 5 Multiple testing procedures for controlling the type 1 error rate



Combo MEDI4736 + tremelimumab combination therapy; Mono MEDI4736 monotherapy.

Note, 2 analyses of OS are planned:

1. At the time of the primary PFS analysis

2. A final OS analysis at 74% maturity

The alpha will be split between the 2 OS analyses using a bespoke spending function where a fixed significance level will be assigned at the first analysis and the remaining significance level assigned to the final analysis, taking account of correlation (Stone 2010):

- 1/10α available allocated to the first analysis,
- The significance level for the final analysis will be calculated using the software package EAST by selecting: boundary family, p-value, and Haybittle-Peto and specifying the fixed significance level used at the first analysis as well as the information fraction for each analysis. The information fraction is calculated as the number of events at the analysis time-point divided by the total number of events at the final analysis time-point.

Any non-statistically significant analyses at the time of the PFS analysis will not preclude further testing of OS. However, if the boundary is not crossed for the primary and key secondary endpoints, AstraZeneca may decide not to perform the final OS analysis.

8.5.1 Analysis of the primary variable - progression-free survival

The primary PFS analysis will be based on the programmatically derived RECIST 1.1 using the Investigator tumor assessments. The analysis will be performed in the ITT population using a stratified log-rank test adjusting for PD-L1 tumor expression (positive versus negative) and histology (squamous versus non-squamous). The effect of MEDI4736 + tremelimumab versus SoC treatment will be estimated by the HR together with its corresponding 95% CI and p-value.

The HR and its CI can be estimated from the stratified log-rank as follows (Berry et al 1991, Collett 2003, Selke and Siegmund 1983):

$$HR = \exp\left(\frac{U}{V}\right)$$

$$95\%\ \text{CI for HR} = \left(\exp\left\{\frac{U}{V} - \frac{1.96}{\sqrt{V}}\right\}, \exp\left\{\frac{U}{V} + \frac{1.96}{\sqrt{V}}\right\}\right)$$

Where $U = \sum_k U_k = \sum_k \sum_i (d_{1ki}, -e_{1ki})$ is the stratified log-rank test statistic obtained from the SAS LIFETEST procedure , $\sqrt{V} = \sqrt{\sum_k V_k}$, is its standard deviation, k denotes the stratum and d_{1ki} and e_{1ki} are the observed and expected events in Group 1, stratum k.

A secondary analysis of PFS will be performed to compare MEDI4736 monotherapy versus SoC as well as to compare MEDI4736 + tremelimumab versus MEDI4736 monotherapy. These analyses will be performed using the same methodology as for primary endpoint described above.

Kaplan-Meier plots of PFS will be presented by treatment arm, and by treatment arm and PD-L1 tumor status subgroup, where appropriate. Summaries of the number and percentage of patients experiencing a PFS event and the type of event (RECIST 1.1 or death) will be provided along with median PFS for each treatment.

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time points. The midpoint between the time of progression and the previous evaluable RECIST assessment will be analyzed using a log-rank test. For patients whose death was treated as PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust even in highly asymmetric assessment schedules (Sun and Chen 2010).

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will

be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the PFS analysis is reversed.

A sensitivity analysis will be performed by analyzing the BICR assessments. The stratified log-rank test will be repeated on these data. The HR and CI will be presented. The methodology and acceptance thresholds will be documented in the SAP.

A secondary analysis of PFS based on the programmatically derived RECIST 1.1 using the Investigator tumor assessments will be performed to compare MEDI4736 + tremelimumab versus SoC in the PD-L1-negative population. This analysis will be performed using a stratified log-rank test adjusting solely for histology (squamous versus non-squamous). The effect of treatment will be estimated by the HR together with its corresponding 95% CI and p-value. The HR and CI will be estimated using the same approach as specified above for the primary analysis of PFS. In addition, a secondary analysis of PFS will be performed similarly to compare MEDI4736 + tremelimumab versus MEDI4736 monotherapy in the PD-L1-negative population as well as MEDI4736 monotherapy versus SoC in the PD-L1-positive population.

An exploratory analysis of PFS using Investigator assessment based on RECIST 1.1 modified for confirmation of progression as well as PFS based on BICR assessments according to irRECIST criteria will be performed. The stratified log-rank test used for the primary analysis of PFS will be repeated.

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1 using Investigator assessments) between MEDI4736 + tremelimumab versus SoC in the following subgroups of the FAS (but not limited to):

- Sex (male versus female)
- Age at randomization (<65 versus ≥65 years of age)
- PD-L1 status (positive versus negative)
- Histology (squamous versus non-squamous)
- Smoking (smoker versus non-smoker [never smoker])
- Race (Asian versus non-Asian)

Other baseline variables may also be assessed if there is clinical justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors.

No adjustment to the significance level for testing of the subgroup and sensitivity analyses will be made since all these analyses will be considered supportive of the analysis of PFS.

Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. Before embarking on more detailed modeling, an initial model will be constructed, containing treatment and the stratification factors alone, to ensure that any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

The presence of quantitative interactions will be assessed by means of an overall global interaction test. This will be performed in the overall population by comparing the fit of a Cox proportional hazards model including treatment, all covariates, and all covariate-by-treatment interaction terms, with one that excludes the interaction terms and will be assessed at the 2-sided 10% significance level. If the fit of the model is not significantly improved, then it will be concluded that, overall, the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% significance level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. Throughout this process, all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

Any quantitative interactions identified using this procedure will then be tested to rule out any qualitative interaction using the approach of Gail and Simon 1985).

Additionally, for each subgroup, the HR (MEDI4736 + tremelimumab: SoC) and 95% CI will be calculated from a single model that contains treatment, factor (only the factor that determines the subgroup) and treatment-by-factor interaction term. These will be presented on a forest plot including the HR and 95% CI from the overall population.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events in a subgroup), the relationship between that subgroup and PFS will not be formally analyzed. In this case, only descriptive summaries will be provided.

8.5.2 Objective response rate

The ORR will be based on the programmatically derived RECIST using the Investigator tumor data. The ORR will be compared between MEDI4736 + tremelimumab versus SoC using logistic regression models adjusting for the same factors as the primary endpoint (PD-L1 tumor expression and histology). The results of the analysis will be presented in terms of an odds ratio together with its associated profile likelihood CI and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model). This analysis will be performed in the ITT population and PD-L1-negative population. The analysis of the PD-L1-negative patients will be performed using a logistic regression model adjusting for only histology.

ORR by irRECIST criteria using BICR assessments will also be reported in the ITT population.

Summaries will be produced that present the number and percentage of patients with a tumor response (CR/PR). Overall visit response data will be listed for all patients (ie, the FAS). For each treatment arm, best overall response (BoR) will be summarized by n (%) for each category (CR, PR, SD, PD, and NE). No formal statistical analyses are planned for BoR.

8.5.3 Duration of response

In order to analyze the DoR between MEDI4736 + tremelimumab and SoC, the expected duration of response (EDoR) will be derived for each treatment arm (Ellis et al 2008) using the Investigator tumor data. The EDoR is the product of the proportion of patients responding to treatment and the mean DoR in responding patients and provides an estimate based on all randomized patients. Treatments will be compared by calculating the ratio of EDoRs, using an appropriate probability distribution (to be specified in the SAP) for DoR in responding patients. Additionally, descriptive data will be provided for the DoR in responding patients, including the associated Kaplan-Meier curves (without any formal comparison of treatment arms or p-value attached). This analysis will be performed in the ITT population.

8.5.4 Proportion of patients alive and progression free at 12 months

The APF12 will be summarized (using the Kaplan-Meier curve) and presented by treatment arm. APF12 will be compared between MEDI4736 + tremelimumab and SoC by using the Kaplan-Meier estimator of PFS at 12 months for each treatment to obtain the HR. The HR and CI will be presented using the following approach (Klein et al 2007):

- The *HR(group1:group2)* is estimated as $\frac{\ln \hat{S}_1(t)}{\ln \hat{S}_2(t)}$
- The variance for *ln(HR)* is estimated as $\frac{\hat{\sigma}_1(t)^2}{\ln^2 S_1(t)} + \frac{\hat{\sigma}_2(t)^2}{\ln^2 S_2(t)}$

where $\hat{\sigma}_i(t)^2 = \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$ is the variance for ln{S(t) derived from greenwood's formula for the variance of *S(t)* and can be estimated from standard software packages, where d_i and n_i refer to the number of events and patients at risk for each risk set.

The *ln(HR)* and its variance in each strata will be estimated and combined by weighting inversely proportionately according to each within stratum variance (Whitehead and Whitehead 1991).

This analysis will be performed in the ITT population.

8.5.5 Time from randomization to second progression

Second progression (PFS2) in the ITT population will be analyzed using a stratified log-rank tests, using the same methodology as described for the primary PFS endpoint. The effect of

MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

For supportive purposes, the time to the start of subsequent therapy will be analyzed using the same methodology and model. The HR for the treatment effect together with its 95% CI will be presented. In addition, a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment arm and the time between progression and starting subsequent therapy will be assessed. No multiplicity adjustment will be applied as these are viewed as supportive endpoints.

A summary table of first subsequent therapies by treatment arm will be provided, as well as response to first subsequent therapy by treatment arm.

This analysis will be performed in the ITT population.

8.5.6 Overall survival

OS in the ITT population will be analyzed using a stratified log-rank tests, using the same methodology as described for the primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. 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 for each treatment.

8.5.7 Patient reported outcomes

8.5.7.1 EORTC QLQ-C30

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 3 symptom scales and the 5 individual symptom items will be produced. Similarly, a summary of QoL/function improvement rate for each of the 5 function scales (physical, role, emotional, cognitive, and social) and global health status/QoL will be produced.

Summaries of original and change from baseline values of each symptom scale/item, the global HRQoL score and each functional domain will be reported by visit for each treatment arm. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal item (in

terms of the proportion of patients in the categories of improvement, no change, and deterioration as defined in Section 8.4.3.1) will also be produced for each treatment arm.

8.5.7.2 EORTC QLQ-LC13

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 6 individual symptom items will be produced.

Summaries of original and change from baseline values of each symptom (dyspnea, cough, hemoptysis, chest pain, arm/shoulder pain, other pain) and each treatment-related side effect (sore mouth, dysphagia, peripheral neuropathy and alopecia) will be reported by visit for each treatment arm. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal symptom item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration as defined in Section 8.4.3.1) will also be produced for each treatment arm.

8.5.7.3 PRO-CTCAE

PRO-CTCAE data will be presented using summaries and descriptive statistics based on the FAS. Further details will be provided in the SAP.

8.5.7.4 Patients' Global Impression of Change

PGIC data will be presented using summaries and descriptive statistics based on the FAS. Further details will be provided in the SAP.

8.5.7.5 EQ-5D-5L

Descriptive statistics, graphs, and listings will be reported for health state utility values and the visual analogue scale by visit, as well as the change in these scores from baseline. To support future economic evaluations, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post-progression, and will be outlined in the payer analysis plan.

8.5.8 Healthcare resource use

An exploratory health economic analysis of hospital episodes including type of contact (hospitalization, outpatient, or day case), reason, length of stay by ward type (including intensive care unit), procedures, and tests may be undertaken to examine the impact of disease and treatment on resource use to primarily support the economic evaluation of MEDI4736 +

tremelimumab in comparison to SoC, and will be outlined in the payer analysis plan. This would include providing descriptive statistics as appropriate, including means, median, and ranges.

8.5.9 Safety data

Safety and tolerability data will be presented by treatment arm using the safety population.

Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of MedDRA preferred terms and CTCAE grade) will be listed individually by patient. The number of patients experiencing each AE will be summarized by treatment arm and CTCAE grade. Additionally, data presentations of the rate of AEs per person-years at risk may be produced. Any safety summaries examining re-treatment with MEDI4736 + tremelimumab and MEDI4736 monotherapy will be produced separately, if required.

Other safety data will be assessed in terms of physical examination, clinical chemistry, hematology, vital signs, and ECGs. Exposure to MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC will be summarized. Time on study, MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC, dose delays/interruptions and dose reductions will also be summarized. At the end of the study, appropriate summaries of all safety data will be produced, as defined in the SAP.

8.5.10 Pharmacokinetic data

PK concentration data will be listed for each patient and each dosing day, and a summary provided for all evaluable patients in the PK analysis population.

8.5.11 Immunogenicity analysis

Immunogenicity results will be listed by patient and a summary will be provided of the number and percentage of patients who develop detectable anti-MEDI4736 and anti-tremelimumab antibodies. The immunogenicity titer and neutralizing ADA data will be listed for samples confirmed positive for the presence of anti-MEDI4736 and anti-tremelimumab antibodies.

The effect of immunogenicity on PK, PDx, efficacy and safety will be evaluated if data allow.

8.5.12 Pharmacokinetic/Pharmacodynamic relationships

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modeling approach.

8.5.13 Biomarker data

The relationship of PD-L1 expression and if appropriate, other exploratory biomarkers to clinical outcomes (including but not restricted to) PFS and ORR may be presented.

PD-L1 expression determined by IHC 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.

8.5.14 Interim analysis

Interim safety monitoring will be conducted by an IDMC. Details of the plan and communication process will be provided in an IDMC Charter. No interim analysis will be performed for efficacy.

9. STUDY AND DATA MANAGEMENT BY ASTRAZENECA

9.1 Training of study site personnel

Before the first patient is enrolled in the study, an AstraZeneca representative will review and discuss the requirements of the clinical study protocol and related documents with the investigational staff and train them in any study-specific procedures and IVRS/IWRS, WBDC, and any electronic PRO systems to be utilized.

The Principal Investigator will ensure that appropriate training relevant to the study is given to all of these staff and that any new information relevant to the performance of this study is forwarded to the staff involved.

The Principal Investigator will maintain a record of all individuals involved in the study (medical, nursing, and other staff).

9.2 Monitoring of the study

During the study, an AstraZeneca representative will have regular contacts with the study site, including visits to

- Provide information and support to the Investigator(s)
- Confirm that facilities remain acceptable
- Confirm that the investigational team is adhering to the protocol, data are being accurately and timely recorded in the eCRFs, biological samples are handled in accordance with the Laboratory Manual, and study drug accountability checks are being performed
- Perform source data verification (a comparison of the data in the eCRFs with the patient's medical records at the hospital or practice and other records relevant to the study), including verification of informed consent of participating patients. This will require direct access to all original records for each patient (eg, clinic charts).

- Ensure that withdrawal of informed consent for the use of the patient's biological samples is reported, biological samples are identified and disposed of or destroyed accordingly, and the action is documented and reported to the patient

The AstraZeneca representative will be available between visits if the Investigators or other staff at the centers need information and advice about the study conduct.

9.2.1 Source data

Refer to the CSA for the location of source data.

9.2.2 Direct access to source data in Japan

The Head of the study site and the Principal Investigator/Investigator will cooperate for monitoring and audit by AstraZeneca and accept inspection by the IRB or regulatory authorities. All study documents such as raw data will be open for direct access to source data at the request of the monitor and the auditor of AstraZeneca, the IRB, or regulatory authorities.

The monitor will verify data from the eCRFs against source data before the Principal Investigator signs the eCRFs to ensure accuracy and completeness of documentation and ensure that the Principal Investigator has submitted the eCRFs to AstraZeneca.

9.2.3 Study agreements

The Principal Investigator at each center should comply with all the terms, conditions, and obligations of the CSA for this study. In the event of any inconsistency between this clinical study protocol and the CSA, the terms of clinical study protocol shall prevail with respect to the conduct of the study and the treatment of patients. In all other respects not relating to study conduct or treatment of patients, the terms of the CSA shall prevail.

Agreements between AstraZeneca and the Principal Investigator should be in place before any study-related procedures can take place or before any patients are enrolled.

9.2.4 Archiving of study documents

The Investigator will follow the principles outlined in the CSA.

9.3 Study timetable and end of study

The end of the study is defined as the "last visit of the last patient undergoing the study." The Investigator will be notified by AstraZeneca when recruitment is complete.

The study is expected to start in and end by .

The study may be terminated at individual centers if the study procedures are not being performed according to Good Clinical Practice (GCP) or if recruitment is slow. AstraZeneca may also terminate the entire study prematurely if concerns for safety arise within this study or in any other study involving MEDI4736.

9.4 Data management by AstraZeneca or delegate

Data management will be performed by a chosen vendor according to the Data Management Plan. AEs and medical/surgical history will be classified according to the terminology of the latest version of MedDRA. Medications will be classified according to the AstraZeneca Drug Dictionary. Classification coding will be performed by the chosen vendor.

The data collected through third party sources will be obtained and reconciled against study data.

Data queries will be raised for inconsistent, impossible, or missing data. All entries to the study database will be available in an audit trail.

The data will be validated as defined in the Data Management Plan. Quality control procedures will be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly. The Data Management Plan will also clarify the roles and responsibilities of the various functions and personnel involved in the data management process.

When all data have been coded, validated, signed, and locked, a clean file will be declared. Any treatment-revealing data may be added thereafter, and the final database will be locked.

Serious adverse event reconciliation

SAE reconciliation reports are produced and reconciled with the Patient Safety database and/or the investigational site.

Data management of genotype data

Any genotype data generated in this study will be stored in the AstraZeneca genotyping database or other appropriate secure system within AstraZeneca and/or a third party contracted to work with AstraZeneca to analyze samples. The results from this genetic research may be reported in the CSR for the main study or in a separate report, as appropriate.

Some or all of the clinical datasets from the main study may be merged with the genetic data in a suitable secure environment separate from the clinical database.

Data associated with human biological samples

Data associated with human biological samples will be transferred from laboratories internal or external to AstraZeneca.

10. ETHICAL AND REGULATORY REQUIREMENTS

10.1 Ethical conduct of the study

The study will be performed in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation (ICH)/GCP, applicable regulatory requirements, and the AstraZeneca policy on Bioethics and Human Biological Samples. The applicable regulatory requirements in Japan are "Good Clinical Practice for Trials on Drugs" (Ministry of Health, Labor, and Welfare [MHLW] Ordinance No. 28, 27 March 1997, partially revised by MHLW Ordinance and their related notifications).

10.2 Patient data protection

The ICF will incorporate wording that complies with relevant data protection and privacy legislation. In some cases, such wording will be in a separate accompanying document.

AstraZeneca will not provide individual genotype results to patients, their family members, their general physician, any insurance company, any employer, or any other third party, unless required to do so by law**.**

Precautions are taken to preserve confidentiality and prevent genetic data from being linked to the identity of the patient. In exceptional circumstances, however, certain individuals might see both the genetic data and the personal identifiers of a patient. For example, in the case of a medical emergency, an AstraZeneca Physician or an Investigator might know a patient's identity and might also have access to his or her genetic data. Also, regulatory authorities may require access to the relevant files. Even so, the patient's medical information and the genetic files would remain physically separate.

10.3 Ethics and regulatory review

An EC/IRB should approve the final study protocol, including the final version of the ICF and any other written information and/or materials to be provided to the patients. The Investigator will ensure the distribution of these documents to the applicable EC/IRB and to the study site staff.

The opinion of the EC/IRB should be given in writing. The Investigator should submit the written approval to AstraZeneca before enrolment of any patient into the study.

The EC/IRB should approve all advertising used to recruit patients for the study.

AstraZeneca should approve any modifications to the ICF that are needed to meet local requirements.

If required by local regulations, the protocol should be re-approved by the EC/IRB annually.

Before enrolment of any patient into the study, the final study protocol, including the final version of the ICF, should be approved by the national regulatory authority or a notification to the national regulatory authority should be approved, according to local regulations.

AstraZeneca will handle the distribution of these documents to the national regulatory authorities.

AstraZeneca will provide regulatory authorities, ECs/IRBs, and Principal Investigators safety updates or reports according to local requirements.

Each Principal Investigator is responsible for providing the EC/IRB with reports of any serious and unexpected adverse drug reactions from any other study conducted with the IP. AstraZeneca will provide this information to the Principal Investigator so that he/she can meet these reporting requirements.

10.4 Informed consent

The Principal Investigator(s) at each center will:

- Ensure that each patient is given full and adequate oral and written information about the nature, purpose, possible risk, and benefit of the study
- Ensure that each patient is notified that he or she is free to discontinue from the study at any time
- Ensure that each patient is given the opportunity to ask questions and allowed time to consider the information provided
- Ensure that each patient provides a signed and dated informed consent before conducting any procedure specifically for the study
- Ensure that the original, signed ICF(s) is/are stored in the Investigator's Study File
- Ensure that a copy of the signed ICF is given to the patient
- Ensure that any incentives for patients who participate in the study as well as any provisions for patients harmed as a consequence of study participation are described in the ICF that is approved by an EC/IRB.

For sites in Japan only

If any new information on the study medication becomes available that may influence the decision of the patient to continue the study, the Investigator should inform the patient of such information immediately, record this in a written form, and confirm with the patient if he or she wishes to continue the participation in the study. In addition, if the Investigator deems it necessary to revise the ICF, he or she should revise it immediately (refer to Section 10.5). The Investigator should re-explain to the patients using the updated ICF even if the patients

have already been informed of the new information verbally. Written informed consent to continue participation in the study should be provided separately.

10.5 Changes to the protocol and informed consent form

For sites outside Japan

Study procedures will not be changed without the mutual agreement of the Principal Investigator and AstraZeneca.

If there are any substantial changes to the study protocol, then these changes will be documented in a study protocol amendment and, where required, in a new version of the study protocol (revised clinical study protocol).

The amendment is to be approved by the relevant EC/IRB and, if applicable, the national regulatory authority, before implementation. Local requirements are to be followed for revised protocols.

AstraZeneca will distribute any subsequent amendments and new versions of the protocol to each Principal Investigator. For distribution to EC/IRB, see Section 10.3.

If a protocol amendment requires a change to a center's ICF, AstraZeneca and the center's EC/IRB are to approve the revised ICF before the revised form is used.

If required by local regulations, any administrative change will be communicated to or approved by each EC/IRB.

10.6 Audits and inspections

Authorized representatives of AstraZeneca, a regulatory authority, or an EC/IRB may perform audits or inspections at the center, including source data verification. The purpose of an audit or inspection is to systematically and independently examine all study-related activities and documents, to determine whether these activities were conducted, and to determine if data were recorded, analyzed, and accurately reported according to the protocol, GCPs, ICH guidelines, and any applicable regulatory requirements. The Investigator will contact AstraZeneca immediately if contacted by a regulatory agency about an inspection at the center.

11. LIST OF REFERENCES

Aaronson et al 1993
Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365-76.

Antonia et al 2014a
Antonia S, Goldberg S, Balmanoukian A, Narwal R, Robbins P, D'Angelo G, et al. A Phase 1 open-label study to evaluate the safety and tolerability of MEDI4736, an anti-programmed cell death-ligand 1 (PD-L1) antibody, in combination with tremelimumab in patients with advanced non-small cell lung cancer (NSCLC). Poster presented at European Society of Medical Oncology (ESMO) Meeting; 2014 Sep 26-30; Madrid, Spain.

Antonia et al 2014b
Antonia S, Ou SI, Khleif SN, Brahmer J, Blake-Haskins A, Robbins PB, et al. Clinical activity and safety of MEDI4736, an anti-programmed cell death ligand-1 (PD-L1) antibody in patients with NSCLC. Poster presented at the European Society for Medical Oncology (ESMO) Meeting; 2014 Sep 26-30; Madrid, Spain.

Bergman et al 1994
Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M; EORTC Study Group on Quality of Life. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. Eur J Cancer 1994;30A(5):635-42.

Berry et al 1991
Berry G, Kitchin RM, Mock PA. A comparison of 2 simple hazard ratio estimators based on the logrank test. Stat Med 1991;10(5):749-55.

Bonomi 2010
Bonomi PD. Implications of key trials in advanced nonsmall cell lung cancer. Cancer 2010;116(5):1155-64.

Brahmer et al 2012
Brahmer JR, Tykodi SS, Chow LQM, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med 2012;366(26):2455-65.

Brahmer et al 2014
Brahmer JR, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA et al. Nivolumab (anti-PD-1, BMS-936558, ONO-4538) in patients (pts) with advanced non-small-cell lung cancer (NSCLC): Survival and clinical activity by subgroup analysis. Journal of Clinical Oncology, 2014 ASCO Annual Meeting Abstracts; 32(15) (2014 May 20 Supplement): 8112.

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009;28:739-61.

Ciuleanu et al 2009
Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomized, double-blind, phase 3 study. Lancet 2009;374(9699):1432-40.

Collett 2003
Collett D. Modelling survival data in medical research: 2nd ed. Chapman and Hall/CRC; 2003.

D'Addario et al 2010
D'Addario G, Fruh M, Reck M, Baumann P, Klepetko W, Felip E, et al. Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2010;21(Suppl 5):116-9.

Drake et al 2013
Drake CG, McDermott DF, Sznol M, Choueiri TK, Kluger HM, Powderly JD et al. Survival, safety, and response duration results of nivolumab (Anti-PD-1; BMS-936558; ONO-4538) in a phase I trial in patients with previously treated metastatic renal cell carcinoma (mRCC): Long-term patient follow-up. J Clin Oncol 31, 2013 (suppl; abstr 4514).

Dunn et al 2004
Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annu Rev Immunol 2004;22:329-60.

Ellis et al 2008
Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008;29(4):456-65.

EuroQol Group 1990
EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208.

EuroQol Group 2013
EuroQol Group. EQ-5D-5L user guide: basic information on how to use the EQ-5D-5L instrument, version 2.0, October 2013. Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L_v2.0_October_2013.pdf. Accessed 21 November 2013.

Fairman et al 2014
Fairman D, Narwal R, Liang M, Robbins PB, Schneider A, Chavez C, et al. Pharmacokinetics of MEDI4736, a fully human anti-PDL1 monoclonal antibody, in patients with advanced solid tumours. Journal of Clinical Oncology, 2014 ASCO Annual Meeting Abstracts;32(5s): (suppl; abstr 2602).

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer: 2001.

FDA Guidance 2011
Food and Drug Administration. Guidance for Industry: Clinical trial endpoints for the approval of non-small cell lung cancer drugs and biologics. June 2011. Available from: URL: http://www.fda.gov/downloads/Drugs/Guidances/UCM259421.pdf. Accessed 21 May 2014.

Forde et al 2014
Forde PM, Kelly RJ, Brahmer JR. New strategies in lung cancer: translating immunotherapy into clinical practice. Clin Cancer Res 2014;20(5):1067-73.

Gail and Simon 1985
Gail M, Simon R. Tests for qualitative interactions between treatment effects and patient subsets. Biometrics 1985;41(2):361–72.

GLOBOCAN 2012
GLOBOCAN. Lung cancer, estimated incidence, mortality and prevalence worldwide in 2012. International Agency for Research on Cancer, World Health Organization; Lyon, 2012. Available at http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed on 16 April 2015.

Gong et al 2009
Gong AY, Zhou R, Hu G, Li X, Splinter PL, O'Hara SP, et al. MicroRNA-513 regulates B7-H1 translation and is involved in IFN-gamma-induced B7-H1 expression in cholangiocytes. J Immunol 2009;182(3):1325-33.

Herbst et al 2013
Herbst RS, Gordon MS, Fine GD, Sosman JA, Soria JC, Hamid O, et al. A study of MPDL3280A, an engineered PD-L1 antibody in patients with locally advanced or metastatic tumours [abstract]. J Clin Oncol 2013;31(Suppl 15):Abstract 3000.

Herdman et al 2011
Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727-36.

Hirano et al 2005
Hirano F, Kaneko K, Tamura H, Dong H, Wang S, Ichikawa M, et al. Blockade of B7-H1 and PD-1 by monoclonal antibodies potentiates cancer therapeutic immunity. Cancer Res 2005;65(3):1089-96.

Hodi et al 2010
Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanan JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma [published erratum appears in N Engl J Med 2010;363(13):1290]. N Engl J Med 2010;363(8):711-23.

Hodi et al 2014
Hodi FS, Sznol M, Kluger HM, McDermott DF, Carvajal RD, Lawrence DP et al. Long-term survival of ipilimumab-naive patients (pts) with advanced melanoma (MEL) treated with nivolumab (anti-PD-1, BMS-936558, ONO-4538) in a phase I trial. J Clin Oncol, 2014 ASCO Annual Meeting Abstracts; 32(15) (2014 May 20 Supplement): 9002.

Howlander et al 2014
Howlander N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, et al. SEER cancer statistics review, 1975-2011. Bethesda (MD): National Cancer Institute,. Available from: URL:http://seer.cancer.gov/csr/1975_2011/ based on November 2013 SEER data submission (posted to the SEER web site, April 2014).

IASLC Staging Manual in Thoracic Oncology
Goldstraw P, ed. Staging manual in thoracic oncology. 7th ed. IASLC 2010. (Available on request)

Iwai et al 2002
Iwai Y, Ishida M, Tanaka Y, Okazaki T, Honjo T, Minato N. Involvement of PD-L1 on tumour cells in the escape from host immune system and tumour immunotherapy by PD-L1 blockade. Proc Natl Acad Sci USA 2002;99(19):12293-7.

Janssen et al 2008a
Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. Value Health 2008;11(2):275-84.

Janssen et al 2008b
Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. Qual Life Res 2008;17(3):463-73.

Kirkwood et al 2010
Kirkwood JM, Lorigan P, Hersey P, Hauschild A, Robert C, McDermott D, et al. Phase II trial of tremelimumab (CP-675,206) in patients with advanced refractory or relapsed melanoma. Clin Cancer Res 2010;16(3):1042-8.

Kitano et al 2014
Kitano S, Postow MA, ZieglerCG, Kuk D, Panageas KS, Cortez C, et al. Computational algorithm-driven evaluation of monocytic myeloid-derived suppressor cell frequency for prediction of clinical outcomes. Cancer Immunol Res 2014;2(8):812-21.

Klein et al 2007
Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007;26(24):4505-19.

Korn et al 2008
Korn EL, Liu PY, Lee SJ, Chapman JA, Niedzwiecki D, Suman VJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. J Clin Oncol 2008;26(4):527-34.

Litwin et al 1998
Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. J Urol 1998;159:1988-92.

Meyer et al 2014
Meyer C, Cagnon L, Costa-Nunes Cm, Baumgaaertner P, Montandon N, Leyvraz L, et al. Frequencies of circulating MDSC correlate with clinical outcome of melanoma patients treated with ipilimumab. Cancer Immunol Immunother 2014;63(3):247-57.

Mu et al 2011
Mu CY, Huang JA, Chen Y, Chen C, Zhang XG. High expression of PD-L1 in lung cancer may contribute to poor prognosis and tumour cells immune escape through suppressing tumour infiltrating dendritic cells maturation. Med Oncol 2011;28(3):682-8.

NCCN 2014
National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines®). Non-Small Cell Lung Cancer. Version 4.2014. www.nccn.org.

Nishino et al 2013
Nishino M, Biobbie-Hurder A, Gargano M, Suda M, Ramaiya NH, Hodi FS. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. Clin Cancer Res 2013;19(14):3936-43.

Oemar and Janssen 2013
Oemar M, Oppe M. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument V5.0 (October 2013). Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

Okudaira et al 2009
Okudaira K, Hokari R, Tsuzuki Y, Okada Y, Komoto S, Watanabe C, et al. Blockade of B7-H1 or B7-DC induces an anti-tumour effect in a mouse pancreatic cancer model. Int J Oncol 2009;35(4):741-9.

Osoba et al 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16(1):139-44.

Pardoll 2012
Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12(4):252-64.

Paz-Ares et al 2013
Paz-Ares LG, de Marinis F, Dediu M, Thomas M, Pujol JL, Bidoli P, et al. PARAMOUNT: Final overall survival results of the phase III study of maintenance pemetrexed versus placebo immediately after induction treatment with pemetrexed plus cisplatin for advanced nonsquamous non-small-cell lung cancer. J Clin Oncol 2013;31(23):2895-902.

Pazdur 2008
Pazdur R. Endpoints for assessing drug activity in clinical trials. Oncologist 2008;13(Suppl 2):19-21.

Peggs et al 2009
Peggs KS, Quezada SA, Allison JP. Cancer immunotherapy: co-stimulatory agonists and coinhibitory antagonists. Clin Exp Immunol 2009;157:9-19.

Pickard et al 2007
Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Med Care 2007;45(3):259-63.

Pisters and LeChevalier 2005
Pisters KM, LeChevalier T. Adjuvant chemotherapy in completely resected non-small-cell lung cancer. J Clin Oncol 2005;23(14):3270-8.

Reck et al 2014
Reck M, Popat S, Reinmuth N, De Ruysscher D, Kerr KM, Peters S; ESMO Guidelines Working Group. Metastatic non-small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014;25(Suppl 3):iii27-39.

Ribas et al 2013
Ribas A, Kefford R, Marshall MA, Punt CJA, Haanen JB, Marmol M, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. J Clin Oncol 2013;31:616-22.

Sandler et al 2006
Sandler A, Gray R, Perry MC, Brahmer J, Schiller JH, Dowlati A, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med 2006;355(24):2524-50.

Scagliotti et al 2008
Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008;26(21)3543-51.

Schadendorf et al 2013
Schadendorf D, Hodi FS, Robert C et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in metastatic or locally advanced, unresectable melanoma. Presented at: European Cancer Congress 2013 (ECCO-ESMO-ESTRO); September 27-October 1, 2013; Amsterdam, The Netherlands. Abstract 24.

Schiller et al 2002
Schiller JH, Harringdon D, Belani CP, Langer C, Sandler A, Krook J, et al. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. N Engl J Med 2002;10(346):92-8.

Schwarzenbach et al 2014
Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical evidence of circulating cell-free microRNAs in cancer. Nat Rev Clin Oncol 2014;11(3):145-56.

Selke and Siegmund 1983
Selke T, Siegmund D. Sequential analysis of the proportional hazards model. Biometrika 1983;70:315-26.

Sprangers and Aaronson 1992
Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol 1992;45:743-60.

Stone 2010
Stone A. The application of bespoke spending functions in group-sequential designs and the effect of delayed treatment switching in survival trials. Pharm Stat 2010;9(2):151-61.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method with the conventional approach for interval-censored data analysis. Stat Biopharm Res 2010;2(1):97-108.

Tarhini and Kirkwood 2008
Tarhini AA, Kirkwood JM. Tremelimumab (CP-675,206): a fully human anticytotoxic T lymphocyte-associated antigen 4 monoclonal antibody for treatment of patients with advanced cancers. Expert Opin Biol Ther 2008;8(10):1583-93.

Topalian et al 2012
Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med 2012;366(26):2443-54.

Topalian et al 2014
Topalian SL, Sznol M, McDermott DF, Kluger HM, Carvajal RD, Sharfman WH et al. Survival, durable tumor remission, and long-term safety in patients with advanced melanoma receiving nivolumab. J Clin Oncol, 2014 Apr 1; 32(10):1020-30.

Wang et al 2013
Wang Z, Han J, Cui Y, Fan K, Zhou X. Circulating microRNA-21 as noninvasive predictive biomarker for response in cancer immunotherapy. Med Hypotheses 2013;81(1):41-3.

Weber et al 2012
Weber J, Kähler KC, Hauschild A. Management of immune-related adverse events and kinetics of response with ipilimumab. J Clin Oncol 2012;30(21):2691-7.

Whitehead and Whitehead 1991
Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.

Wolchok et al 2013
Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, et al. Nivolumab plus ipilimumab in advanced melanoma. N Engl J Med 2013;369(2):122-33.

Yuan et al 2011
Yuan J, Adamow M, Ginsberg BA, Rasalan TS, Ritter E, Gallardo HF, et al. Integrated NY-ESO-1 antibody and CD8+ T-cell responses correlated with clinical benefit in advanced melanoma patients treated with ipilimumab. Proc Natl Acad Sci USA 2011;108(40):16723-8.

Zhang et al 2008
Zhang C, Wu S, Xue X, Li M, Qin X, Li W, et al. Anti-tumour immunotherapy by blockade of the PD-1/PD-L1 pathway with recombinant human PD-1-IgV. Cytotherapy 2008;10(7):711-9.

Clinical Study Protocol Appendix A

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Protocol Dated	

Appendix A
Signatures

ASTRAZENECA SIGNATURE(S)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

This Clinical Study Protocol has been subjected to an internal AstraZeneca peer review.

I agree to the terms of this study protocol.

AstraZeneca Research and Development site representative

_____

Date
(Day Month Year)

ASTRAZENECA SIGNATURE(S)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

This Clinical Study Protocol has been subjected to an internal AstraZeneca peer review.

I agree to the terms of this study protocol.

AstraZeneca Research and Development site representative

Date
(Day Month Year)

ASTRAZENECA SIGNATURE(S)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

This Clinical Study Protocol has been subjected to an internal AstraZeneca peer review.

I agree to the terms of this study protocol.

AstraZeneca Research and Development site representative

Date
(Day Month Year)

SIGNATURE OF INTERNATIONAL CO-ORDINATING INVESTIGATOR

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First Line Treatment of Patients with Locally Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

This Clinical Study Protocol has been subjected to an internal AstraZeneca peer review.

I agree to the terms of this study protocol. I will conduct the study according to the procedures specified herein, and according to the principles of Good Clinical Practice and local regulations, and I ensure that all relevant site staff follows the instructions given in the latest version of the Laboratory Manual for Investigators.

Center No.:

Signature:

Date
(Day Month Year)

Appendix B
Additional Safety Information

FURTHER GUIDANCE ON THE DEFINITION OF A SERIOUS ADVERSE EVENT (SAE)

Life threatening

'Life-threatening' means that the patient was at immediate risk of death from the AE as it occurred or it is suspected that use or continued use of the product would result in the patient's death. 'Life-threatening' does not mean that had an AE occurred in a more severe form it might have caused death (eg, hepatitis that resolved without hepatic failure).

Hospitalization

Outpatient treatment in an emergency room is not in itself a serious AE, although the reasons for it may be (eg, bronchospasm, laryngeal edema). Hospital admissions and/or surgical operations planned before or during a study are not considered AEs if the illness or disease existed before the patient was enrolled in the study, provided that it did not deteriorate in an unexpected way during the study.

Important medical event or medical intervention

Medical and scientific judgement should be exercised in deciding whether a case is serious in situations where important medical events may not be immediately life threatening or result in death, hospitalization, disability, or incapacity but may jeopardize the patient or may require medical intervention to prevent 1 or more outcomes listed in the definition of serious. These should usually be considered as serious.

Simply stopping the suspect drug does not mean that it is an important medical event; medical judgement must be used.

Examples of such events are:

- Angioedema not severe enough to require intubation but requiring iv hydrocortisone treatment
- Hepatotoxicity caused by paracetamol (acetaminophen) overdose requiring treatment with N-acetylcysteine
- Intensive treatment in an emergency room or at home for allergic bronchospasm
- Blood dyscrasias (eg, neutropenia or anemia requiring blood transfusion, etc) or convulsions that do not result in hospitalization
- Development of drug dependency or drug abuse

A GUIDE TO INTERPRETING THE CAUSALITY QUESTION

When making an assessment of causality consider the following factors when deciding if there is a 'reasonable possibility' that an AE may have been caused by the drug.

- Time Course. Exposure to suspect drug. Has the patient actually received the suspect drug? Did the AE occur in a reasonable temporal relationship to the administration of the suspect drug?

- Consistency with known drug profile. Was the AE consistent with the previous knowledge of the suspect drug (pharmacology and toxicology) or drugs of the same pharmacological class? OR could the AE be anticipated from its pharmacological properties?

- Dechallenge experience. Did the AE resolve or improve on stopping or reducing the dose of the suspect drug?

- No alternative cause. The AE cannot be reasonably explained by another etiology such as the underlying disease, other drugs, or other host or environmental factors.

- Rechallenge experience. Did the AE reoccur if the suspected drug was reintroduced after having been stopped? AstraZeneca would not normally recommend or support a rechallenge.

- Laboratory tests. A specific laboratory investigation (if performed) has confirmed the relationship?

In difficult cases, other factors could be considered such as:

- Is this a recognized feature of overdose of the drug?

- Is there a known mechanism?

Causality of 'related' is made if following a review of the relevant data, there is evidence for a 'reasonable possibility' of a causal relationship for the individual case. The expression 'reasonable possibility' of a causal relationship is meant to convey, in general, that there are facts (evidence) or arguments to suggest a causal relationship.

The causality assessment is performed based on the available data including enough information to make an informed judgment. With limited or insufficient information in the case, it is likely that the event(s) will be assessed as 'not related'.

Causal relationship in cases where the disease under study has deteriorated due to lack of effect should be classified as no reasonable possibility.

MEDI4736 AND TREMELIMUMAB

There is no information to date on drug-drug interactions with MEDI4736 or tremelimumab either pre-clinically or in patients. As MEDI4736 and tremelimumab are monoclonal antibodies and therefore proteins, they will be degraded to small peptides and amino acids and will be eliminated by renal and reticuloendothelial clearance. It is therefore not expected that MEDI4736 or tremelimumab will induce or inhibit the major drug metabolizing cytochrome P450 pathways. As a result, there are no expected pharmacokinetic drug-drug interactions.

No formal drug-drug interaction studies have been conducted with tremelimumab. However, in renal cell carcinoma studies, acute renal failure has been reported with the combination of tremelimumab and sunitinib.

The mechanism of action of MEDI4736 involves binding to PD-L1, and the mechanism of action of tremelimumab involves binding to CTLA-4; therefore, significant pharmacodynamic drug interactions with the commonly administered concomitant medications are not expected. Despite this, appropriate clinical monitoring in all of the planned clinical studies will be conducted to evaluate any potential drug-drug interactions.

AstraZeneca

Clinical Study Protocol Appendix C

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Date	

Appendix C
International Airline Transportation Association (IATA) 6.2 Guidance Document

LABELLING AND SHIPMENT OF BIOHAZARD SAMPLES

International Airline Transportation Association (IATA) classifies biohazardous agents into 3 categories **(http://www.iata.org/whatwedo/cargo/dangerous_goods/infectious_substances.htm)**. For transport purposes, the classification of infectious substances according to risk groups was removed from the Dangerous Goods Regulations (DGR) in the 46th edition (2005). Infectious substances are now classified either as Category A, Category B, or Exempt. There is no direct relationship between Risk Groups and Categories A and B.

Category A Infectious Substances are infectious substances in a form that, when exposure to it occurs, is capable of causing permanent disability, life-threatening, or fatal disease in otherwise healthy humans or animals. Category A pathogens are, eg, Ebola, Lassa fever virus:

- Are to be packed and shipped in accordance with IATA Instruction 602.

Category B Infectious Substances are infectious substances that do not meet the criteria for inclusion in Category A. Category B pathogens are, eg, Hepatitis A, B, C, D, and E viruses, Human immunodeficiency virus (HIV) types 1 and 2. They are assigned the following UN number and proper shipping name:

- UN 3373 – Biological Substance, Category B
- Are to be packed in accordance with UN3373 and IATA 650

Exempt - all other materials with minimal risk of containing pathogens

- Clinical trial samples will fall into Category B or exempt under IATA regulations
- Clinical trial samples will routinely be packed and transported at ambient temperature in IATA 650 compliant packaging **(http://www.iata.org/whatwedo/cargo/dangerous_goods/infectious_substances.htm)**
- **Biological samples transported in dry ice require additional dangerous goods specification for the dry-ice content**
- IATA compliant courier and packaging materials should be used for packing and transportation and packing should be done by an IATA certified person, as applicable

- Samples routinely transported by road or rail are subject to local regulations which require that they are also packed and transported in a safe and appropriate way to contain any risk of infection or contamination by using approved couriers and packaging / containment materials at all times. The IATA 650 biological sample containment standards are encouraged wherever possible when road or rail transport is used.

Clinical Study Protocol Appendix D

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Date	

Appendix D
Pharmacogenetics Research

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or special term	Explanation
DNA	Deoxyribonucleic acid
LIMS	Laboratory information management system
NSCLC	Non-small-cell lung cancer

1. BACKGROUND AND RATIONALE

AstraZeneca intends to perform genetic research in the MEDI4736 clinical development program to explore how genetic variations may affect the clinical parameters associated with this drug combination. Collection of DNA samples from populations with well-described clinical characteristics may lead to improvements in the design and interpretation of clinical trials and, possibly, to genetically guided treatment strategies.

Future research may suggest other genes or gene categories as candidates for influencing not only response to MEDI4736, but also susceptibility to NSCLC. Thus, this genetic research may involve study of additional un-named genes or gene categories, but only as related to NSCLC and MEDI4736 treatment.

2. GENETIC RESEARCH OBJECTIVES

The objective of this research is to collect and store DNA, derived from a blood sample, for future exploratory research into genes/genetic variations that may influence response, ie, distribution, safety, tolerability, and efficacy of MEDI4736, and/or susceptibility to NSCLC.

3. GENETIC RESEARCH PLAN AND PROCEDURES

3.1 Selection of genetic research population

3.1.1 Study selection record

All enrolled patients who take part in the main study will be asked to participate in this genetic research. Participation is voluntary, and if a patient declines to participate, there will be no penalty or loss of benefit. The patient will not be excluded from any aspect of the main study.

3.1.2 Inclusion criteria

For inclusion in this genetic research, patients must fulfil all of the inclusion criteria described in the main body of the Clinical Study Protocol **and**:

- Provide informed consent for the genetic sampling and analyses.

3.1.3 Exclusion criteria

Exclusion from this genetic research may be for any of the exclusion criteria specified in the main study or any of the following:

- Previous allogeneic bone marrow transplant
- Non-leukocyte depleted whole blood transfusion in 120 days of genetic sample collection

3.1.4 Discontinuation of patients from this genetic research

Specific reasons for discontinuing a patient from this genetic research are:

Withdrawal of consent for genetic research: patients may withdraw from this genetic research at any time, independent of any decision concerning participation in other aspects of the main study. Voluntary discontinuation will not prejudice further treatment. Procedures for discontinuation are outlined in Section 3.9 of the main Clinical Study Protocol.

3.2 Collection of samples for genetic research

Blood samples will ideally be collected during the screening/baseline period. If for any reason the sample is not drawn during the screening/baseline period, it should be taken as soon as possible, but not later than the last study visit. Although genotype is a stable parameter, early sample collection is preferred to avoid introducing bias through excluding patients who may withdraw due to an adverse event, as these patients would be important to include in any genetic analysis. Only 1 sample should be collected per patient for genetics during the study.

If the patient agrees to participate, an 9-mL blood sample will be collected into a tube containing reagents that coagulate blood and stabilize blood cell DNA and gently inverted a minimum of 5 times to mix thoroughly. Tubes will be identified with the protocol study number, center number, enrollment code, and date of sample collection. No personal identifiers (patient name, initials, or date of birth) will be placed on the tube or accompanying documentation. A record of the date of the patient consent to the host genetic research and the date of the blood sample collection will be recorded.

AstraZeneca/MedImmune, or its designee, will act as the central laboratory for sample logistics. This will include the supply of site material and all transport arrangements.

A single blood sample will be stored frozen (-20°C or below) at the site and sent to the central laboratory. The central laboratory will then send the samples to AstraZeneca/MedImmune, or its designee laboratory, for DNA extraction. Samples must remain frozen at all times. Further details on the processing of the samples are outlined in the Laboratory Manual for Investigators.

3.3 Coding and storage of DNA samples

The processes adopted for the coding and storage of samples for genetic analysis are important to maintain patient confidentiality. Samples will be stored for a maximum of 15 years from the date of the last patient's last visit, after which they will be destroyed. DNA is a finite resource that may be used up during analyses. Samples will be stored and used until no further analyses are possible or the maximum storage time has been reached.

For all samples irrespective of the type of coding used the DNA will be extracted from the blood sample. The DNA sample will be assigned a unique number replacing the information on the sample tube. Thereafter, the DNA sample will be identifiable by the unique DNA number only. The DNA number is used to identify the sample and corresponding data at the AstraZeneca/MedImmune genetics laboratories, or at the designated contract laboratory. No personal details identifying the individual will be available to any person (AstraZeneca/MedImmune employee or contract laboratory staff working with the DNA).

The samples and data for genetic analysis in this study will be single coded. The link between the patient enrollment/randomization code and the DNA number will be maintained and stored in a secure environment, with restricted access WITHIN the Clinical Genotyping Group Laboratory Information Management System (LIMS) at AstraZeneca. The link will be used to identify the relevant DNA samples for analysis, facilitate correlation of genotypic results with clinical data, allow regulatory audit and to trace samples for destruction in the case of withdrawal of consent when the patient has requested disposal/destruction of collected samples not yet analyzed.

4. ETHICAL AND REGULATORY REQUIREMENTS

The principles for ethical and regulatory requirements for the study, including this genetics research component, are outlined in Section 10 of the main Clinical Study Protocol.

The study will be performed in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation/Good Clinical Practice, applicable regulatory requirements, and the AstraZeneca policy on Bioethics.

4.1 Informed consent

The portion of this study evaluating genetic alterations in blood samples is optional, and the patient may participate in other components of the main study without participating in this specific genetic analysis. To participate in this genetic component of the study, the patient must sign and date both the consent form for the main study and the genetic component of the study. Copies of both signed and dated consent forms must be given to the patient and the original filed at the study center. The principal investigator(s) is responsible for ensuring that consent is given freely and that the patient understands that they may freely discontinue from the genetic aspect of the study at any time.

4.2 Subject data protection

AstraZeneca/MedImmune will not provide individual genotype results to patients, any insurance company, any employer, their family members, their general physician, or any other third party, unless required to do so by law.

Extra precautions are taken to preserve confidentiality and prevent genetic data being linked to the identity of the patient. Individual patients will not be identified in any report or publication resulting from this work. The data and results of this study may be reviewed with collaborators and published, but neither the patient's name nor any other personal identifiers will appear in any publication or report. In exceptional circumstances, however, certain individuals might see both the genetic data and the personal identifiers of a patient. For example, in the case of a medical emergency, an AstraZeneca/MedImmune physician or an investigator might know a patient's identity and also have access to his or her genetic data. Regulatory authorities may also require access to the relevant files, though the patient's medical information and the genetic files would remain physically separate.

5. DATA MANAGEMENT

Any genotype data generated in this study will be stored in the AstraZeneca genotyping LIMS database, or other appropriate secure system within AstraZeneca and/or third party contracted to work with AstraZeneca/MedImmune to analyze the samples.

The results from this genetic research will be reported separately from the clinical study report for the main study.

Some or all of the clinical datasets from the main study may be merged with the genetic data in a suitable secure environment separate from the clinical database.

6. STATISTICAL METHODS AND DETERMINATION OF SAMPLE SIZE

The number of patients that will agree to participate in the genetic research is unknown. It is therefore not possible to establish whether sufficient data will be collected to allow a formal statistical evaluation or whether only descriptive statistics will be generated. A statistical analysis plan will be prepared where appropriate.

7. LIST OF REFERENCES

None

AstraZeneca

Clinical Study Protocol Appendix E

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Date	

Appendix E
Actions Required in Cases of Combined Increase of Aminotransferase and Total Bilirubin - Hy's Law

TABLE OF CONTENTS PAGE

1. INTRODUCTION

This Appendix describes the process to be followed in order to identify and appropriately report cases of Hy's Law. It is not intended to be a comprehensive guide to the management of elevated liver biochemistries. Specific guidance on the managing liver abnormalities can be found in Section 6.7 of the protocol.

During the course of the study the Investigator will remain vigilant for increases in liver biochemistry. The investigator is responsible for determining whether a patient meets potential Hy's Law (PHL) criteria at any point during the study.

The Investigator participates, together with AstraZeneca clinical project representatives, in review and assessment of cases meeting PHL criteria to agree whether Hy's Law (HL) criteria are met. HL criteria are met if there is no alternative explanation for the elevations in liver biochemistry other than Drug Induced Liver Injury (DILI) caused by the Investigational Medicinal Product (IMP).

The Investigator is responsible for recording data pertaining to PHL/HL cases and for reporting Adverse Events (AE) and Serious Adverse Events (SAE) according to the outcome of the review and assessment in line with standard safety reporting processes.

2. DEFINITIONS

Potential Hy's Law (PHL)

Aspartate Aminotransferase (AST) or Alanine Aminotransferase (ALT) ≥3 × Upper Limit of Normal (ULN) **together with** Total Bilirubin (TBL) ≥2 × ULN at any point during the study following the start of study medication irrespective of an increase in Alkaline Phosphatase (ALP).

Hy's Law (HL)

AST or ALT ≥3 × ULN **together with** TBL ≥2 × ULN, where no other reason, other than the IMP, can be found to explain the combination of increases, eg, elevated ALP indicating cholestasis, viral hepatitis, another drug.

For PHL and HL to be met the elevation in transaminases must precede or be coincident with (ie, on the same day) the elevation in TBL, but there is no specified timeframe within which the elevations in transaminases and TBL must occur.

3. IDENTIFICATION OF POTENTIAL HY'S LAW CASES

In order to identify cases of PHL it is important to perform a comprehensive review of laboratory data for any patient who meets any of the following identification criteria in isolation or in combination:

- ALT ≥3 × ULN
- AST ≥3 × ULN
- TBL ≥2 × ULN

When a patient meets any of the identification criteria, in isolation or in combination, the central laboratory will immediately send an alert to the Investigator (also sent to AstraZeneca representative).

The Investigator will also remain vigilant for any local laboratory reports where the identification criteria are met, where this is the case the Investigator will:

- Notify the AstraZeneca representative
- Request a repeat of the test (new blood draw) by the central laboratory
- Complete the appropriate unscheduled laboratory CRF module(s) with the original local laboratory test result

When the identification criteria are met from central or local laboratory results the Investigator will without delay:

- Determine whether the patient meets PHL criteria (see Section 2 of this Appendix for definition) by reviewing laboratory reports from all previous visits (including both central and local laboratory results)

The Investigator will, without delay, review each new laboratory report and if the identification criteria are met will:

- Determine whether the patient meets PHL criteria (see Section 2 of this Appendix for definition) by reviewing laboratory reports from all previous visits
- Promptly enter the laboratory data into the laboratory CRF

4. FOLLOW-UP

4.1 Potential Hy's Law Criteria not met

If the patient does not meet PHL criteria the Investigator will:

- Perform follow-up on subsequent laboratory results according to the guidance provided in the Clinical Study Protocol.

4.2 Potential Hy's Law Criteria met

If the patient does meet PHL criteria the Investigator will:

- Determine whether PHL criteria were met at any study visit prior to starting study treatment in the presence of liver metastases (see Section 6)
- Notify the AstraZeneca representative who will then inform the central Study Team

The Study Physician contacts the Investigator, to provide guidance, discuss, and agree an approach for the study patients' follow-up and the continuous review of data. Subsequent to this contact the Investigator will:

- Monitor the patient until liver biochemistry parameters and appropriate clinical symptoms and signs return to normal or baseline levels, or as long as medically indicated
- Investigate the etiology of the event and perform diagnostic investigations as discussed with the Study Physician
- Complete the 3 Liver CRF Modules as information becomes available
- If at any time (in consultation with the Study Physician) the PHL case meets serious criteria, report it as an SAE using standard reporting procedures

5. REVIEW AND ASSESSMENT OF POTENTIAL HY'S LAW CASES

The instructions in this section should be followed for all cases where PHL criteria are met.

No later than 3 weeks after the biochemistry abnormality was initially detected, the Study Physician contacts the Investigator in order to review available data and agree on whether there is an alternative explanation for meeting PHL criteria other than DILI caused by the IMP. The AstraZeneca Medical Science Director and Global Safety Physician will also be involved in this review together with other subject matter experts as appropriate.

According to the outcome of the review and assessment, the Investigator will follow the instructions below.

If there **is** an agreed alternative explanation for the ALT or AST and TBL elevations, a determination of whether the alternative explanation is an AE will be made and subsequently whether the AE meets the criteria for a SAE:

- If the alternative explanation is **not** an AE, record the alternative explanation on the appropriate CRF

- If the alternative explanation is an AE/SAE, record the AE /SAE in the CRF accordingly and follow the AZ standard processes

If it is agreed that there is **no** explanation that would explain the ALT or AST and TBL elevations other than the IMP:

- Report an SAE (report term 'Hy's Law') according to AstraZeneca standard processes.

  – The 'Medically Important' serious criterion should be used if no other serious criteria apply

  – As there is no alternative explanation for the HL case, a causality assessment of 'related' should be assigned.

If, there is an unavoidable delay, of over 3 weeks, in obtaining the information necessary to assess whether or not the case meets the criteria for HL, then it is assumed that there is no alternative explanation until such time as an informed decision can be made:

- Report an SAE (report term 'Potential Hy's Law') applying serious criteria and causality assessment as per above

- Continue follow-up and review according to agreed plan. Once the necessary supplementary information is obtained, repeat the review and assessment to determine whether HL criteria are met. Update the SAE report according to the outcome of the review

6. ACTIONS REQUIRED WHEN POTENTIAL HY'S LAW CRITERIA ARE MET BEFORE AND AFTER STARTING STUDY TREATMENT

This section is applicable to patients who meet PHL criteria on study treatment (including the 30-day follow-up period post discontinuation of study treatment) having previously met PHL criteria at a study visit prior to starting study treatment.

At the first on study treatment occurrence of PHL criteria being met the Investigator will:

- Determine if there has been a significant change in the patients' condition[#] compared with the last visit where PHL criteria were met[#]
  - If there is no significant change no action is required
  - If there is a significant change notify the AstraZeneca representative, who will inform the central Study Team, then follow the subsequent process described in Section 4.2 of this Appendix

[#] A 'significant' change in the patient's condition refers to a subsequent clinically relevant change in any of the individual liver biochemistry parameters (ALT, AST, or total bilirubin) in isolation or in combination, or a clinically relevant change in associated symptoms such as fatigue, vomiting, rash, right upper quadrant pain, jaundice, or eosinophilia. The determination of whether there has been a significant change will be at the discretion of the Investigator; this may be in consultation with the Study Physician if there is any uncertainty.

7. ACTIONS REQUIRED FOR REPEAT EPISODES OF POTENTIAL HY'S LAW

This section is applicable when a patient meets PHL criteria on study treatment (including the 30-day follow-up period) and has already met PHL criteria at a previous on study treatment visit.

The requirement to conduct follow-up, review, and assessment of a repeat occurrence(s) of PHL is based on the nature of the alternative cause identified for the previous occurrence.

The Investigator should determine the cause for the previous occurrence of PHL criteria being met and answer the following question:

- Was the alternative cause for the previous occurrence of PHL criteria being met chronic or progressing malignant disease or did the patient meet PHL criteria prior to starting study treatment and at their first on study treatment visit as described in Section 6?

If No: follow the process described in Section 4.2 of this Appendix

If Yes:

Determine if there has been a significant change in the patient's condition[#] compared with when PHL criteria were previously met

- If there is no significant change no action is required

- If there is a significant change follow the process described in Section 4.2 of this Appendix

[#] A 'significant' change in the patient's condition refers to a subsequent clinically relevant change in any of the individual liver biochemistry parameters (ALT, AST, or total bilirubin) in isolation or in combination, or a clinically relevant change in associated symptoms, such as fatigue, vomiting, rash, right upper quadrant pain, jaundice, or eosinophilia. The determination of whether there has been a significant change will be at the discretion of the Investigator; this may be in consultation with the Study Physician if there is any uncertainty.

8. REFERENCES

FDA Guidance for Industry (issued July 2009) 'Drug-induced liver injury: Premarketing clinical evaluation':

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf

AstraZeneca

Clinical Study Protocol Appendix F

Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Appendix Date	

Appendix F
Guidelines for Evaluation of Objective Tumor Response Using RECIST 1.1 Criteria (Response Evaluation Criteria in Solid Tumors)

TABLE OF CONTENTS PAGE

LIST OF TABLES

1. INTRODUCTION

This appendix details the implementation of Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 guidelines (Eisenhauer et al 2009) for the D419AC00001 study with regards to Investigator assessment of tumor burden including protocol-specific requirements for this study.

2. DEFINITION OF MEASURABLE, NON-MEASURABLE, TARGET AND NON-TARGET LESIONS

Only patients with measurable disease at baseline should be included in the study. Measurable disease is defined by the presence of at least one measurable lesion which has not been previously irradiated.

Tumor lesions selected for screening biopsy must not be used as index lesions, unless there are no other lesions suitable for biopsy.

Measurable:

A lesion, not previously irradiated or biopsied per the protocol prior to randomisation, that can be accurately measured at baseline as ≥10 mm in the longest diameter (except lymph nodes which must have short axis ≥15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements.

Non-measurable:

- All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with ≥10 mm to <15 mm short axis at baseline[1]).
- Truly non-measurable lesions include the following: bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical examination that is not measurable by CT or MRI.
- Tumour lesions selected for screening biopsy
- Previously irradiated lesions[2]

[1] Nodes with <10 mm short axis are considered non-pathological and should not be recorded or followed as non-target lesions (NTLs).
[2] Localised post-radiation changes which affect lesion sizes may occur. Therefore, lesions that have been previously irradiated will not be considered measurable and must be selected as NTL at baseline and followed up as part of the NTL assessment.

- Skin lesions assessed by clinical examination
- Brain metastasis

Special cases:

- Lytic bone lesions or mixed lytic–blastic lesions, with identifiable soft tissue components, can be considered measurable if the soft tissue component meets the definition of measurability. Blastic lesions are considered non-measurable.
- Cystic metastases can be considered measurable lesions if they meet the criteria for measurability from a radiological point of view, but if non-cystic lesions are present in the same patient, these should be selected as target lesions (TLs).

Target lesions:

A maximum of 5 measurable lesions (with a maximum of 2 lesions per organ), representative of all lesions involved suitable for accurate repeated measurement, should be identified as TLs at baseline.

Non-target lesions:

All other lesions (or sites of disease) not recorded as TL should be identified as NTL at baseline.

3. METHODS OF ASSESSMENT

The same method of assessment and the same technique should be used to characterize each identified and recorded lesion at baseline and during follow-up visits.

A summary of the methods to be used for RECIST assessment is provided in Table 1, and those excluded from tumor assessments for this study are highlighted with the rationale provided.

Table 1 Summary of methods of assessment

Target lesions	Non-target lesions	New lesions
CT (preferred)	CT (preferred)	CT (preferred)
MRI	MRI	MRI
	Clinical examination	Clinical examination
	X-ray, Chest X-ray	X-ray, Chest X-ray
		Ultrasound
		Bone scan
		FDG-PET

CT Computed tomography; FDG-PET 18-Fluoro-deoxyglucose positron emission tomography; MRI Magnetic resonance imaging.

3.1 CT and MRI

CT and MRI are generally considered to be the best currently available and reproducible methods to measure TL selected for response assessment and to assess NTL and identification of any new lesions.

In the D419AC00001 study, it is recommended that CT examinations of the chest and abdomen (including liver and adrenal glands) will be used to assess tumour burden at baseline and follow-up visits. CT examination with intravenous contrast media administration is the preferred method. MRI should be used where CT is not feasible or it is medically contra-indicated. For brain lesion assessment, MRI is the preferred method.

3.2 Clinical examination

In the D419AC00001 study, clinical examination will not be used for assessment of TL. Clinically detected lesions can be selected as TLs if they are assessed by CT or MRI scans. Clinical examination can be used to assess NTL and to identify the presence of new lesions.

3.3 X-ray

3.3.1 Chest X-ray

In the D419AC00001 study, chest X-ray assessment will not be used for assessment of TL as they will be assessed by CT examination or MRI examination. Chest X-ray can, however, be used to assess NTL and to identify the presence of new lesions.

3.3.2 Plain X-ray

In the D419AC00001 study plain X-ray may be used as a method of assessment for bone NTL and to identify the presence of new bone lesions.

3.4 Ultrasound

In the D419AC00001 study, ultrasound examination will not be used for assessment of TL and NTL as it is not a reproducible method, does not provide an accurate assessment of tumor size, and it is subjective and operator dependent. Ultrasound examination can, however, be used to identify the presence of new lesions. If new clinical symptoms occur and an ultrasound is performed then new lesions should be confirmed by CT or MRI examination.

3.5 Endoscopy and laparoscopy

In the D419AC00001 study, endoscopy and laparoscopy will not be used for tumor assessments as they are not validated in the context of tumor assessment.

3.6 Tumor markers

In the D419AC00001 study, tumor markers will not be used for tumor response assessments as per RECIST 1.1.

3.7 Cytology and histology

In the D419AC00001 study histology will not be used as part of the tumor response assessment as per RECIST 1.1.

Cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment is required when the measurable tumour has met criteria for response or stable disease. In such circumstances, the cytology is necessary to differentiate between response/stable disease (an effusion may be a side effect of the treatment) and progressive disease (if the neoplastic origin of the fluid is confirmed). Where cytology findings are not available, any effusion that significantly worsens (from trace to large) or appearance of clinically significant effusion (requiring change in drug therapy) during the study treatment will be considered to be progression of NTL, or disease progression due to new lesions.

3.8 Isotopic bone scan

Bone lesions identified on an isotopic bone scan at baseline and confirmed by CT, MRI, or X-ray at baseline should be recorded as NTL and followed by the same method as per baseline assessment.

In the D419AC00001 study, isotopic bone scans may be used as a method of assessment to identify the presence of new bone lesions at follow-up visits. New lesions will be recorded where a positive hot-spot that was not present on the baseline bone scan assessment is identified on a bone scan performed at any time during the study. The Investigator should consider the positive hot-spot to be a significant new site of malignant disease and represent true disease progression in order to record the new lesion. Confirmation by CT, MRI and X-ray is recommended where bone scan findings are equivocal.

3.9 FDG-PET scan

In the D419AC00001 study, 18-Fluoro-deoxyglucose positron emission tomography (FDG-PET) scans may be used as a method for identifying new lesions, according with the following algorithm: New lesions will be recorded where there is positive 18-Fluoro-deoxyglucose uptake[3] not present on baseline FDG-PET scan or in a location corresponding to a new lesion on CT/MRI at the same follow-up visit. If there is no baseline FDG-PET scan available, and no evidence of new lesions on CT/MRI scans then follow-up CT/MRI assessments should be continued, scheduled as per protocol or clinical indicated, in order to confirm new lesions.

[3] A positive FDG-PET scan lesion should be reported only when an uptake greater than twice that of the surrounding tissue is observed.

4. TUMOR RESPONSE EVALUATION

4.1 Schedule of evaluation

The methods of assessment of tumor burden used at baseline CT/MRI scans of the chest and abdomen (including liver and adrenal glands) must be used at each subsequent follow-up assessment. Additional imaging may be performed based on the signs and symptoms of the patient, eg, new lesions at follow up.

Baseline assessments should be performed no more than 28 days before start of study treatment, and ideally should be performed as close as possible to the start of investigational product. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; see Section 3.1 of the Clinical Study Protocol), then every 8 weeks thereafter until confirmed objective disease progression as defined by RECIST 1.1 (irrespective of the reason for stopping treatment/or subsequent therapy). 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.

For patients who discontinue study drug due to toxicity in the absence of confirmed objective progression, objective tumor assessments should be continued every 6 weeks for 48 weeks (relative to the date of randomization) then every 8 weeks until confirmed objective disease progression.

Disease progression requires confirmation; the confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD in the absence of clinically significant deterioration.

If progression is not confirmed then the patient should continue on study treatment and on treatment assessments.

If an unscheduled assessment was 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.

Additional assessments will be performed post confirmed objective disease progression for patients remaining on IMT treatment, re-treatment, or until subsequent cancer therapy according to the clinical study protocol.

4.2 Target lesions

4.2.1 Documentation of target lesions

A maximum of 5 measurable lesions, with a maximum of 2 lesions per organ (including lymph nodes), representative of all lesions involved should be identified as TL at baseline. Target lesions should be selected on the basis of their size (longest diameter for non-nodal

lesions or short axis for nodal lesions), but in addition should be those that lend themselves to reproducible repeated measurements. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

The site and location of each TL should be documented as well as the longest diameter for non-nodal lesions (or short axis for lymph nodes). All measurements should be recorded in millimeters. At baseline the sum of the diameters for all TL will be calculated and reported as the baseline sum of diameters. At follow-up visits the sum of diameters for all TL will be calculated and reported as the follow-up sum of diameters.

Special cases:

- For TL measurable in 2 or 3 dimensions, always report the longest diameter. For pathological lymph nodes measurable in 2 or 3 dimensions, always report the short axis.
- If the CT/MRI slice thickness used is >5 mm, the minimum size of measurable disease at baseline should be twice the slice thickness of the baseline scan.
- If a lesion has completely disappeared, the longest diameter should be recorded as 0 mm.
- If a TL splits into two or more parts, then record the sum of the diameters of those parts.
- If two or more TLs merge then the sum of the diameters of the combined lesion should be recorded for one of the lesions and 0 mm recorded for the other lesion(s).
- If a TL is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned. If an accurate measure can be given, this should be recorded, even if it is below 5 mm.
- If a TL cannot be measured accurately due to it being too large, provide an estimate of the size of the lesion.
- When a TL has had any intervention eg, radiotherapy, embolization, surgery, during the study, the size of the TL should still be provided where possible.

4.2.2 Evaluation of target lesions

This section provides the definitions of the criteria used to determine objective tumor visit response for TL (see Table 2).

Table 2 Evaluation of target lesions

Complete Response (CR)	Disappearance of all TLs since baseline. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10 mm.
Partial Response (PR)	At least a 30% decrease in the sum of the diameters of TL, taking as reference the baseline sum of diameters
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Progression of disease (PD)	At least a 20% increase in the sum of diameters of TLs, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm.
Not Evaluable (NE)	Only relevant if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit. Note: if the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response

CR Complete response; PR Partial response; PD Progression of disease; NE Not evaluable; SD Stable disease; TL Target lesion.

4.3 Non-target lesions

4.3.1 Evaluation of non-target lesions

All other lesions (or sites of disease) not recorded as TL should be identified as NTL at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits. At each visit an overall assessment of the NTL response should be recorded by the Investigator. This section provides the definitions of the criteria used to determine and record overall response for NTL at the investigational site at each visit (see Table 3).

Table 3 Evaluation of non-target lesions

Complete response (CR)	Disappearance of all NTLs since baseline. All lymph nodes must be non-pathological in size (<10 mm short axis).
Non CR/Non PD	Persistence of one or more NTL.
Progression (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.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the Investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: for patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met.

CR Complete response; PR Partial response; PD Progression of disease; NE Not evaluable; NTL Non-target lesion; TL Target lesion.

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

4.4 New lesions

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: ie, not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumor.

If a new lesion is equivocal, for example because of its small size, the treatment and tumor assessments should be continued until the new lesion has been confirmed. If repeat scans confirm there is a new lesion, then the progression date should be declared using the date of the initial scan.

4.5 Symptomatic deterioration

Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping study therapy.

Patients with 'symptomatic deterioration' 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.

4.6 Evaluation of overall visit response

The overall visit response will be derived using the algorithm shown in Table 4.

Table 4 Overall visit response

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	NA	No	CR
CR	Non CR/Non PD	No	PR
CR	NE	No	PR
PR	Non PD or NE	No	PR
SD	Non PD or NE	No	SD
NE	Non PD or NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR Complete response, PR Partial response, SD Stable disease, PD Progression of disease, NE Not evaluable, NA Not applicable (only relevant if there were no non-target lesions at baseline).

5. CONFIRMATION OF PROGRESSION

Disease progression requires confirmation; the confirmatory scan should occur no earlier than 4 weeks after the initial assessment of progression of disease (PD) in the absence of clinical deterioration.

Progression would be considered confirmed if the following criteria are met:

- ≥20% increase in the sum diameters of TLs compared with the nadir at 2 consecutive visits
- And/or significant progression (worsening) of NTLs or new lesions at the confirmatory PD time-point compared with the first time point where progression of NTLs or new lesions identified
- And/or additional new unequivocal lesions at the confirmatory PD time-point compared with the first time point new lesions identified.

In the absence of significant clinical deterioration, the Investigator should continue study treatment until progression is confirmed.

If progression is not confirmed, then the patient should continue on study treatment and on treatment assessments.

If a patient discontinues treatment (and/or receives a subsequent cancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

6. CENTRAL REVIEW

The Contract Research Organization appointed by AstraZeneca to perform the independent central review for this study will provide specification for radiological imaging protocols in standard acquisition guidelines documentation.

7. REFERENCES

Eisenhauer et al 2009
Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228-47.

AstraZeneca

Clinical Study Protocol Appendix G	
Drug Substance	MEDI4736 and tremelimumab
Study Code	D419AC00001
Edition Number	01
Date	

Appendix G
Patient Reported Outcomes: EORTC QLQ-C30, EORTC QLQ-LC13, PRO-CTCAE, EQ-5D-5L, and PGIC

EORTC QLQ-C30 (version 3)

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

Please fill in your initials: ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

Your birthdate (Day, Month, Year): ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

Today's date (Day, Month, Year): 31 ⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋⌊ ⌋

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

During the past week:

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

Please go on to the next page

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

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

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

1	2	3	4	5	6	7
Very poor						Excellent

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

1	2	3	4	5	6	7
Very poor						Excellent

EORTC QLQ - LC13

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

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

NCI- PRO-CTCAE ITEMS

As individuals go through treatment for their cancer they sometimes experience different symptoms and side effects. For each question, please check or mark an ⓧ in the one box that best describes your experiences over the past 7 days…

RASH	
Did you have any RASH?	
○ Yes	○ No

HAIR LOSS				
Did you have any HAIR LOSS?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

HAND-FOOT SYNDROME (A RASH OF THE HANDS OR FEET THAT CAN CAUSE CRACKING, PEELING, REDNESS OR PAIN)				
What was the SEVERITY of your HAND-FOOT SYNDROME (A RASH OF THE HANDS OR FEET THAT CAN CAUSE CRACKING, PEELING, REDNESS OR PAIN) at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

ITCHY SKIN				
What was the SEVERITY of your ITCHY SKIN at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

ARM OR LEG SWELLING				
How often did you have ARM OR LEG SWELLING?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly
What was the SEVERITY of your ARM OR LEG SWELLING at their WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did ARM OR LEG SWELLING INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

NAUSEA				
How often did you have NAUSEA?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly
What was the SEVERITY of your NAUSEA at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

VOMITING				
How often did you have VOMITING?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly
What was the SEVERITY of your VOMITING at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

LOOSE OR WATERY STOOLS (DIARRHEA)				
How often did you have LOOSE OR WATERY STOOLS (DIARRHEA)?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly

NCI- PRO-CTCAE ITEMS

Please think back over the past 7 days…

NUMBNESS OR TINGLING IN YOUR HANDS OR FEET				
What was the SEVERITY of your NUMBNESS OR TINGLING IN YOUR HANDS OR FEET at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did NUMBNESS OR TINGLING IN YOUR HANDS OR FEET INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

MOUTH AND THROAT SORES				
What was the SEVERITY of your MOUTH AND THROAT SORES at their WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did MOUTH AND THROAT SORES INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

SHIVERING OR SHAKING CHILLS				
How often did you have SHIVERING OR SHAKING CHILLS?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly
What was the SEVERITY of your SHIVERING OR SHAKING CHILLS at their WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe

PAIN, SWELLING, REDNESS AT A SITE OF DRUG INJECTION OR IV	
Did you have any PAIN, SWELLING, REDNESS AT A SITE OF DRUG INJECTION OR IV?	
○ Yes	○ No

DIZZINESS				
What was the SEVERITY of your DIZZINESS at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did DIZZINESS INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

Health Questionnaire

English version for the UK

Under each heading, please tick the ONE box that best describes your health TODAY

MOBILITY

I have no problems in walking about ❑
I have slight problems in walking about ❑
I have moderate problems in walking about ❑
I have severe problems in walking about ❑
I am unable to walk about ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN / DISCOMFORT

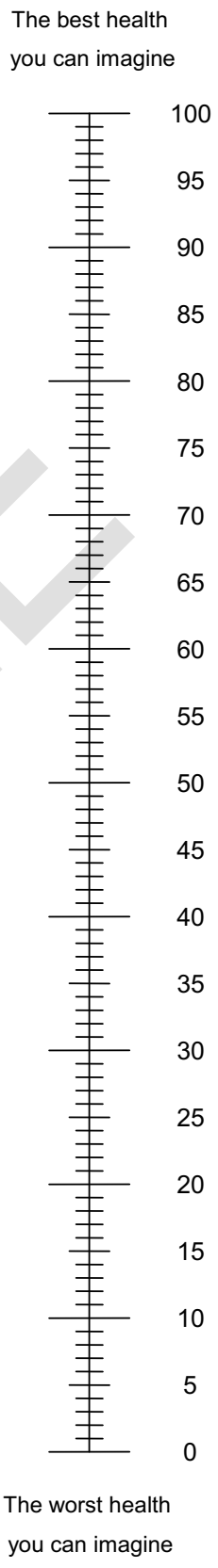
I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

ANXIETY / DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



PATIENT GLOBAL IMPRESSION OF CHANGE (PGIC)

Since the start of the treatment I have received in this study, my overall health status is:

Please tick (✓) one box only:

- ☐ Very Much Improved
- ☐ Much Improved
- ☐ Minimally Improved
- ☐ No Change
- ☐ Minimally Worse
- ☐ Much Worse
- ☐ Very Much Worse

AstraZeneca

Clinical Study Protocol

Drug Substance	Durvalumab (MEDI4736) and tremelimumab
Study Code	D419AC00001
Version	08
Date	

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

Sponsor: AstraZeneca AB, 151 85 Södertälje, Sweden

EudraCT Number: 2015-001279-39

VERSION HISTORY

Version 8.0, Changes to the protocol are summarised below.
Synopsis, Sections 6.1.9, 7.2.2 Updated to include further clarification on post final Data Cut Off (DCO) procedures. **Section 1.3.2, Potential Risks** Since the MYSTIC study was started, the durvalumab programme has developed and the emerging safety profile has become more established. As such, the "Potential Risks" section of this clinical study protocol has updated to align and be consistent with the broader durvalumab programme, including the Investigator Brochure and the clinical study protocol format of more recent studies. **Section 3.10.3, Survival Status for Withdrawn Consent and Lost to Follow Up Patients** Updated to clarify which analysis sets require survival status data and how and when to obtain this data. **Sections 6.3.1, Time period for collection of adverse events, 6.3.2, Follow-up of unresolved adverse events** Updated to clarify time period for collection of adverse events, and to clarify the follow-up of adverse events unresolved at the patient's last visit in the study or at study completion. **6.7, Management of Investigational Product-related Toxicities** Updated to include the web link to the most current Toxicity Management Guidelines. **6.7.1, MEDI4736 + tremelimumab and MEDI4736 monotherapy adverse events of special interest (AESI)** Updated to align with the most current Investigator Brochure **Appendix G** Updated to include the most current version of the Dosing Modification and Toxicity Management Guidelines. **Various** Updated terminology from immune related Adverse Event (irAE) to immune mediated Adverse Event (imAE).

Version 7.0,

Changes to the protocol are summarised below.

Synopsis, Section 1.4
Clarification of the expected proportion of patients with ≥25% PD-L1 membrane-expression in tumoral tissue. With the study fully recruited there is no change in the overall patient number.

Synopsis, Glossary, Figure 1, Section 1.4
The definitions for the PD-L1 tumor membrane-expression sub-groups are provided.

Synopsis, Sections 1.1.5, 1.2.6, 1.3.1.1, 1.3.1.3, 2.1, 2.2, 2.3
For MEDI4733 in combination with tremelimumab the assessment of progression-free survival (PFS) and overall survival (OS) were nominated as co-primary objectives in NSCLC patients with ≥25% PD-L1 membrane-expression in tumoral tissue.

For MEDI4733 monotherapy the assessment of overall survival (OS) is nominated as a co-primary objective in NSCLC patients with ≥25% PD-L1 membrane-expression in tumoral tissue .

Primary and secondary objectives, endpoints and rational for endpoints were modified accordingly.

Synopsis, Section 8, Table 9, Table 12, Figure 5
The statistical considerations and analyses have been updated to accommodate the nominated co-primary objectives.

Section 3.1, Inclusion Criteria #3
Correction of spelling error.

Section 4
Removed the set duration for treatment periods as they can be variable per protocol.

Table 2
Footnote O updated for clarity.

Table 4
Footnote J updated for clarity.

Sections 6.7.1, 6.7.2
Updated to align with the current IB and ICF.

Section 7.7
In the section on prohibited medications, the following changes have been made:

- A temporary period of steroids will be allowed if clinically indicated and considered to be essential for the management of the patient (eg, for chronic obstructive pulmonary disease, radiation, nausea, etc).
- EGFR TKIs should not be given concomitantly whilst the patient is on study treatment. In addition they should be used with caution in the 90 days after the last dose of durvalumab. (Increased incidences of pneumonitis (with third generation EGFR TKIs) and increased

Version 7.0,

Changes to the protocol are summarised below.

incidence of transaminase increases (with 1st generation EGFR TKIs) has been reported when durvalumab has been given concomitantly).

- Inactivated viruses, such as those in the influenza vaccine, are permitted

Section 8.4.4 – 8.4.4.3
Statement that the clinical meaningfulness threshold of the PRO analyses will be described in the Statistical Analysis Plan (SAP)

Appendix E – RECIST
Text edits have been inserted for clarity

Appendix G
Appendix G updated with the new version of the Dosing Modification and Toxicity Management Guidelines for Immune-mediated, Infusion-related, and Nonimmune-mediated Reactions (MEDI4736 Monotherapy, Tremelimumab Monotherapy, or MEDI4736 + Tremelimumab Combination Therapy.

Version 6.0,

Changes to the protocol are summarised below.

Section 3.8 Restrictions
Spermicide was removed as it is not a highly effective method of contraception.

Synopsis, Section 7.2.2 Duration of treatment and criteria for retreatment

To clarify that treatment through progression only applies to the immunotherapy groups (monotherapy and combination therapy arms).

Updated wording in section 7.2.2. to match synopsis, indicating patients in the immunotherapy groups (rather than only the MEDI4736 group) will not be permitted to continue immunotherapy if progression occurs in a target lesion that has previously shown a confirmed response.

Section 8.4.1.1 Progression-free survival To clarify that in the SoC group treatment can be continued, at investigator's discretion, until disease progression in confirmed. Patients in the SoC are not allowed to continue treatment once disease progression in confirmed

Version 5.0,

Changes to the protocol are summarized below

Title pages

MEDI4736 is identified as Durvalumab in Drug substance in the header box, and referred to as MEDI4736 thereafter.

Synopsis, Section 1.2.6 – Rationale for Endpoints, Section 2 – Study Objectives, Section 5.1 – Efficacy Assessments, Section 8.4 – Outcome Measures for Analysis, Section 8.5 – Methods for Statistical Analysis

The primary and secondary objective endpoints have been updated to reflect changes to endpoint measures - Blinded Independent Central Review (BICR) tumor assessments rather than Investigator assessments as well as to assess the treatment benefit and efficacy of MEDI4736 as suggested by emerging immuno-oncology data. Investigator assessments will be used for sensitivity analysis.

Synopsis, Section 1.4 – Study design and Section 8.2 – Sample size estimate

Numbers of subjects enrolled, randomized, and per treatment group were updated to 1850, 1092, and 364, respectively. Figure 1 (overall study design) and Figure 2 (study flow chart) were updated accordingly.

Synopsis, Section 1.4 – Study design

The timing for patients to provide a tumor tissue sample was clarified as being the enrollment visit.

Synopsis, Section 1.4 – Study design and Section 2 – Study objectives

The assessment of progression-free survival (PFS) and overall survival (OS) were nominated as co-primary objectives. Primary and secondary objectives, endpoints and rational for endpoints were modified accordingly.

Synopsis, Section 7.2.2 Duration of treatment and criteria for retreatment

Duration of treatment was modified so that patients in all groups can continue therapy until disease progression rather than stopping at 12 months. Emerging data from ongoing MEDI4736 studies is suggestive of some patients losing clinical benefit after they complete the 12 months of therapy.

In all groups, patients with PD (unconfirmed and confirmed) who, in the Investigator's opinion, would continue to receive benefit from their assigned treatment, and who meet the criteria for treatment in the setting of PD, may continue to receive treatment. It was also clarified that patients on MEDI4736 alone will not be permitted to continue immunotherapy, if the progression occurs after confirmed response to immunotherapy treatment within the target lesion, and if progression events occurred in the target lesions while the patient was receiving immunotherapy during the same treatment period.

Synopsis, Section 8 – Statistical methods

The statistical considerations and analyses have been updated to accommodate the new co-primary objectives and the modifications in provisions for duration of treatment and progression during treatment.

List of Appendices, Section 6.7 – Management of investigational product-related toxicities, and Appendices

Appendix G was updated with the new version of the Dosing Modification and Toxicity Management Guidelines for Immune-mediated, Infusion-related, and Nonimmune-mediated Reactions (MEDI4736 Monotherapy, Tremelimumab Monotherapy, or MEDI4736 + Tremelimumab Combination Therapy. A new version of European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire was included in Appendix F.

Section 1.2.5 - Rationale of retreatment option

The rationale for the retreatment option is amended to enable patients in the MEDI4736 + tremelimumab combination group who complete 4 dosing cycles (providing clinical benefit per Investigator judgement), and subsequently have PD during treatment with MEDI4736 alone, to restart combination treatment, if they also meet eligibility criteria.

Section 1.3.2.1 - MEDI4736 + tremelimumab

The potential risks, based on the mechanism of action of MEDI4736 and related molecules, are updated to "colitis, pneumonitis, hepatitis/hepatotoxicity, neuropathy/neuromuscular toxicity, endocrinopathy, dermatitis, and nephritis. In addition, pancreatitis is an important potential risk particularly with MEDI4736 and tremelimumab combination therapy. Other inflammatory responses with potential immune-mediated etiology reported with MEDI4736 and similar molecules include myocarditis, pericarditis, and uveitis" in accordance with the current Investigator Brochure.

Section 3.2 - Exclusion criteria

- Sarcomatoid variant of non-small-cell lung cancer (NSCLC) is added as an exclusion criteria.
- The part of the exclusion criteria of brain metastases or spinal cord compression and off steroids and anticonvulsants for at least 1 month prior to study treatment is amended to and off steroids for at least 14 days prior to study treatment. In addition, following radiotherapy and/or surgery, patients with brain metastases must wait 4 weeks after the intervention and must confirm stability with imaging before randomization.
- Tuberculosis (clinical evaluation) has been added to hepatitis B, hepatitis C and human immunodeficiency virus as part of the active infection exclusion criterion. The exclusion criterion Known history of clinical diagnosis of tuberculosis has been deleted. In addition supplementary information on the diagnoses are given:
  - HIV diagnosis requires positive HIV 1 or 2 antibodies.

- Active hepatitis B virus (HBV) is defined by a known positive HBV surface antigen (HBsAg) result. Patients with a past or resolved HBV infection (defined as the presence of hepatitis B core antibody and absence of HBsAg) are eligible.
- Patients positive for hepatitis C virus (HCV) antibody are eligible only if polymerase chain reaction is negative for HCV ribonucleic acid (RNA).

Section 3.8 - Restrictions

- The restrictions for female patients of childbearing potential are strengthened from 2 methods of effective contraception to at least 1 highly effective method (ie, low failure rate of <1% per year). Additionally the male partners of a female patient of child bearing potential must use a male condom plus spermicide (except in countries where spermicides are not approved).
- Male patients with a female partner of childbearing potential must use a male condom plus spermicide (except in countries where spermicides are not approved), and it is highly recommended for the female partner of a male patient to also use a highly effective method of contraception.

Section 3.9 - Discontinuation of IP

The stipulation for discontinuation of IP of any adverse event (AE) that meets the criteria for discontinuation was removed.

Section 4 – Study plan and timing of procedures

In Table 2 (Schedule of Assessments) the following amendments were:

- The window for Cycle 1 was amended from +1 to +3 days
- Activated partial thromboplastin time (APTT) and international normalized ratio (INR) assessments, which will be performed as clinically indicated and at Screening, were added to the table.
- For the pharmacokinetic samples, it is now specified that they will take place on the same day as the infusion and within 1 hour of end of infusion, as opposed to up to 1 hour predose and with 10 minutes of the end of the infusion.

Section 5.1 - Efficacy Assessments

The timing for performing the baseline assessment was modified to no more than 28 days before randomization rather than no more than 28 days before the start of IP treatment.

Section 5.1.2 - Survival assessments

Patients on treatment or in survival follow-up will now be contacted following the data cut-off for the primary analysis of the PFS and for each interim and final analysis of OS rather than following

the data cut-off for the primary analysis and all subsequent survival analyses to provide complete survival data.

Section 5.2.1 - Laboratory safety assessments

- In Table 5, provision is made that if the amylase and lipase analyses could not be performed in a local laboratory, then 1 or the other would be performed in line with local practice.
- In Table 6, APPT and INR are added to the table, and it is stipulated that they will be assessed at Screening rather than Baseline as indicated in the table footnote.

Section 5.2.4 – Vital signs

Blood pressure may now be measured in a supine or semi-supine position rather than just a supine position.

Section 5.3.1.3 - PRO-CTCAE

The language of "symptom" was replaced with "side effect."

Section 6 – Safety reporting and medical management

In accordance with the new protocol template, the following first 3 sub-sections of Section 6 were reordered **from**:

6.1 Definition of serious adverse events
6.2 Recording of adverse events
6.3 Definition of adverse events

to

6.1 Definition of adverse events
6.2 Definition of serious adverse events
6.3 Recording of adverse events

Section 6.7.1 - MEDI4736 + tremelimumab and MEDI4736 monotherapy adverse events of special interest (AESI)

This section was updated to remove the detailed explanation of MEDI4736's main AESIs and refer the reader to the current IB where these are explained in full, together with specific guidelines for their evaluation and treatment.

Section 7.1 - Identity of IP

The section for MEDI4736 and tremelimumab is updated with the current recommendations for preparation and dose calculations.

Section 7.7 – Concomitant and other treatments

In the section on prohibited medications, the following amends are made:

- Systematic corticosteroid will be permitted for the prevention of chemotherapy-related toxicities (nausea/vomiting prevention and prophylaxis)
- Drugs with laxative properties should be used with caution rather than avoided
- Herbal and natural remedies in general, rather than those for constipation only, should be avoided for 90 days after the last dose of MEDI4736 monotherapy or combination treatment (MEDI4736 + tremelimumab) during the study

Section 8.3 - Definition of analysis set

The outcome variables for the efficacy data were updated in Table 9 to align with the revised co-primary endpoints.

Section 8.4.1 – Calculation or derivation of efficacy variables

BICR of Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST 1.1)-based assessments were moved above Investigator RECIST 1.1-based assessments. Text of BICR-based assessments was changed to reflect BICR-RECIST 1.1 analysis on all radiological scans of all patients rather than a random sample.

Section 8.4.1.2 – Calculation or derivation of efficacy variables and Section 8.5.2 – Objective response rate

Text was added to explain that objective response rate analysis of Investigator tumor data is based upon RECIST 1.1 as a sensitivity analysis.

Section 8.5.1.1 – Progression-free survival, and Section 11 – References

Text was added to describe Investigator assessment (ascertainment bias) and the BICR assessments and method of dealing with discrepancies. New reference was added.

Version 4.0,

Changes to the protocol are summarised in Amendment 3.

Version 3.0,

Changes to the protocol are summarised in Amendment 2.

Version 2.0,
Changes to the protocol are summarised in Amendment 1.

Version 1.0,
Initial document.

This submission document contains confidential commercial information, disclosure of which is prohibited without providing advance notice to AstraZeneca and opportunity to object.

This clinical study protocol has been subject to a peer review according to AstraZeneca standard procedures. The clinical study protocol is publicly registered and the results are disclosed and/or published according to the AstraZeneca Global Policy on Bioethics and in compliance with prevailing laws and regulations.

PROTOCOL SYNOPSIS

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

International Coordinating Investigators

Study site(s) and number of patients planned

The study will enroll approximately 1850 patients to identify approximately 1092 patients who will be randomized to receive MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or platinum-based Standard of Care (SoC) therapy (364 in each group, including approximately 160 patients in each treatment group with programmed cell death ligand 1 [PD-L1]–positive non-small-cell lung cancer [NSCLC], defined as PD-L1 expression ≥25%).

Study period	Phase of development
Estimated date of first patient enrolled	III
Estimated date of last patient completed	III

Study design

This is a randomized, open-label, multi-center, global, Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) wild-type advanced or metastatic NSCLC.

Patients will provide a tumor tissue sample at enrollment to determine PD-L1 expression status (defined by the SP263 PD-L1 immunohistochemistry assay in which:

- ≥25% PD-L1& >1% PD-L1 membrane-expression in tumoral tissue are considered as relevant positive sub-groups
- <25% PD-L1 is considered low/negative
- <1% is considered negative;

For clarity these sub-groups are referred to hereafter as patients with PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$, PD-L1-low/negative or PD-L1-negative tumors, respectively.

Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (PD-L1 positive$_{25\%}$ versus PD-L1-low/negative) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy.

Tumor assessments will be performed every 6 weeks for the first 48 weeks and then every 8 weeks until confirmed disease progression, with categorization of objective tumor response by Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST 1.1).

Objectives

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$, PD-L1-low/negative and/or PD-L1–negative NSCLC. The assessment of progression-free survival (PFS) and overall survival (OS) in patients with PD-L1 positive$_{25\%}$ NSCLC will be considered the co-primary objectives.

Primary Objective:	**Outcome Measure:**
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with PD-L1 positive$_{25\%}$ NSCLC	PFS using Blinded Independent Central Review (BICR) assessments according to RECIST 1.1
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC terms of OS in patients with PD-L1 positive$_{25\%}$ NSCLC	OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of OS in patients with PD-L1 positive$_{25\%}$ NSCLC	OS

Secondary Objectives:	**Outcome Measures:**
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, OS, objective response rate (ORR), duration of response (DoR), proportion of patients alive and progression free at 12 months from randomization (APF12), and time to second progression (PFS2)	OS in PD-L1 positive$_{1\%}$ patients and all patients PFS in PD-L1 positive$_{1\%}$ patients and all patients using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1 positive$_{25\%}$ NSCLC, and patients with PD-L1 positive$_{1\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1 positive$_{25\%}$ NSCLC, and patients with PD-L1 positive$_{1\%}$ NSCLC and all patients using local standard clinical practice[a]
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, ORR, DoR, APF12, and PFS2	PFS in patients with PD-L1 positive$_{25\%}$ NSCLC using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1 positive$_{25\%}$ NSCLC using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1 positive$_{25\%}$ NSCLC using local standard clinical practice[a]
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS, OS, and ORR	PFS and ORR in patients with PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 OS in patients with PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ NSCLC and all patients
To assess disease-related symptoms and health-related quality of life (HRQoL) in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC 30-item core quality of life questionnaire, Version 3 (QLQ-C30 v3) and the 13-item lung cancer quality of life questionnaire (QLQ-LC13) module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.
To assess the pharmacokinetics (PK) of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)

Secondary Objectives:	Outcome Measures:
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of anti-drug antibodies (ADAs) for MEDI4736 and tremelimumab

[a] PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death, based on Investigator tumor assessments.

Safety Objective:	Outcome Measures:
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of advanced or metastatic patients with NSCLC	Adverse events (AEs), physical examinations, laboratory findings, and vital signs

Target patient population

Adult patients (age ≥18 years) with advanced or metastatic (Stage IV) histologically or cytologically documented EGFR and ALK wild-type NSCLC who are treatment naive.

Duration of treatment

Unless specific treatment discontinuation criteria are met, patients in all groups will continue therapy until disease progression.

Progression during treatment

Patients in all groups may continue receiving therapy in the setting of unconfirmed progressive disease (PD) until PD is confirmed, at the Investigator's discretion. According to RECIST 1.1, modified for confirmation of progression, a confirmatory scan will be required following an overall time-point assessment of progression, preferably at the next scheduled visit, and no earlier than 4 weeks after the previous assessment of PD.

Patients in all groups, excluding the SoC arm, with PD according to RECIST 1.1 (unconfirmed and confirmed) who, in the Investigator's opinion, would continue to receive benefit from their assigned treatment and who meet the criteria for treatment in the setting of PD, may continue to receive their assigned treatment for as long as they are still gaining clinical benefit. However, patients in the immunotherapy group(s) will not be permitted to continue immunotherapy if progression occurs after confirmed response (CR, defined by RECIST 1.1) to immunotherapy treatment in the target lesions, ie, the response and progression events both occurred in the target lesions while the patient was receiving immunotherapy during the same treatment period.

Patients in the MEDI4736 + tremelimumab group may restart treatment with the combination therapy if they complete the 4 dosing cycles with MEDI4736 + tremelimumab (with clinical

benefit per Investigator judgment) but subsequently have PD during treatment with MEDI4736 alone and if they meet eligibility criteria for retreatment.

Post final Data Cut Off (DCO)

Patients who continue to receive benefit from their assigned treatment at the final DCO and database closure may continue to receive their assigned treatment for as long as they and their physician feel they are gaining clinical benefit. For patients continuing to receive durvalumab treatment following the final DCO and database closure, it is recommended that the patients continue the scheduled site visits and investigators monitor the patient's safety laboratory results prior to and periodically during treatment with durvalumab in order to manage AEs in accordance with the durvalumab toxicity management guidelines (please see Section 6.1.9).

In the event that a roll-over or safety extension study is available at the time of the final DCO and database closure, patients currently receiving treatment with durvalumab may be transitioned to such a study, and the current study would reach its end. The roll-over or safety extension study would ensure treatment continuation with visits assessment per its protocol. Any patient that would be proposed to move to such study would be given a new Informed Consent.

Investigational product, dosage, and mode of administration

MEDI4736 + tremelimumab combination therapy

- MEDI4736 20 mg/kg via IV infusion every 4 weeks (q4w), starting on Week 0, for up to a total of 4 doses/cycles, and then continue MEDI4736 20 mg/kg via IV infusion q4w, starting on Week 16
- Tremelimumab 1 mg/kg via IV infusion q4w, starting on Week 0, for up to 4 doses/cycles

MEDI4736 monotherapy

- MEDI4736 20 mg/kg via IV infusion q4w, starting on Week 0

Standard of Care therapy

Patients randomized to SoC therapy will receive 1 of the following:

  - Paclitaxel + carboplatin: Paclitaxel 200 mg/m^2 and carboplatin area under the curve (AUC) 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
  - Gemcitabine + cisplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + cisplatin 75 or 80 mg/m2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
  - Gemcitabine + carboplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle +

carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD

- Pemetrexed + cisplatin (non-squamous patients only): Pemetrexed 500 mg/m2 and cisplatin 75 mg/m2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.
- Pemetrexed + carboplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.

Statistical methods

The primary objectives of this study are to assess the efficacy of MEDI4736 + tremelimumab treatment compared with SoC in terms of PFS and OS and of MEDI4736 monotherapy compared with SoC in terms of OS in patients with EGFR and ALK wild-type advanced or metastatic NSCLC with PD-L1 positive$_{25\%}$ tumors. PFS (per RECIST 1.1, using BICR tumor assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression. OS will be defined as the time from the date of randomization until death due to any cause. Thus, primary endpoints in this study are PFS and OS, in patients with PD-L1 positive$_{25\%}$ NSCLC. To control for type I error, an alpha of 0.03 will be used for the OS comparison of MEDI4736 monotherapy versus SoC, an alpha of 0.015 will be used for the OS comparison of MEDI4736 + tremelimumab versus SoC and an alpha of 0.005 will be used for the PFS comparison of MEDI4736 + tremelimumab versus SoC. The study will be considered positive (a success) if any of the OS or PFS analysis results are statistically significant.

Secondary efficacy variables include PFS and OS for MEDI4736 + tremelimumab versus SoC in patients with PD-L1 positive$_{1\%}$,and all randomized patients, PFS for MEDI4736 monotherapy versus SoC in patients with PD-L1 positive$_{25\%}$ tumours, and PFS and OS for MEDI4736+tremelimumab versus MEDI4736 monotherapy in patients with PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ tumors and all randomized patients, as well as ORR, DoR, APF12, and PFS2. All tumor-assessment-related endpoints are as assessed by BICR

Efficacy data will be summarized and analyzed on an intent-to-treat (ITT) basis and the treatment groups will be compared on the basis of randomized treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the ITT population.

Approximately 1092 patients (including approximately 160 patients per arm with PD-L1 positive$_{25\%}$ NSCLC) will be randomized 1:1:1 to MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC. The randomization will be stratified based on PD-L1 tumor expression status (≥25% versus <25%) and histology (squamous versus non-squamous). The primary PFS

analysis for superiority will be performed when both of the following conditions have been met:

- Approximately 231 BICR PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (72% maturity) AND
- Approximately 44 weeks follow up from last patient randomized to the study

The final (primary) OS analysis for superiority will be performed when the following conditions have been met:

- Approximately 225 OS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity) AND
- Approximately 225 OS events have occurred across the MEDI4736 monotherapy and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity)

Two interim analyses to assess the superiority of the MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy group (compared to the SoC group) in terms of OS will be performed at the time of the primary PFS analysis, and when approximately 80% of the target OS events have been reached, respectively.

MEDI4736 + tremelimumab versus SoC (PFS in PD-L1 positive$_{25\%}$ population))

If PFS at 12 months was 34.1% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.59, with 231 PFS events, the trial will have 88% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.690. With an 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (OS in PD-L1 positive$_{25\%}$ population))

If OS at 18 months was 53.2% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average OS HR is 0.62, an estimated 225 death events (70% maturity) are expected to have occurred at 33 months from "first patient in." With at minimum 225 deaths, the study will have 86% power to demonstrate statistical significance at the 2-sided alpha level of 1.32% (with overall alpha for OS 1.5%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.72.

MEDI4736 monotherapy versus SoC (OS in PD-L1 positive$_{25\%}$ population)

If OS at 18 months was 53.2% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS and assuming the true average OS HR is 0.62, an estimated 225 death

events (70% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum of 225 deaths, the study will have 90% power to demonstrate statistical significance at the 2-sided alpha level of 2.58% (with overall alpha for OS 3%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.74.

PFS, based on the programmatically derived PFS from BICR, and OS will be analyzed using a stratified log-rank test The stratification will be by histology alone (squamous, non-squamous) for the primary comparisons using the PD-L1 positive$_{25\%}$ population, and will be by PD-L1 tumor expression status (≥25%, <25%) and histology for the secondary analyses using the PD-L1 positive$_{1\%}$ population and full ITT population. The effect of treatment will be estimated by the HR together with corresponding two-sided appropriately sized CI, (adjusted for interim analyses) and p-value.

Safety data will be summarized descriptively and formal statistical comparisons will not be made.

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

The following abbreviations and special terms are used in this study Clinical Study Protocol.

Abbreviation or special term	Explanation
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse events of special interest
ALK	Anaplastic lymphoma kinase
ALT	Alanine aminotransferase
APTT	Activated partial thromboplastin time;
APF12	Proportion of patients alive and progression free at 12 months from randomization
AST	Aspartate aminotransferase
AUC	Area under the plasma drug concentration-time curve
AUC_{ss}	Area under the curve at steady state
BICR	Blinded Independent Central Review
BoR	Best objective response
BP	Blood pressure
CD	Cluster of differentiation
CI	Confidence interval
C_{max}	Maximum plasma concentration
$C_{max,ss}$	Maximum plasma concentration at steady state
CR	Complete response
CSA	Clinical study agreement
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Event
CTLA-4	Cytotoxic T–lymphocyte-associated antigen 4
$C_{trough,ss}$	Trough plasma concentration at steady state
CXCL	Chemokine (C-X-C motif) ligand
DCR	Disease control rate
DLT	Dose-limiting toxicity
DoR	Duration of response

Abbreviation or special term	Explanation
EC	Ethics Committee, synonymous to Institutional Review Board and Independent Ethics Committee (IEC)
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDoR	Expected duration of response
EGFR	Epidermal growth factor receptor
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQol 5-Dimension, 5-Level health state utility index
ESMO	European Society for Medical Oncology
EU	European Union
FAS	Full Analysis Set
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GI	Gastrointestinal
GMP	Good Manufacturing Practice
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
hCG	Human chorionic gonadotropin
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRQoL	Health-related quality of life
IB	Investigator's Brochure
ICF	Informed consent form
ICH	International Conference on Harmonisation
IDMC	Independent Data Monitoring Committee
IFN	Interferon
IgG	Immunoglobulin G
IHC	Immunohistochemistry
IL	Interleukin
IMT	Immunomodulatory therapy

Abbreviation or special term	Explanation
INR	International normalized ratio
IP	Investigational product
imAE	Immune-mediated adverse event
IRB	Internal Review Board
irRECIST 1.1	Immune-related response criteria, Version 1.1
ITT	Intent-to-Treat
IV	Intravenous
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
mAb	Monoclonal antibody
MDSC	Myeloid-derived suppressor cell
MedDRA	Medical Dictionary for Regulatory Activities
MHLW	Minister of Health, Labor, and Welfare
miRNA	Micro-ribonucleic acid
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NE	Not evaluable
NSCLC	Non–small-cell lung cancer
OAE	Other significant adverse event
ORR	Objective response rate
OS	Overall survival
PBMC	Peripheral blood mononuclear cell
PD	Progressive disease
PD-1	Programmed cell death 1
PD-L1	Programmed cell death ligand 1
PD-L1 positive$_{25\%}$	≥25% of tumor cells with membrane staining for PD-L1 at any intensity
PD-L1-positive$_{1\%}$	≥1% of tumor cells with membrane staining for PD-L1 at any intensity
PD-L1-low/negative	<25% of tumor cells with membrane staining for PD-L1 at any intensity
PD-L1-negative	<1% of tumor cells with membrane staining for PD-L1 at any intensity
PD-L2	Programmed cell death ligand 2
PFS	Progression-free survival

Abbreviation or special term	Explanation
PFS2	Time to second progression
PGIC	Patient's Global Impression of Change
PGx	Pharmacogenetic research
PK	Pharmacokinetic(s)
PR	Partial response
PRO	Patient-reported outcome
PRO-CTCAE	Patient-reported outcomes version of the CTCAE
q2w	Every 2 weeks
q3w	Every 3 weeks
q4w	Every 4 weeks
q8w	Every 8 weeks
q12w	Every 12 weeks
QLQ-C30 v3	30-item core quality of life questionnaire, Version 3
QLQ-LC13	13-item lung cancer quality of life questionnaire
QoL	Quality of life
QTcF	QT interval corrected for heart rate using Fridericia's formula
RECIST 1.1	Response Evaluation Criteria in Solid Tumors, Version 1.1
RNA	Ribonucleic acid
RR	Response rate
RT-QPCR	Reverse transcription quantitative polymerase chain reaction
SAE	Serious adverse event
SAP	Statistical analysis plan
SAS	Safety analysis set
SD	Stable disease
SNP	Single nucleotide polymorphism
SoC	Standard of Care
sPD-L1	Soluble programmed cell death ligand 1
T_3	Triiodothyronine
T_4	Thyroxine
TSH	Thyroid-stimulating hormone
ULN	Upper limit of normal

Abbreviation or special term	Explanation
US	United States
WBDC	Web-Based Data Capture
WHO	World Health Organization

1. INTRODUCTION

1.1 Background and rationale for conducting this study

Lung cancer has been the most common cancer in the world for several decades, with an estimated 1.8 million new cases in 2012 (12.9% of all new cancers), and was also the most common cause of death from cancer in 2012, with 1.59 million deaths (19.4% of the total cancer deaths; GLOBOCAN 2012). Non-small-cell lung cancer (NSCLC) represents approximately 80% to 85% of all lung cancers. Unfortunately, at the time of diagnosis, approximately 70% of patients with NSCLC already have advanced or metastatic disease not amenable to surgical resection. Furthermore, a significant percentage of patients with early stage NSCLC who have undergone surgery subsequently develop distant recurrence and die as a result of their lung cancer (Pisters and LeChevalier 2005).

Despite advances in the diagnosis, imaging, staging, and treatment of NSCLC, the estimated overall 5-year survival for patients in Europe and the United States (US) continues to be low (11% and 17%, respectively; D'Addario et al 2010, Howlander et al 2014). Patients presenting with advanced NSCLC have a median overall survival (OS) of 10 to 12 months (Bonomi 2010). Patients without a targetable mutation (ie, epidermal growth factor receptor [EGFR] or anaplastic lymphoma kinase [ALK] mutation) demonstrate responses to systemic treatment of approximately 20% to 30% and progression-free survival (PFS) of 4 to 5 months (Scagliotti et al 2008, Schiller et al 2002). The durations of responses (DoRs) are also limited, and toxicities can be a major limiting factor. The 1-year survival rate is 30% to 40% for patients with a good performance status. Maintenance therapy, with either continuation or switch, has also been recommended for certain histologic subtypes of NSCLC; for example, maintenance with pemetrexed has been shown to improve OS and PFS, particularly in non-squamous histologies (Ciuleanu et al 2009, Paz-Ares et al 2013).

Common first-line treatment regimens for advanced NSCLC in major global markets are typically platinum-based doublets and include carboplatin and paclitaxel, carboplatin and gemcitabine (squamous only), carboplatin and pemetrexed (non-squamous only), cisplatin and gemcitabine (squamous only), and cisplatin and pemetrexed (non-squamous only). Platinum-based doublet chemotherapy regimens vary to some extent with regard to convenience, associated toxicities, and cost, with the selection of a specific regimen often dictated by local practice and individualized on a case-by-case basis.

1.1.1 Immunotherapies

It is increasingly understood that cancers are recognized by the immune system, and, under some circumstances, the immune system may control or even eliminate tumors (Dunn et al 2004). Studies in mouse models of transplantable tumors have demonstrated that manipulation of co-stimulatory or co-inhibitory signals can amplify T-cell responses against tumors (Peggs et al 2009). This amplification may be accomplished by blocking co-inhibitory molecules, such as cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) or programmed cell death 1(PD-1) , from binding with their ligands, B7 or B7-H1 programmed cell death ligand 1 [PD-L1]).

1.1.2 MEDI4736

MEDI4736 is a human monoclonal antibody (mAb) of the immunoglobulin G (IgG) 1 kappa subclass that inhibits binding of PD-L1 and is being developed by AstraZeneca/MedImmune for use in the treatment of cancer. (MedImmune is a wholly owned subsidiary of AstraZeneca; AstraZeneca/MedImmune will be referred to as AstraZeneca throughout this document.) As MEDI4736 is an engineered mAb, it does not induce antibody-dependent cellular cytotoxicity or complement-dependent cytotoxicity. The proposed mechanism of action for MEDI4736 is interference of the interaction of PD-L1.

PD-L1 is expressed in a broad range of cancers with a high frequency, up to 88% in some types of cancers. In a number of these cancers, including lung, the expression of PD-L1 is associated with reduced survival and an unfavorable prognosis. In lung cancer, only 12% of patients with tumors expressing PD-L1 survived for more than 3 years, compared with 20% of patients with tumors lacking PD-L1 (Mu et al 2011). Based on these findings, an anti-PD-L1 antibody could be used therapeutically to enhance anti-tumor immune responses in patients with cancer. Results of several non-clinical studies using mouse tumor models support this hypothesis, where antibodies directed against PD-L1 or its receptor PD-1 showed anti-tumor activity (Hirano et al 2005, Iwai et al 2002, Okudaira et al 2009, Zhang et al 2008).

MEDI4736 has been given to humans as part of ongoing studies as a single drug or in combination with other drugs. As of 14 July 2014, 509 patients have been enrolled and treated in 10 ongoing clinical studies of MEDI4736 (5 employing MEDI4736 as monotherapy and 5 as combination therapy). Details on the safety profile of MEDI4736 monotherapy are summarized in Section 1.3.2.1. Refer to Appendix G for a complete summary of non-clinical and clinical information and guidance on management of MEDI4736-related toxicities.

MEDI4736 monotherapy exhibits nonlinear (dose-dependent) pharmacokinetics (PK) approaching linearity with a ≥3-mg/kg dose, likely due to saturable target-mediated clearance, and has a half-life of approximately 21 days. Of the 220 patients who received MEDI4736 monotherapy for whom PK/anti-drug antibody (ADA) data were available from Study CD-ON-MEDI4736-1108 (referred to hereafter as Study 1108) as of 14 July 2014, 5 patients (1 patient each in the 0.1- and 3-mg/kg cohorts and 3 patients in the 10-mg/kg cohort) were detected ADA-positive, with an impact on PK/pharmacodynamics in 1 patient in the 3-mg/kg cohort.

1.1.3 Tremelimumab

Tremelimumab, a CTLA-4 mAb of the IgG 2 kappa isotype, is an immunomodulatory therapy (IMT) that is being developed by AstraZeneca for use in the treatment of cancer. Tremelimumab is a human IgG2 mAb directed against CTLA-4.

Binding of CTLA-4 to its target ligands (B7-1 and B7-2) provides a negative regulatory signal, which limits T-cell activation. Anti-CTLA-4 inhibitors antagonize the binding of CTLA-4 to B7 ligands and enhance human T-cell activation as demonstrated by increased cytokine (interleukin [IL]-2 and interferon [IFN] gamma) production in vitro in whole blood

or peripheral blood mononuclear cell (PBMC) cultures (Tarhini and Kirkwood 2008). In addition, blockade of CTLA-4 binding to B7 by anti-CTLA-4 antibodies results in markedly enhanced T-cell activation and anti-tumor activity in animal models, including killing of established murine solid tumors and induction of protective anti-tumor immunity. (Refer to the current version of the tremelimumab investigator brochure [IB] for more information.) Therefore, it is expected that treatment with an anti-CTLA-4 antibody, such as tremelimumab, will lead to increased activation of the human immune system, increasing anti-tumor activity in patients with solid tumors.

An extensive program of non-clinical and clinical studies has been conducted for tremelimumab both as monotherapy and combination therapy with conventional anticancer agents to support various cancer indications using different dose schedules. As of the data cut-off date of 12 November 2014 (for all studies except D4190C00006, which has a cut-off date of 4 December 2014), 973 patients have received tremelimumab monotherapy (not including 497 patients who have been treated in the blinded Phase 2b study, D4880C00003) and 208 patients have received tremelimumab in combination with other agents. Details on the safety profile of tremelimumab monotherapy are summarized in Section 1.3.2.3. Refer to the current tremelimumab IB for a complete summary of non-clinical and clinical information and see Appendix G for guidance on management of tremelimumab-related toxicities.

Tremelimumab exhibited a biphasic PK profile with a long terminal phase half-life of 22 days. Overall, a low incidence of ADAs (<6%) was observed for treatment with tremelimumab.

1.1.4 MEDI4736 in combination with tremelimumab

Targeting both PD-1 and CTLA-4 pathways may have additive or synergistic activity (Pardoll 2012) because the mechanisms of action of CTLA-4 and PD-1 are non-redundant; therefore, AstraZeneca is also investigating the use of MEDI4736 + tremelimumab combination therapy for the treatment of cancer.

Study D4190C00006 is a Phase Ib dose-escalation study to establish safety, PK/pharmacodynamics, and preliminary anti-tumor activity of MEDI4736 + tremelimumab combination therapy in patients with advanced NSCLC. The dosing schedule utilized is MEDI4736 every 2 or 4 weeks (q2w; q4w) up to Week 50 and 48 (12 months), combined with tremelimumab q4w up to Week 24 for 7 doses then every 12 weeks (q12w) for 2 additional doses for up to 12 months. The study is ongoing and continues to accrue.

As of 27 January 2015, a total of 74 patients have been treated in the study, including 58 patients on the q4w dosing schedule and 16 patients on the q2w dosing schedule. Patients have received between 1 and 13 doses of MEDI4736 and between 1 and 9 doses of tremelimumab. Details on the safety profile of MEDI4736 + tremelimumab combination therapy are summarized in Sections 1.2.1 and 1.3.2.4. Refer to the current versions of the MEDI4736 IB and the tremelimumab IB for a complete summary of non-clinical and clinical information and Appendix G for guidance on management of MEDI4736 + tremelimumab-related toxicities.

As of 27 January 2015 in Study D4190C00006, an approximately dose-proportional increase in PK exposure (maximum plasma concentration [C_{max}] and area under the plasma drug concentration-time curve (AUC) from time zero to Day 28 post-dose) of both MEDI4736 and tremelimumab was observed over the dose range of 3 to 20 mg/kg MEDI4736 q4w or q2w and 1 to 10 mg/kg tremelimumab q4w. Four of 60 patients with ADA data available were ADA positive for either anti-MEDI4736 or anti-tremelimumab antibodies post-treatment; MEDI4736 PK was impacted in only 2 of these 4 patients. Complete soluble programmed cell death ligand 1 (sPD-L1) suppression (surrogate for PD-L1 targeting) was observed in almost all patients over the dose range of 3 to 20 mg/kg of MEDI4736 q4w or q2w.

1.1.5 Rationale for conducting this study

Current therapies for advanced NSCLC have poor outcomes (low 5-year survival of 17% for the US), with responses to systemic chemotherapy in the first-line setting of approximately 20% to 30%, and a median OS of approximately 10 to 12 months (Bonomi 2010, D'Addario et al 2010, Scagliotti et al 2008, Schiller et al 2002). Responses are also limited in duration. Systemic chemotherapy is associated with significant side effects, including neutropenia, nausea, vomiting and dehydration, and alopecia (Sandler et al 2006, Scagliotti et al 2008). There is still a significant unmet medical need for additional treatment options for use in this patient population as the 1-year survival rate is 30% to 40%.

As an antibody that blocks the interaction between PD-L1 and its receptors, MEDI4736 may relieve PD-L1-dependent immunosuppressive effects and, therefore, enhance the cytotoxic activity of anti-tumor T-cells. This hypothesis is supported by emerging clinical data from other mAbs targeting the PD-L1/PD-1 pathway, which provide early evidence of clinical activity and a manageable safety profile (Brahmer et al 2012, Topalian et al 2012). Responses have been observed in patients with PD-L1-positive tumors and patients with PD-L1-negative tumors. In addition, MEDI4736 monotherapy has shown durable responses in NSCLC in Study 1108 (see Section 1.3.1.1).

The rationale for combining MEDI4736 and tremelimumab is that the mechanisms of CTLA-4 and PD-1 are non-redundant, suggesting that targeting both pathways may have additive or synergistic activity (Pardoll 2012). In fact, combining immunotherapy agents has been shown to result in improved response rates relative to monotherapy. For example, the concurrent administration of nivolumab and ipilimumab to patients with advanced melanoma induced higher objective response rates (ORRs) than those obtained with single-agent therapy. Importantly, responses appeared to be deep and durable (Wolchok et al 2013). Similar results have been observed in an ongoing study of MEDI4736 + tremelimumab in NSCLC (Antonia et al 2014a), with further updated details presented in this clinical study protocol.

Based on the preliminary clinical efficacy and safety data observed in patients with NSCLC in Study D4190C00006 (with MEDI4736 + tremelimumab) and in Study 1108 (with MEDI4736 monotherapy), AstraZeneca plans to determine the activity of MEDI4736 in combination with tremelimumab and MEDI4736 monotherapy as first-line treatment in patients with NSCLC. The preliminary efficacy, safety, and tolerability data of the MEDI4736 + tremelimumab combination in Study D4190C00006 and MEDI4736 monotherapy in Study 1108 support the

development of these treatments in NSCLC. The primary endpoints of this Phase III study are to determine the activity of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to Standard of Care (SoC) in patients with EGFR and ALK wild-type NSCLC when used as first-line treatment in patients with PD-L1 positive$_{25\%}$ tumors and secondary endpoints will include the comparison of the activity in patients with PD-L1 positive$_{1\%}$ tumours and all randomised patients. In addition to quantifying the clinical effects of MEDI4736 in combination with tremelimumab, this will also allow comparison between the relative efficacy and tolerability of MEDI4736 + tremelimumab and MEDI4736 alone.

1.2 Rationale for study design, doses, and control groups

This study will utilize an open-label design due to the different treatment administration schedules and treatment durations.

1.2.1 MEDI4736 + tremelimumab combination therapy dose rationale

The MEDI4736 + tremelimumab doses and regimen selected for this study are based on the goal of selecting an optimal combination dose of MEDI4736 and tremelimumab that would yield sustained target suppression (sPD-L1), demonstrate promising efficacy, and have an acceptable safety profile.

As of 27 January 2015, a total of 74 patients with advanced NSCLC have been treated in Study D4190C00006. The 74 patients have received between 1 and 9 doses of tremelimumab and between 1 and 13 doses of MEDI4736. Various dose combinations were explored, with doses of tremelimumab ranging from 1 to 10 mg/kg and doses of MEDI4736 ranging from 3 to 20 mg/kg. Fifty-eight of these patients were in the q4w dosing schedule and 16 patients were in the q2w dosing schedule, with the goal of identifying the dose combination that best optimizes the risk:benefit profile in an acceptable range of PK and pharmacodynamic values.

In order to reduce the dosing frequency of MEDI4736 to align with the q4w dosing of tremelimumab, while ensuring an acceptable PK/pharmacodynamic, safety, and efficacy profile, cohorts were narrowed to 15 and 20 mg/kg MEDI4736 q4w. PK simulations from the MEDI4736 monotherapy data indicated that a similar area under the plasma drug concentration-time curve at steady state (AUC_{ss}; 4 weeks) was expected following both 10 mg/kg q2w and 20 mg/kg q4w MEDI4736. The observed MEDI4736 PK data from the D4190C00006 study were well in line with the predicted monotherapy PK data developed preclinically. This demonstrates similar exposure of MEDI4736 20 mg/kg q4w and 10 mg/kg q2w, with no alterations in PK when MEDI4736 and tremelimumab (doses ranging from 1 to 3 mg/kg) are dosed together. While the median C_{max} at steady state ($C_{max,ss}$) is expected to be higher with 20 mg/kg q4w (approximately 1.5 fold) and median trough concentration at steady state ($C_{trough,ss}$) is expected to be higher with 10 mg/kg q2w (approximately 1.25 fold), this is not expected to impact the overall safety and efficacy profile, based on existing preclinical and clinical data.

Monotonic increases in pharmacodynamic activity were observed with increasing doses of tremelimumab relative to the activity observed in patients treated with MEDI4736

monotherapy. There was evidence of augmented pharmacodynamic activity relative to MEDI4736 monotherapy with combination doses containing 1 mg/kg tremelimumab, inclusive of both the 15 and 20 mg/kg MEDI4736 plus 1 mg/kg tremelimumab combinations.

Patients treated with doses of tremelimumab above 1 mg/kg had a higher rate of adverse events (AEs), including discontinuations due to AEs, serious AEs (SAEs), and severe AEs. Between the 10 mg/kg MEDI4736 + 1 mg/kg tremelimumab and 10 mg/kg MEDI4736 + 3 mg/kg tremelimumab cohorts treated at the q2w schedule, the number of patients reporting any AE, Grade 3 AEs, SAEs, and treatment-related AEs was higher in the 10 mg/kg MEDI4736 + 3 mg/kg tremelimumab cohort than the 10 mg/kg MEDI4736 + 1 mg/kg tremelimumab cohort. A similar pattern was noted in the q4w regimens, suggesting that, as the dose of tremelimumab increased above 1 mg/kg, a higher rate of treatment-related events may be anticipated. Further, the SAEs frequently attributed to immunotherapy, pneumonitis and colitis, were more commonly seen in cohorts using either 3 mg/kg or 10 mg/kg of tremelimumab compared to the 1-mg/kg dose cohorts. Together, these data suggest that a combination using a tremelimumab dose of 1 mg/kg appeared to minimize the rate of toxicity when combined with MEDI4736. As a result, all combination doses utilizing either the 3 or 10 mg/kg doses of tremelimumab were eliminated in the final dose selection.

In contrast, cohorts assessing higher doses of MEDI4736 with a constant dose of tremelimumab did not show an increase in the rate of AEs. The data suggested that increasing doses of MEDI4736 may not impact the safety of the combination as much as the tremelimumab dose. Further, safety data between the 10-mg/kg and 20-mg/kg cohorts were similar, with no change in safety events with increasing dose of MEDI4736.

In Study D4190C00006, of all treatment cohorts, the cohort of 11 patients treated in the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab group had the fewest AEs, Grade ≥3 AEs, SAEs, and treatment discontinuations due to AEs, but still showed strong evidence of clinical activity. This cohort had a lower number of treatment-related Grade ≥3 AEs or treatment related SAEs. No dose-limiting toxicities (DLTs) were reported.

Preliminary clinical activity of the MEDI4736 and tremelimumab combination did not appear to change with increasing doses of tremelimumab. The 15- and 20-mg/kg MEDI4736 q4w cohorts demonstrated objective responses at all doses of tremelimumab, and increasing doses of tremelimumab did not provide deeper or more rapid responses.

Efficacy data suggested that the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose cohort may demonstrate equivalent clinical activity to other dose combinations. A total of 5 of 11 patients in the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab cohort were evaluable for efficacy with at least 8 weeks of follow-up. Of these, there were 2 patients (40%) with partial response (PR), 1 patient (20%) with stable disease (SD), and 1 patient (20%) with progressive disease (PD). (The fifth patient had only a single scan, which was conducted outside the window for these evaluations.)

Additionally, of all cohorts, the 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose cohort had the fewest AEs, Grade ≥3 AEs, SAEs, and treatment discontinuations due to AEs, but still

showed some evidence of clinical activity. All together, the data suggested that a 20 mg/kg MEDI4736 + 1 mg/kg tremelimumab dose combination should be selected for further development.

1.2.2 Rationale for 4 cycles of combination therapy followed by MEDI4736 monotherapy

Long-term follow up on melanoma patients treated with ipilimumab, an anti-CTLA-4 targeting antibody (dosed every 3 weeks for 4 doses and then discontinued), shows that patients responding to ipilimumab derive long-term benefit, with a 3-year OS rate of approximately 22%. Furthermore, the survival curve in this population reached a plateau at 3 years and was maintained through 10 years of follow up (Schadendorf et al 2013).

Similar data have been presented for other anti-PD-1/PD-L1 targeting antibodies:

- Nivolumab (anti-PD-1) was dosed q2w for up to 96 weeks in a large Phase I dose-escalation and expansion study, and showed responses were maintained for a median of 22.94 months for melanoma (doses 0.1 mg/kg to 10 mg/kg), 17 months for NSCLC (doses 1, 3, and 10 mg/kg), and 12.9 months for renal cell carcinoma patients (doses 1 and 10 mg/kg) at the time of data analysis (Hodi et al 2014, Brahmer et al 2014, Drake et al 2013). Furthermore, responses were maintained beyond treatment discontinuation in the majority of patients who stopped nivolumab treatment (either due to protocol specified end of treatment, complete response [CR], or toxicity) for up to 56 weeks at the time of data analysis (Topalian et al 2014).

- MPDL3280a (anti-PD-L1) and the combination of nivolumab with ipilimumab, in which patients were dosed for a finite time period and responses maintained beyond treatment discontinuation have been reported (Herbst et al 2013, Wolchok et al 2013).

Similar long term results may be expected with use of other immune-mediated cancer therapeutics including anti-CTLA-4 antibodies such as tremelimumab, anti PD-L1 antibodies such as MEDI4736, or the combination of the two.

The MEDI4736 + tremelimumab combination regimen will be administered for 4 doses followed by monotherapy MEDI4736 20 mg/kg q4w.

1.2.3 MEDI4736 monotherapy dose rationale

A dose of MEDI4736 dose of 20 mg/kg q4w is supported by in-vitro data, non-clinical activity, clinical PK/pharmacodynamics, biomarkers, and activity data from Study 1108 in patients with advanced solid tumors (ongoing first-time-in-humans study).

As of 14 July 2014, 393 patients enrolled in Study 1108 have received MEDI4736, predominantly at 10 mg/kg q2w (either in the dose-escalation or dose-expansion phase of the study). Data presented at the European Society for Medical Oncology (ESMO) meeting 2014

with a later cut-off of 21 August 2014 showed that MEDI4736 was well tolerated at all doses in the NSCLC subset of patients enrolled into Study 1108, with drug-related Grade ≥3 AEs reported in 3% of patients and drug-related AEs leading to discontinuation reported in 1% of patients. No drug-related colitis or hyperglycemia of any grade, no Grade ≥3 pneumonitis reported, and no drug-related AEs leading to death were reported (Antonia et al 2014b). No DLTs were observed up to a dose of 10 mg/kg q2w or 15 mg/kg every 3 weeks (q3w).

Efficacy data on the patients with NSCLC in Study 1108, presented at ESMO 2014 (cut-off date of 21 August 2014), showed a disease control rate (DCR) at 12 weeks of 41% and ORR of 16% among 162 evaluable patients, with activity observed in both squamous and non-squamous histologies. The ORR was higher (25%; 12 CR/PR; n=48) in patients with PD-L1 positive tumors, defined as those with ≥25% of tumor cells with membrane staining for PD-L1, compared to patients with PD-L1 negative tumors (10%; 7 CR/PR; n=74) (Antonia et al 2014b).

Based on available PK/pharmacodynamic data from ongoing Study 1108 with doses ranging from 0.1 to 10 mg/kg q2w or 15 mg/kg q3w, MEDI4736 exhibited non-linear (dose-dependent) PK consistent with target-mediated drug disposition. The PK approached linearity at ≥3 mg/kg q2w, suggesting near complete target saturation (membrane-bound and sPD-L1), and further shows that the MEDI4736 dosing frequency can be adapted to a particular regimen given the linearity seen at higher doses than 3 mg/kg. The expected half-life with doses ≥3 mg/kg q2w is approximately 21 days. A dose-dependent suppression in peripheral sPD-L1 was observed over the dose range studied, consistent with engagement of MEDI4736 with PD-L1. Dose-related changes in a variety of peripheral biomarkers have been observed over the dose range of 0.1 to 15 mg/kg. A low level of immunogenicity has been observed. No patients have experienced immune-complex disease following exposure to MEDI4736. Of the 220 patients who received MEDI4736 monotherapy and for whom PK/ADA data were available as of14 July 2014, 5 were ADA positive, with an impact on PK/pharmacodynamics reported in 1 patient at 3 mg/kg.

Data from Study 006 (Phase I trial in patients with NSCLC using the combination of MEDI4736 and tremelimumab), also show an approximately dose-proportional increase in PK exposure for MEDI4736 over the dose range of 3 to 20 mg/kg MEDI4736 q4w or q2w. The observed MEDI4736 PK data from the combination study were well in line with the predicted monotherapy PK data (5^{th} median and 95^{th} percentiles) for a q4w regimen.

A population PK model was developed using the data from Study 1108 (doses=0.1 to 10 mg/kg q2w or 15 mg/kg q3w; Fairman et al 2014). Multiple simulations indicate that a similar overall exposure is expected following both 10 mg/kg q2w and 20 mg/kg q4w regimens, as represented by AUC_{ss}; 4 weeks. Median $C_{max,ss}$ is expected to be higher with 20 mg/kg q4w (~1.5 fold) and median $C_{trough,ss}$ is expected to be higher with 10 mg/kg q2w (~1.25 fold). Clinical activity with the 20 mg/kg q4w dosing regimen is anticipated to be consistent with 10 mg/kg q2w with the proposed similar dose of 20 mg/kg q4w expected to (a) achieve complete target saturation in majority of patients; (b) account for anticipated variability in PK, pharmacodynamic, and clinical activity in diverse cancer populations; (c)

maintain sufficient PK exposure in case of ADA impact; and (d) achieve PK exposure that yielded maximal anti-tumor activity in animal models.

As of 8 April 2015, there is initial safety data for 16 patients receiving the 20 mg/kg q4w dosing regimen (12 patients from Study 1108 and 4 patients from the Japan Phase I trial). The toxicities observed with 20 mg/kg q4w are consistent with the 10 mg q2w regimen, and there were no DLTs observed. Of the 12 patients in Study 1108, 42% of patients have experienced any grade AE, with 2 being Grade 3 and above (17%). None of the Grade 3 and higher events was considered treatment-related. No patients on the Japan Phase I trial have experienced a Grade 3 or above AE. At present, the data do not suggest that the safety profile of MEDI4736 will be different in the 20 mg/kg q4w dosing regimen when compared to 10 mg/kg q2w regimen.

Given the similar AUC and modest differences in median peak and trough levels at steady state, the observation that both regimens maintain complete sPD-L1 suppression at trough, and the available clinical data, the 20 mg/kg q4w and 10 mg/kg q2w regimens are expected to have similar efficacy and safety profiles, supporting further development with a dose of 20 mg/kg q4w.

1.2.4 Rationale for Standard of Care as a comparator

The choice of SoC options provided in this study includes carboplatin and paclitaxel, carboplatin (or cisplatin) and gemcitabine (squamous only), carboplatin (or cisplatin) and pemetrexed (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only). Patients in the SoC group will receive treatment determined by the Investigator, from the SoC agents approved for use in NSCLC in their local market, until progression per standard practice. The SoC options provided in this study include agents that are commonly used in advanced or metastatic NSCLC and allow sufficient flexibility for Investigators and patients to select the agents that reflect their normal clinical practice and national guidelines (NCCN 2014 and Reck et al 2014).

1.2.5 Rationale for retreatment option

In contrast to patients treated with chemotherapy, who are unlikely to respond to rechallenge with the same agent upon progression, responses have been observed upon retreatment with IMTs. Several potential mechanisms of resistance to IMT exist, including loss of T-cell "memory" or recurrence of immune escape, which suggest retreatment for patients who initially respond or demonstrate SD is reasonable. Preliminary data in patients previously treated with IMTs suggest that responses are similar to those observed following initial treatment (Forde et al 2014; Hodi et al 2010).

Patients in the MEDI4736 + tremelimumab group who complete 4 dosing cycles (with clinical benefit per Investigator judgement), and subsequently have PD during treatment with MEDI4736 alone may restart treatment if they meet eligibility criteria for retreatment.

1.2.6 Rationale for endpoints

The primary objectives of this study are to assess the efficacy of:

- MEDI4736 + tremelimumab treatment compared with SoC in terms of PFS in patients with EGFR and ALK wild-type advanced or metastatic NSCLC with PD-L1 positive$_{25\%}$ tumors
- MEDI4736 + tremelimumab treatment compared with SoC in terms of OS in patients with EGFR and ALK wild-type advanced or metastatic NSCLC with PD-L1 positive$_{25\%}$ tumors
- MEDI4736 treatment compared with SoC in terms of OS in patients with EGFR and ALK wild-type advanced or metastatic NSCLC with PD-L1 positive$_{25\%}$ tumors

Within the statistical hierarchy, a key secondary endpoint of the study is to assess the efficacy in the overall study population:

- MEDI4736 + tremelimumab treatment compared with SoC in terms of PFS and OS in patients with EGFR and ALK wild-type advanced or metastatic NSCLC with PD-L1 positive$_{1\%}$ tumors and in the overall NSCLC population

Emerging data in immuno-oncology suggest that the treatment benefit of immunotherapies can more strongly manifest in OS compared to PFS (Borghaei et al 2015, Brahmer et al 2015, Fehrenbacher et al 2016), and therefore, both PFS and OS are co-primary endpoints. Additionally, the emerging immuno-oncology data indicate an improved treatment effect in patients with NSCLC tumors that express PD-L1 for both combination and monotherapy (Hellman et al 2016, Reck et al 2016). Therefore, the study characterizes the efficacy of MEDI4736 + tremelimumab and MEDI4736 monotherapy in the treatment in the PD-L1-positive NSCLC population (≥25%, ≥1%) and in the overall NSCLC population.

The secondary efficacy endpoints of ORR, DoR, proportion of patients alive and progression-free and time to second progression (PFS2) will be examined to further evaluate the anti-tumor effect and survival benefit of MEDI4736 + tremelimumab versus SoC. The secondary efficacy endpoints of PFS, ORR and PFS2 will also be examined to further evaluate the anti-tumor effect and survival benefit of MEDI4736 monotherapy compared to SoC. PFS and ORR will be examined to further evaluate the anti-tumor effect of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy. ORR and DoR will be assessed using Blinded Independent Central Review (BICR) assessments according to Response Evaluation Criteria in Solid Tumors, Version 1.1 (RECIST 1.1).

Anti-tumor activity will be evaluated by irRECIST 1.1 using BICR assessments as an exploratory endpoint in order to assess the utility of modified RECIST criteria in capturing the true clinical benefit of MEDI4736 and tremelimumab. Using irRECIST 1.1 will allow taking into account the unique kinetics of responses that have been observed and well characterized with this class of agents, including in ongoing monotherapy studies.

The secondary health-related quality of life (HRQoL) assessments (the European Organisation for Research and Treatment of Cancer [EORTC] 30-item core quality of life questionnaire,

Version 3 [QLQ-C30 v3] and 13-item lung cancer quality of life questionnaire [QLQ-LC13]) will show the overall influence of the benefits and toxicity of the treatment from a patient's perspective and will aid in understanding of the benefit/risk evaluation. These patient-reported outcome (PRO) questionnaires are well-established instruments that have been previously included in lung cancer clinical trials.

The PK and immunogenicity of MEDI4736 and tremelimumab are being examined to assess the PK of both agents when administered in combination, and to assess their potential impact on PK, pharmacodynamics, and safety and efficacy parameters. Biological samples will be used to explore potential biomarkers in tumor, plasma, and/or serum that may influence pathogenesis, response, and clinical characteristics.

1.3 Benefit/risk and ethical assessment

The following sections include summaries of the potential benefits and risks associated with MEDI4736 monotherapy, tremelimumab monotherapy, and MEDI4736 + tremelimumab combination therapy, respectively, prior to the overall benefit:risk assessment.

1.3.1 Potential benefits

1.3.1.1 MEDI4736

The majority of the safety and efficacy data currently available for MEDI4736 are based on the first time in-human, single-agent study (Study 1108) in patients with advanced solid tumors. Data from Study 1108 were presented at the ESMO 2014 Congress. As of 21 August 2014, 162 patients with NSCLC were evaluable for response analysis. The DCR at 12 weeks in patients receiving 10 mg/kg MEDI4736 q2w was 39%, and the ORR was 15% (26% [12 out of 47 patients] with known PD-L1-positive NSCLC [ie, ≥25% PD-L1 expression] and 10% [7 out of 74 patients] with known PD-L1 low/negative NSCLC [ie, <25% PD-L1 expression]). A total of 24% of patients receiving 10 mg/kg MEDI4736 q2w had SD for ≥12 weeks (including 21% [10 out of 47 patients] with known PD-L1-positive NSCLC and 32% [24 out of 74 patients] with known PD-L1-negative NSCLC). Responses were ongoing in 88% of patients with NSCLC, with an objective response duration ranging from 0.1 to 32.4 weeks (Antonia et al 2014b).

1.3.1.2 Tremelimumab

In a single-group, Phase II study (Study A3671008) of tremelimumab administered at 15 mg/kg every 90 days to patients with refractory melanoma, a response rate (RR) of 7% and a median OS of 10 months in the second-line setting (as compared to approximately 6 months with best supportive care reported from a retrospective analysis, Korn et al 2008) were observed (Kirkwood et al 2010). In a randomized, open-label, first-line Phase III study of tremelimumab (administered at 15 mg/kg every 90 days) versus chemotherapy (dacarbazine or temozolomide) in advanced melanoma (Study A3671009), results of the final analysis showed an RR of 11% and a median OS of 12.58 months in this first-line setting as compared to 10.71 months with standard chemotherapy; however, these results were not statistically significant (Ribas et al 2013). Additionally, a Phase II maintenance study (Study A3671015) in patients with Stage IIIB or IV NSCLC who had responded or remained stable failed to

achieve statistical significance. The primary endpoint of PFS at 3 months was 22.7% in the tremelimumab group (15 mg/kg) compared with 11.9% in the best supportive care group (Study A3671015).

1.3.1.3 MEDI4736 + tremelimumab

The preclinical and clinical justification for this combination as noted in Section 1.2.1 also supports the synergy of this combination. Available data, such as those presented by Wolchok et al, suggest that the combination of agents targeting PD-1/PD-L1 and CTLA-4 may have profound and durable benefits in patients with melanoma (Wolchok et al 2013). Further, preliminary efficacy data from Study D4190C00006 has demonstrated that this combination is clinically active and well tolerated. As of 27 January 2015, 53 patients were evaluable for response across various MEDI4736 + tremelimumab dose regiments. Of these, 12 patients (23%) had a best response of PR and 14 patients (26%) had a best response of SD. In the MEDI4736 20 mg/kg plus tremelimumab 1 mg/kg q4w cohort, a total of 5 of 11 patients were evaluable for efficacy with at least 8 weeks of follow-up. Of these, there were 2 patients (40%) with PR, 1 patient (20%) with SD, and 1 patient (20%) with PD. (The fifth patient had only a single scan, which was conducted outside the window for these evaluations.)

Current experience with single agent IMT studies suggests that clinical responses may be restricted to a subset of any given patient population and that it might be beneficial to enrich the patient population by selecting patients likely to respond to therapy. To date, no assay has been established or validated, and no single approach has proven accurate, for patient enrichment for IMTs. However, independent data from multiple sources using different assays and scoring methods suggests that PD-L1 expression on tumor cells and/or tumor infiltrating cells may be associated with greater clinical benefit.

Data from ongoing studies with MEDI4736 and other agents targeting the PD-1/PD-L1 pathway suggest, as shown in a number of tumor types (eg, NSCLC, renal cell carcinoma, and melanoma), that monotherapy may be more efficacious (in terms of ORR) in patients who are PD-L1-positive.

Given the findings, a number of ongoing studies are assessing the activity of agents in patients with PD-L1–positive tumors. Though biomarker development is ongoing, and the final boundaries of these populations are yet to be established, there is also an unmet medical need in patients with PD-L1–low/negative tumors that needs to be addressed. Data, as of 27 January 2015 from Study 006 show that with the addition of tremelimumab to MEDI4736, the ORR can be increased to 25% in patients with PD-L1-low/negative NSCLC (as defined by the SP263 PD-L1 immunohistochemistry [IHC] assay). As patients with PD-L1 positive(≥25%) tumors can also have an increase in ORR, from 25% with MEDI4736 monotherapy, to 36% with the combination of MEDI4736 and tremelimumab, the study will enroll all patients with NSCLC, with efficacy analysis performed in the overall study population, and pre-defined PD-L1 subgroups.

1.3.2 Potential risks

1.3.2.1 Overall risks

Monoclonal antibodies directed against immune checkpoint proteins, such as programmed cell death ligand 1 (PD-L1) as well as those directed against programmed cell death-1 (PD-1) or cytotoxic T-lymphocyte antigen-4 (CTLA-4), aim to boost endogenous immune responses directed against tumor cells. By stimulating the immune system however, there is the potential for adverse effects on other tissues.

Most adverse drug reactions seen with the immune checkpoint inhibitor class of agents are thought to be due to the effects of inflammatory cells on specific tissues. These risks are generally events with a potential inflammatory or immune mediated mechanism and which may require more frequent monitoring and/or unique interventions such as immunosuppressants and/or endocrine therapy. These immune mediated effects can occur in nearly any organ system, and are most commonly seen as gastrointestinal AEs such as colitis and diarrhoea, pneumonitis/interstitial lung disease (ILD), hepatic AEs such as hepatitis and liver enzyme elevations, skin events such as rash and dermatitis and endocrinopathies including hypo- and hyper-thyroidism.

1.3.2.2 MEDI4736

Risks with durvalumab include, but are not limited to, diarrhea/colitis and intestinal perforation, pneumonitis/ILD, endocrinopathies (hypo- and hyper-thyroidism, type I diabetes mellitus, hypophysitis and adrenal insufficiency) hepatitis/increases in transaminases, nephritis/increases in creatinine, pancreatitis/increases in amylase and lipase, rash/pruritus/dermatitis, myocarditis, myositis/polymyositis, other rare or less frequent inflammatory events including neurotoxicities, infusion-related reactions, hypersensitivity reactions and infections/serious infections.

For information on all identified and potential risks with durvalumab please always refer to the current version of the durvalumab IB.

In monotherapy clinical studies AEs (all grades) reported very commonly (≥ 10% of patients) are fatigue, nausea, decreased appetite, dyspnea, cough, constipation, diarrhea, vomiting, back pain, pyrexia, asthenia, anemia, arthralgia, peripheral edema, headache, rash, and pruritus. Approximately 9% of patients experienced an AE that resulted in permanent discontinuation of durvalumab and approximately 6% of patients experienced an SAE that was considered to be related to durvalumab by the study investigator.

The majority of treatment-related AEs were manageable with dose delays, symptomatic treatment, and in the case of events suspected to have an immune basis, the use of established treatment guidelines for immune-mediated toxicity (please see Section 6.7.2).

A detailed summary of durvalumab monotherapy AE data can be found in the current version of the durvalumab IB.

1.3.2.3 Tremelimumab

Risks with tremelimumab monotherapy include, but are not limited to, GI effects (colitis, diarrhoea, enterocolitis and intestinal perforation), endocrine disorders (hypo and hyperthyroidism, hypophysitis and adrenal insufficiency), skin effects (rash, and pruritus), elevations in lipase and amylase and clinical manifestations of pancreatitis, other gastrointestinal events e.g. ulcerative colitis, dehydration, nausea and vomiting; hepatic events including hepatitis, and liver enzyme elevations; pneumonitis and ILD; nervous system events including encephalitis, peripheral motor and sensory neuropathies, Guillain-Barre and proximal muscle weakness; cytopenias including thrombocytopenia, anemia and neutropenia; infusion-related reactions, anaphylaxis, and allergic reactions; renal events including renal failure, acute kidney injury, nephritis, nephrotic syndrome, autoimmune nephritis and electrolyte abnormalities such as hypokalemia; autoimmune diseases including autoimmune arthritis, Sjogren's syndrome and giant cell temporal arteritis; hyperglycemia and diabetes mellitus; and pyrexia.

For information on all identified and potential risks with tremelimumab please always refer to the current version of the tremelimumab IB.

Using pooled data from monotherapy clinical studies AEs (all grades) reported very commonly (≥ 10% of patients) were diarrhea, nausea, fatigue, pruritus, decreased appetite, rash, vomiting, dyspnoea, constipation, cough, pyrexia, abdominal pain, decreased weight, headache, asthenia, and anaemia. Approximately 16% of patients experienced an AE that resulted in permanent discontinuation of tremelimumab and approximately 45% of patients experienced an SAE.

A detailed summary of tremelimumab monotherapy AE data can be found in the current version of the tremelimumab IB.

1.3.2.4 MEDI4736 + tremelimumab

The safety of durvalumab + tremelimumab combination therapy was initially evaluated in the ongoing dose escalation and dose expansion Study 006, in patients with NSCLC, and is being studied in a number of other ongoing clinical trials, in a number of different indications, and has to date shown a manageable safety and tolerability profile.

The types of risks with the combination of durvalumab + tremelimumab (based on an equivalent durvalumab dose of 20m/kg and a tremelimumab dose of 1mg/kg) are similar to those for durvalumab and tremelimumab monotherapy. Emerging data from study 006, other studies evaluating the combination, and from combinations of other agents in the same class indicate an increased frequency and/or severity of some of these immune-mediated toxicities.

For information on all identified and potential risks with the durvalumab+tremelimumab combination please always refer to the current version of the durvalumab IB

In durvalumab+tremelimumab combination studies at the dose of durvalumab 20mg/kg and tremelimumab 1mg/kg AEs (all grades) reported very commonly (≥ 10% of patients) are fatigue, diarrhoea,nausea, dyspnea, decreased appetite, pruritus, vomiting, anaemia, constipation, cough, abdominal pain, pyrexia, back pain, arthralgia, hypothyroidism, asthenia, oedema peripheral, weight, decreased hyponatraemia and rash.

Approximately 15% of patients experienced an AE that resulted in permanent discontinuation of study drug and approximately 15% of patients experienced an SAE that was considered to be related to durvalumab and tremelimumab by the study investigator.

A detailed summary of durvalumab + tremelimumab combination AE data can be found in the current version of the durvalumab IB.

1.3.3 Overall benefit-risk and ethical assessment

There remains a significant unmet medical need for additional treatment options for patients with EGFR and ALK wild-type, advanced or metastatic NSCLC who have not received prior chemotherapy or any systemic therapy for advanced or metastatic NSCLC.

The study design aims to minimize potential risks; intensive monitoring, including early safety assessment, is in place for those risks deemed to be most likely based on prior experience with the IPs (ie, MEDI4736 + tremelimumab, MEDI4736, and SoC).

Based upon the available non-clinical and clinical safety data, the limited survival benefit provided by the currently available treatment options to patients, the limited life expectancy due to malignant disease, the activity seen with MEDI4736 in this tumor type, and the strength of the scientific hypotheses under evaluation, the MEDI4736 + tremelimumab combination and the MEDI4736 monotherapy treatments proposed for evaluation in this study may have the potential to provide meaningful clinical benefit with a manageable safety and tolerability profile by generating durable clinical responses, thereby improving quality of life (QoL) and potentially extending survival.

Therefore, the investigation of the potential therapeutic efficacy of the combination of MEDI4736 with tremelimumab and MEDI4736 monotherapy in patients with PD-L1-positive and -negative tumors is acceptable, and the overall benefit/risk assessment supports the proposed study design.

1.4 Study design

This is a randomized, open-label, multi-center, global Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with EGFR and ALK wild-type advanced or metastatic NSCLC. A schematic diagram of the overall study design is shown in Figure 1, and a detailed study flow chart is shown in Figure 2.

This study will enroll approximately 1850 patients at sites in North America, Asia, Australia, and Europe to randomize approximately 1092 patients (including approximately 160 patients with PD-L1 positive$_{25\%}$ NSCLC in each treatment group) to treatment.

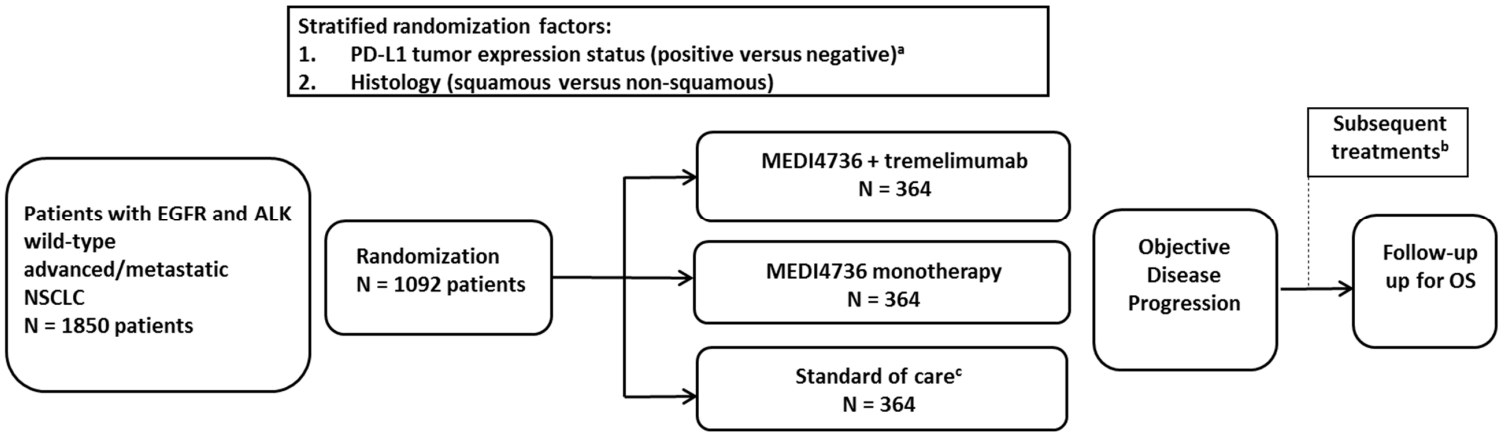
Patients will provide a tumor tissue sample at enrollment (newly acquired or archived sample <3 months old) to determine PD-L1 expression status at 25% (defined by the SP263 PD-L1 IHC assay in which:

- ≥25% PD-L1& >1% PD-L1 membrane-expression in tumoral tissue are considered as relevant positive sub-groups
- <25% PD-L1 is considered low/negative
- <1% is considered negative;

Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (PD-L1 positive$_{25\%}$ versus PD-L1-low/negative) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy.

Doses and treatment regimens are described in Section 7.2. Assessments will be conducted as indicated in Table 2, Table 3, and Table 4.

Figure 1 Overall study design

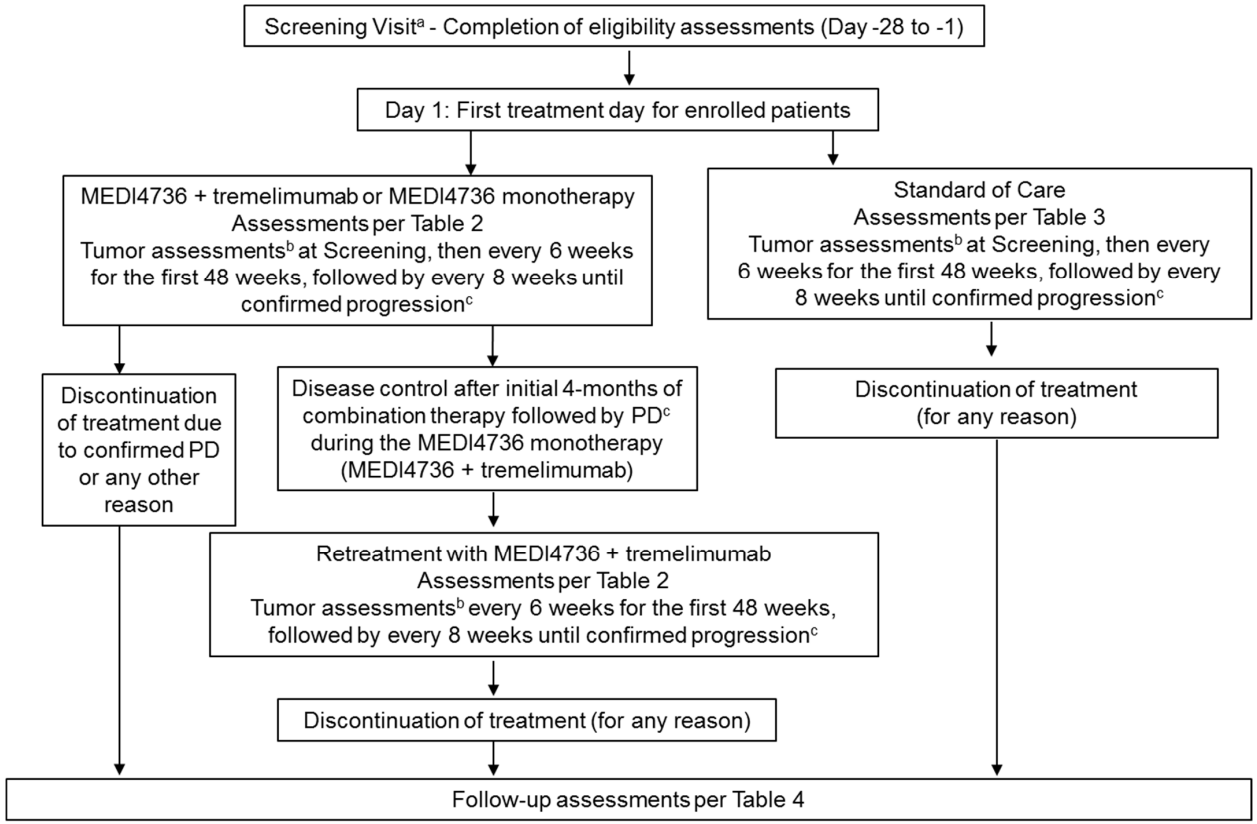


a Statification by PD-L1 membrane-expression in tumoral tissue (≥25%, <25%). Sites will be supplied with PD-L1 status upon request at disease progression

b Offer of standard chemotherapy per Investigator discretion

c Standard of Care is an Investigator choice from the following: paclitaxel + carboplatin, gemcitabine + cisplatin (or carboplatin) (squamous only), pemetrexed + cisplatin (or carboplatin) (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only following pemetrexed/platinum induction)

Figure 2 Study flow chart



[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization.
[b] Tumor assessments were performed using RECIST 1.1
[c] A confirmatory scan is always required following the initial demonstration of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). (See Section 5.1 for more information.)

2. STUDY OBJECTIVES

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1-positive and/or PD-L1–negative NSCLC. The assessment of PFS and OS in all patients and PFS in patients with PD-L1-positive NSCLC will be considered co-primary objectives.

2.1 Primary objective

Primary Objective:	**Outcome Measure:**
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with PD-L1 positive$_{25\%}$ NSCLC	PFS using Blinded Independent Central Review (BICR) assessments according to RECIST 1.1
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC terms of OS in patients with PD-L1 positive$_{25\%}$ NSCLC	OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of OS in patients with PD-L1 positive$_{25\%}$ NSCLC	OS

2.2 Secondary objectives

Secondary Objectives:	**Outcome Measures:**
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, OS, objective response rate (ORR), duration of response (DoR), proportion of patients alive and progression free at 12 months from randomization (APF12), and time to second progression (PFS2)	OS in PD-L1 positive$_{1\%}$ patients and all patients PFS in PD-L1 positive$_{1\%}$ patients and all patients using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1 positive$_{25\%}$ NSCLC, and patients with PD-L1 positive$_{1\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1 positive$_{25\%}$ NSCLC, and patients with PD-L1 positive$_{1\%}$ NSCLC and all patients using local standard clinical practice[a]
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, ORR, DoR, APF12, and PFS2	PFS in patients with PD-L1 positive$_{25\%}$ NSCLC using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1 positive$_{25\%}$ NSCLC using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1 positive$_{25\%}$ NSCLC using local standard clinical practice[a]
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS, OS, and ORR	PFS and ORR in patients with PD-L1 positive$_{25\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 OS in patients with PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ NSCLC and all patients
To assess disease-related symptoms and health-related quality of life (HRQoL) in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC 30-item core quality of life questionnaire, Version 3 (QLQ-C30 v3) and the 13-item lung cancer quality of life questionnaire (QLQ-LC13) module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.
To assess the pharmacokinetics (PK) of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of anti-drug antibodies (ADAs) for MEDI4736 and tremelimumab

[a] PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death, based on Investigator tumor assessments.

2.3 Safety objective

Safety Objective:	Outcome Measures:
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of advanced or metastatic patients with NSCLC	Adverse events (AEs), physical examinations, laboratory findings, and vital signs

2.4 Exploratory objectives

Exploratory Objectives:	Outcome Measures:
To explore irRECIST 1.1 as an assessment methodology for clinical benefit of MEDI4736 + tremelimumab compared to SoC with assessment by BICR	PFS and ORR using BICR assessment according to irRECIST 1.1
To assess AEs by patient self-reporting of specific CTCAE symptoms	Collection of approximately 20 patient-reported outcomes version of CTCAE (PRO-CTCAE) symptoms via an electronic device solution
To assess patients' overall impression of the change in their health status since the start of study treatment	Patients' Global Impression of Change (PGIC) item will be collected directly from patients via an electronic device solution
To investigate the relationship between PK exposure and clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze PK exposure and the relationship with clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, as deemed appropriate
To describe and evaluate resource use associated with assigned treatments and underlying disease during assigned treatment	Health resource utilization measures including hospitalization, outpatient visits, or emergency department visits
To explore the impact of treatment and disease state on health state utility using the EQ-5D-5L during assigned treatment	EQ-5D-5L health state utility index will be used to derive health state utility based on patient-reported data

Exploratory Objectives:	Outcome Measures:
To investigate associations between pre-treatment peripheral myeloid-derived suppressor cells (MDSCs) measures and clinical activity	A graphical and/or a data modeling approach will be used to analyze the relationship between MDSC counts with clinical outcomes and/or with tumor lesion measurements
To investigate the relationship between biomarkers and clinical outcomes, efficacy, AEs, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze biomarkers (eg, IFNγ and/or PD-L1 expression defined under alternative methods) and the relationship with clinical outcomes, efficacy, AEs, and/or safety parameters, as deemed appropriate

Note: Exploratory objective analyses may be reported separately from the main clinical study report.

3. PATIENT SELECTION, ENROLLMENT, RANDOMIZATION, RESTRICTIONS, DISCONTINUATION, AND WITHDRAWAL

Each patient must meet all of the inclusion criteria (Section 3.1) and none of the exclusion criteria (Section 3.2) for this study. Under no circumstances will there be exceptions to this rule.

3.1 Inclusion criteria

For inclusion in the study, patients should fulfill the following criteria:

1. Age ≥18 years at the time of screening

2. Written informed consent and any locally required authorization (eg, Health Insurance Portability and Accountability Act in the US, European Union [EU] Data Privacy Directive in the EU) obtained from the patient/legal representative prior to performing any protocol-related procedures, including screening evaluations. (For patients aged <20 years and enrolling in Japan, a written informed consent should be obtained from the patient and his or her legally acceptable representative.)

3. Histologically or cytologically documented Stage IV NSCLC not amenable to curative surgery or radiation (according to Version 7 of the International Association for the Study of Lung Cancer Staging Manual in Thoracic Oncology; IASLC Staging Manual in Thoracic Oncology).

4. Patients must have tumors that lack sensitizing EGFR mutation (eg, exon 19 deletion or exon 21 L858R, exon 21 L861Q, exon 18 G719X, or exon 20 S768I mutation) and ALK rearrangement. (If a patient has squamous histology or is known to have a tumor with a KRAS mutation, then EGFR and ALK testing is not required).

5. No prior chemotherapy or any other systemic therapy for advanced or metastatic NSCLC. Patients who have received prior platinum-containing adjuvant, neoadjuvant, or definitive chemoradiation for advanced disease are eligible, provided that progression has occurred >6 months from last therapy.

6. Tumor PD-L1 status, with the SP263 PD-L1 IHC assay confirmed by a reference laboratory, must be known prior to randomization. As such, all patients must be able to undergo a fresh tumor biopsy during screening or to provide an available tumor sample taken <3 months prior to enrollment . Tumor lesions used for fresh biopsies should not be target lesions, unless there are no other lesions suitable for biopsy. Fine needle aspirate specimens are not acceptable. Specimens from metastatic bone lesions are typically unacceptable unless there is a significant soft tissue component. The tumor specimen submitted to establish eligibility should be of sufficient quantity to allow for PD-L1 IHC and other exploratory biomarker analyses and is preferred in formalin-fixed paraffin embedded blocks.

7. WHO/ECOG performance status of 0 or 1 at enrollment.

8. At least 1 lesion, not previously irradiated, that can be accurately measured at baseline as ≥10 mm in the longest diameter (except lymph nodes which must have a short axis ≥15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and that is suitable for accurate repeated measurements as per RECIST 1.1 guidelines.

9. No prior exposure to immune-mediated therapy including, but not limited to, other anti-CTLA-4, anti-PD-1, anti-PD-L1, and anti-programmed cell death ligand 2 anti-(PD-L2) antibodies, excluding therapeutic anticancer vaccines.

10. Adequate organ and marrow function as defined below:

    - Hemoglobin ≥9.0 g/dL
    - Absolute neutrophil count $\geq 1.5 \times 10^9$ /L
    - Platelet count $\geq 100 \times 10^9$/L
    - Serum bilirubin ≤1.5 × the upper limit of normal (ULN). This will not apply to patients with confirmed Gilbert's syndrome, who will be allowed in consultation with their physician.
    - Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) ≤2.5 × ULN; for patients with hepatic metastases, ALT and AST ≤5 × ULN
    - Calculated creatinine clearance ≥50 mL/min as determined by Cockcroft-Gault (using actual body weight) or 24 hour urine collection

Males:

$$\text{Creatinine clearance (mL/min)} = \frac{\text{Weight (kg)} \times (140 - \text{Age})}{72 \times \text{serum creatinine (mg/dL)}}$$

Females:

$$\text{Creatinine clearance (mL/min)} = \frac{\text{Weight (kg)} \times (140 - \text{Age})}{72 \times \text{serum creatinine (mg/dL)}} \times 0.85$$

11. Evidence of post-menopausal status or negative urinary or serum pregnancy test for female pre-menopausal patients. Women will be considered post-menopausal if they have been amenorrheic for 12 months without an alternative medical cause. The following age-specific requirements apply:

    – Women <50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of exogenous hormonal treatments and if they have luteinizing hormone and follicle-stimulating hormone levels in the post-menopausal range for the institution or underwent surgical sterilization (bilateral oophorectomy or hysterectomy).

    – Women ≥50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of all exogenous hormonal treatments, had radiation-induced oophorectomy with last menses >1 year ago, had chemotherapy-induced menopause with >1 year interval since last menses, or underwent surgical sterilization (bilateral oophorectomy or hysterectomy).

3.2 Exclusion criteria

Patients should not enter the study if any of the following exclusion criteria are fulfilled:

1. Involvement in the planning and/or conduct of the study (applies to both AstraZeneca staff and/or staff at the study site)

2. Concurrent enrollment in another clinical study, unless it is an observational (non-interventional) clinical study or the follow-up period of an interventional study

3. Mixed small-cell lung cancer and NSCLC histology, sarcomatoid variant

4. Any concurrent chemotherapy, IP, biologic, or hormonal therapy for cancer treatment. Concurrent use of hormonal therapy for non-cancer-related conditions (eg, hormone replacement therapy) is acceptable.

5. Radiotherapy treatment to more than 30% of the bone marrow or with a wide field of radiation within 4 weeks of the first dose of study drug. Note: Local treatment of isolated lesions, excluding target lesions, for palliative intent is acceptable.

6. Major surgical procedure (as defined by the Investigator) within 28 days prior to the first dose of IP. Note: Local surgery of isolated lesions for palliative intent is acceptable.

7. History of allogenic organ transplantation

8. Active or prior documented autoimmune or inflammatory disorders (including inflammatory bowel disease [eg, colitis or Crohn's disease], diverticulitis with the exception of diverticulosis, celiac disease or other serious gastrointestinal chronic conditions associated with diarrhea), systemic lupus erythematosus, Sarcoidosis syndrome, or Wegener syndrome (granulomatosis with polyangiitis), Graves' disease, rheumatoid arthritis, hypophysitis, uveitis, etc within the past 3 years prior to the start of treatment. The following are exceptions to this criterion:

   – Patients with vitiligo or alopecia

   – Patients with hypothyroidism (eg, following Hashimoto syndrome) stable on hormone replacement or psoriasis not requiring systemic treatment

9. Any condition that, in the opinion of the Investigator, would interfere with the evaluation of IP or interpretation of patient safety or study results, including, but not limited to, ongoing or active infection, symptomatic congestive heart failure, uncontrolled hypertension, unstable angina pectoris, cardiac arrhythmia, interstitial lung disease, or psychiatric illness/social situations that would limit compliance with study requirement, substantially increase risk of incurring AEs from MEDI4736 or tremelimumab, or compromise the ability of the patient to give written informed consent

10. Medical contraindication to platinum (cisplatin or carboplatin)-based doublet chemotherapy

11. History of another primary malignancy except for

   – Malignancy treated with curative intent and with no known active disease ≥5 years before the first dose of study drug and of low potential risk for recurrence

   – Adequately treated non-melanoma skin cancer or lentigo maligna without evidence of disease

   – Adequately treated carcinoma in situ without evidence of disease (eg, cervical cancer in situ)

12. History of leptomeningeal carcinomatosis

13. Brain metastases or spinal cord compression unless the patient is stable (asymptomatic, no evidence of new or emerging brain metastases) and off steroids for at least 14 days prior to start of study treatment. Following radiotherapy and/or surgery, patients with brain metastases must wait 4 weeks following the intervention and must confirm stability with imaging before randomization. Patients with suspected brain metastases at screening should have a CT/MRI of the brain prior to study entry.

14. Mean QT interval corrected for heart rate using Fridericia's formula (QTcF) ≥470 ms

15. History of active primary immunodeficiency

16. Active infection, including tuberculosis (clinical evaluation), hepatitis B, hepatitis C, or human immunodeficiency virus (HIV, positive HIV 1 or 2 antibodies). Active hepatitis B virus (HBV) is defined by a known positive HBV surface antigen (HBsAg) result. Patients with a past or resolved HBV infection (defined as the presence of hepatitis B core antibody and absence of HBsAg) are eligible. Patients positive for hepatitis C virus (HCV) antibody are eligible only if polymerase chain reaction is negative for HCV ribonucleic acid (RNA).

17. Current or prior use of immunosuppressive medication within 14 days before the first dose of MEDI4736 or tremelimumab. The following are exceptions to this criterion:

    - Intranasal, inhaled, topical steroids, or local steroid injections (eg, intra-articular injection).

    - Systemic corticosteroids at physiologic doses not to exceed 10 mg/day of prednisone or its equivalent

    - Steroids as premedication for hypersensitivity reactions (eg, CT scan premedication)

18. Receipt of live, attenuated vaccine within 30 days prior to the first dose of IP. Note: Patients, if enrolled, should not receive live vaccine during the study and up to 30 days after the last dose of IP.

19. Female patients who are pregnant or breast-feeding or male or female patients of reproductive potential who are not willing to employ effective birth control from screening to 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy.

20. Known allergy or hypersensitivity to IP or any excipient or to other humanized mAbs

Procedures for withdrawal of incorrectly enrolled patients are presented in Section 3.4.

3.3 Patient enrollment and randomization

Investigators should keep a record (ie, the patient screening log) of patients who entered screening.

At screening (Days -28 to -1), the Investigators or suitably trained delegate will:

1. Obtain signed informed consent before any study specific procedures are performed. (Informed consent of study procedures may be obtained prior to the 28-day screening window in order to permit tumor biopsy sample acquisition which must be analyzed prior to randomization.)

2. Obtain a unique 7-digit enrollment number (E-code), through the Interactive Voice Response System (IVRS)/(Interactive Web Response System) IWRS in the following format (ECCNNXXX: CC being the country code, NN being the center number, and XXX being the patient enrollment code at the center). This number is the patient's unique identifier and is used to identify the patient on the electronic case report forms (eCRFs).

3. Obtain tumor sample and send for PD-L1 expression (Obtaining the tumor biopsy sample should be given the highest priority and, as such, the sample may be obtained and sent for PD-L1 expression status evaluation prior to the 28-day screening window in order to permit analysis prior to randomization.)

   The sample should be sent only for the patient with known EGFR and ALK status. **If a patient has squamous histology or is known to have a tumor with a KRAS mutation, then EGFR and ALK testing is not required.** If EGFR and ALK status is unknown, then the tumor sample (archive or fresh, primary or metastatic) should be used firstly for (local or central) EGFR and ALK mutation testing in accordance to inclusion criterion 4.

4. Determine patient eligibility (see Sections 3.1 and 3.2)

5. Obtain signed informed consent for genetic research study (optional)

At randomization, once the patient is confirmed to be eligible, the Investigator or suitably trained delegate will:

1. Define the SoC treatment (based on the most appropriate option for the patient) that the patient would receive if randomized to the SoC group prior to randomization of

the patient. This must be completed for all patients. The information will be recorded in the IVRS/IWRS system.

Note, for all patients with non-squamous tumor histology who would be scheduled to receive pemetrexed if randomized to the SoC group, folic acid and vitamin B12 should commence prior to randomization for up to 7 days, in line with local practice. This is to ensure SoC treatment can begin on Day 1.

2. Obtain a unique randomization number via the IVRS/IWRS. Numbers will start at 001 and will be assigned strictly sequentially by IVRS/IWRS as patients are eligible for entry into the study. The system will randomize the eligible patient to 1 of the 3 treatment groups. PD-L1 expression status results must be received from the central laboratory by the IVRS/IWRS prior to randomization.

If the patient is ineligible and not randomized, the IVRS/IWRS should be contacted to terminate the patient in the system.

Patients will begin treatment on Day 1. Patients must not be treated unless all eligibility criteria have been met.

If a patient withdraws from participation in the study, then his or her enrollment/randomization code cannot be reused. Withdrawn patients will not be replaced.

3.4 Procedures for handling incorrectly enrolled patients

Patients who fail to meet the eligibility criteria should not, under any circumstances, be randomized or receive study medication. There can be no exceptions to this rule. Patients who are enrolled but found to not meet all the eligibility criteria must not be randomized, and must not be initiated on treatment and must be withdrawn from the study as a screen failure.

When a patient does not meet all the eligibility criteria but is randomized in error or incorrectly started on treatment, the Investigator should inform the Study Physician immediately, and the Study Physician and the Investigator should discuss whether to continue or discontinue the patient from treatment. The Study Physician must ensure that all decisions are appropriately documented.

3.5 Methods for assigning treatment groups

The actual treatment given to patients will be determined by the randomization scheme in the IVRS/IWRS. The randomization scheme will be produced by a computer software program that incorporates a standard procedure for generating randomization numbers.
One randomization list will be produced for each of the randomization strata. A blocked randomization will be generated, and all centers will use the same list in order to minimize any imbalance in the number of patients assigned to each treatment group.

Patients will be identified to the IVRS/IWRS per country regulations. Randomization codes will be assigned strictly sequentially, within each stratum, as patients become eligible for

randomization. The IVRS/IWRS will provide the kit identification number to be allocated to the patient at the randomization visit.

3.6 Methods for ensuring blinding

Not applicable; this study is not blinded.

3.7 Methods for unblinding

Not applicable; this study is not blinded.

3.8 Restrictions

The following restrictions apply while the patient is receiving study treatment and for the specified times before and after:

1. Female patients of child-bearing potential

   – Females of childbearing potential who are sexually active with a non-sterilized male partner must use at least 1 **highly** effective method of contraception (Table 1) from the time of screening and must agree to continue using such precautions for 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy. Male partners of a female patient must use a male condom plus spermicide (except in countries where spermicides are not approved) throughout this period. Cessation of birth control after this point should be discussed with a responsible physician. Not engaging in sexual activity for the total duration of the drug treatment and the drug washout period is an acceptable practice; however, periodic abstinence, the rhythm method, and the withdrawal method are not acceptable methods of birth control. Female patients should also refrain from breastfeeding throughout this period.

2. Male patients with a female partner of childbearing potential

   – Non-sterilized males who are sexually active with a female partner of childbearing potential must use a male condom plus spermicide (except in countries where spermicides are not approved) from screening through 180 days after receipt of the final dose of MEDI4736 + tremelimumab combination therapy or 90 days after receipt of the final dose of MEDI4736 monotherapy (see Table 1). Not engaging in sexual activity is an acceptable practice; however, abstinence, the rhythm method, and the withdrawal method are not acceptable methods of contraception. Male patients should refrain from sperm donation throughout this period.

   – It is strongly recommended for the female partner of a male patient to also use a highly effective method of contraception throughout this period.

Note - Females of childbearing potential are defined as those who are not surgically sterile (ie, bilateral tubal ligation, bilateral oophorectomy, or complete hysterectomy) or post-menopausal.

Women will be considered post-menopausal if they have been amenorrheic for 12 months without an alternative medical cause. The following age-specific requirements apply:

- Women <50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of exogenous hormonal treatments and if they have luteinizing hormone and follicle-stimulating hormone levels in the post-menopausal range for the institution or underwent surgical sterilization (bilateral oophorectomy or hysterectomy).

- Women ≥50 years of age would be considered post-menopausal if they have been amenorrheic for 12 months or more following cessation of all exogenous hormonal treatments, had radiation-induced menopause with last menses >1 year ago, had chemotherapy-induced menopause with last menses >1 year ago, or underwent surgical sterilization (bilateral oophorectomy, bilateral salpingectomy or hysterectomy).

Highly effective methods of contraception, defined as 1 that results in a low failure rate (ie, less than 1% per year) when used consistently and correctly are described in Table 1. Note that some contraception methods are not considered highly effective (eg, male or female condom with or without spermicide; female cap, diaphragm, or sponge with or without spermicide; non-copper containing intrauterine device; progestogen-only oral hormonal contraceptive pills where inhibition of ovulation is not the primary mode of action [excluding Cerazette/desogestrel, which is considered highly effective]; and triphasic combined oral contraceptive pills).

Patients in the SoC group: Follow the local prescribing information relating to contraception, the time limits for such precautions, and any additional restrictions for agents in the SoC group.

Table 1 Highly effective[a] methods of contraception (<1% failure rate)	
Barrier/Intrauterine methods	**Hormonal methods**
• Copper T intrauterine device • Levonorgesterel-releasing intrauterine system (eg, Mirena®)[b] •	• Etonogestrel implants: eg, Implanon or Norplan • Intravaginal device: eg, ethinylestradiol and etonogestrel • Medroxyprogesterone injection: eg, Depo-Provera • Normal and low-dose combined oral contraceptive pill • Norelgestromin/ethinylestradiol transdermal system • Cerazette (desogestrel)

[a] Highly effective (i.e. failure rate of <1% per year)
[b] This is also considered a hormonal method

3. All patients: Patients should not donate blood or blood components while participating in this study and for 90 days following the last dose of IP.

4. Restrictions relating to concomitant medications are described in Section 7.7

3.9 Discontinuation of investigational product

An individual patient will not receive any further IP (MEDI4736 + tremelimumab combination, MEDI4736 monotherapy, or SoC) if any of the following occur in the patient in question:

- Withdrawal of consent from further treatment with IP. The patient is, at any time, free to discontinue treatment, without prejudice to further treatment. A patient who discontinues treatment is normally expected to continue to participate in the study unless they specifically withdraw their consent to further participation in any study procedures and assessments (see Section 3.10.2).
- An AE that, in the opinion of the Investigator or AstraZeneca, contraindicates further dosing
- Pregnancy or intent to become pregnant
- Non-compliance with the study protocol that, in the opinion of the Investigator or AstraZeneca, warrants withdrawal from study treatment (eg, refusal to adhere to scheduled visits)
- Initiation of alternative anticancer therapy including another investigational agent

- Confirmed PD and Investigator determination that the patient is no longer benefiting from treatment with IP

3.9.1 Procedures for discontinuation of a patient from investigational product

At any time, patients are free to discontinue IP without prejudice to further treatment. A patient who decides to discontinue IP will always be asked about the reason(s) for discontinuation and the presence of any AE. If possible, they will be seen and assessed by an Investigator. AEs will be followed up (see Section 6). The Study Physician should be notified of any ongoing AE that may delay treatment or necessitate permanent discontinuation of treatment.

Patients who are permanently discontinued from further receipt of IP, regardless of the reason, will be identified as having permanently discontinued treatment. Patients who are permanently discontinued will enter follow-up (see Table 4). All patients will be followed for survival until the end of the study. Patients who decline to return to the site for evaluations should be contacted by telephone as an alternative.

Any patient who discontinues study treatment for reasons other than objective disease progression should have tumor assessments performed as scheduled in Table 4 until objective disease progression is documented or death occurs, unless consent is withdrawn.

Patients who have permanently discontinued from further receipt of IP will need to be discontinued from the IVRS/IWRS.

3.10 Criteria for withdrawal of the patient from the study

3.10.1 Screen failures

Screen failures are patients who do not fulfill the eligibility criteria for the study, and therefore must not be randomized. These patients should have the reason for study withdrawal recorded as “eligibility criteria not fulfilled” (ie, patient does not meet the required inclusion/exclusion criteria). This reason for study withdrawal is only valid for screen failures (ie, not randomized patients). Patients can be rescreened a single time, but they cannot be re-randomized.

3.10.2 Withdrawal of the informed consent

Patients are free to withdraw from the study at any time (IP and assessments) without prejudice to further treatment.

Patients who withdraw consent for further participation in the study will not receive any further IP or further study observation, with the exception of follow-up for survival, which will continue until the end of the study unless the patient has expressly withdrawn their consent to survival follow-up. Note that the patient may be offered additional tests or tapering of treatment to withdraw safely.

A patient who withdraws consent will always be asked about the reason(s) for withdrawal and the presence of any AE. The Investigator will follow up AEs outside of the clinical study. The patient will return electronic PRO devices, if applicable.

Patients will be considered lost to follow-up only if no contact has been established by the time the study is completed (see Section 9.3), such that there is insufficient information to determine the patient's status at that time. Patients who refuse to continue participation in the study, including telephone contact, should be documented as "withdrawal of consent" rather than "lost to follow-up." Investigators should document attempts to re-establish contact with missing patients throughout the study period. If contact with a missing patient is re-established, the patient should not be considered lost to follow-up and evaluations should resume according to the protocol.

3.10.3 Survival status for withdrawn consent and lost to follow up patients

In order to support key end points of PFS and OS analyses, the survival status of all patients in the full analysis and the safety analysis sets should be re-checked, this includes those patients who withdrew consent or are classified as "lost to follow up."

- Lost to Follow up – site personnel should check hospital records, the patients' current physician, and a publicly available death registry (if available) to obtain a current survival status. (The SURVIVE module will be updated.)
- In the event that the patient has actively withdrawn consent to the processing of their personal data, the survival status of the patient can be obtained by site personnel from publicly available death registries (if available) where it is possible to do so under applicable local laws to obtain a current survival status. (The SURVIVE module will be updated.)

3.11 Discontinuation of the study

The study may be stopped if, in the judgment of AstraZeneca, study patients are placed at undue risk because of clinically significant findings that meet any of the following criteria:

- Meet individual stopping criteria or are otherwise considered significant
- Are assessed as causally related to study drug
- Are not considered to be consistent with continuation of the study

In addition, the study may be stopped based on the findings of the interim safety analysis conducted by the Independent Data Monitoring Committee (IDMC) (see Section 6.8).

Regardless of the reason for termination, all data available for the patients at the time of discontinuation of follow-up must be recorded in the eCRFs. All reasons for discontinuation of treatment must be documented.

In terminating the study, AstraZeneca will ensure that adequate consideration is given to the protection of the patients' interests. If this study is discontinued, all other studies involving MEDI4736 or tremelimumab will remain open to enrollment and screening, if deemed appropriate by AstraZeneca.

4. STUDY PLAN AND TIMING OF PROCEDURES

The procedures for the screening and the treatment periods in this study are presented for the MEDI4736 + tremelimumab combination therapy group and MEDI4736 monotherapy group in Table 2and for the SoC therapy group in Table 3. The procedures for the follow-up period are presented in Table 4.

Table 2 Schedule of assessments for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy treatment and retreatment periods

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13[a] etc	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48 etc**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337 etc**	
Window (days)	**NA**	**+3[b]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Informed consent											
Informed consent: study procedures	X[c]										4.1, 10.4
Consent: genetic sample and analysis (optional)	X										3.3
Study procedures											
Physical exam (full)	X										5.2.2
Targeted physical exam (based on symptoms)		X	X	X	X	X	X	X	X	X	5.2.2
Vital signs[d]	X	X	X	X	X	X	X	X	X	X	5.2.4
ECG[e]	X	As clinically indicated									5.2.3
Concomitant medications	<---------- --->										7.7
Demography, including baseline characteristics and tobacco use	X										4.1
Eligibility criteria	X										3.1, 3.2
Laboratory assessments[f]											
Clinical chemistry	X	X	X	X	X	X	X	X	X	X	Table 5
Hematology	X	X	X	X	X	X	X	X	X	X	Table 6
APTT and INR	X	As clinically indicated									Table 6
TSH, free T_3, and free T_4[g]	X	X	X	X	X	X	X	X	X	X	5.2.1
Urinalysis	X	As clinically indicated									Table 7
Hepatitis B and C and HIV	X										5.2.1

Table 2 Schedule of assessments for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy treatment and retreatment periods

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13[a] etc	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48 etc**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337 etc**	
Window (days)	**NA**	**+3[b]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Pregnancy test[h]	X	X	X	X	X	X	X	X	X	X	5.2.1
Pharmacokinetics											
MEDI4736 PK sample (serum)		X[i]			X[i]			X[i]			5.4.1
Tremelimumab PK sample (serum; combination therapy group only)		X[i]			X[i]						5.4.1
Monitoring											
WHO/ECOG performance status	X	X	X	X	X	X	X	X	X	X	5.3.3
AE/SAE assessment	<---------- --->										6.3.1
Drug accountability		X	X	X	X	X	X	X	X	X	7.6
Pre-randomization medication											
Folic acid[j]	X	Discontinue as randomized to monotherapy or combination therapy									3.3
IM Vitamin B12[j]	X	Discontinue as randomized to monotherapy or combination therapy									3.3
IP administration											
Combination therapy group											
MEDI4736 (combination therapy)[kl]		X	X	X	X	X	X	X	X	X	7.2.1
Tremelimumab[i,j]		X	X	X	X						7.2.1
Monotherapy group											
MEDI4736 (monotherapy)[l]		X	X	X	X	X	X	X	X	X	7.2.1
PRO assessments[m]											

Table 2 Schedule of assessments for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy treatment and retreatment periods

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13[a] etc	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48 etc**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337 etc**	
Window (days)	**NA**	**+3[b]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
EORTC QLQ-C30, EQ-5D-5L	X	Every 4 weeks for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)									5.3.1.1, 5.3.1.5
EORTC QLQ-LC13, PRO-CTCAE[n]	X	Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)									5.3.1.2, 5.3.1.3
PGIC			X	X	X			X		X	5.3.1.4
Other laboratory assessments and assays											
Immunogenicity assessment (ADA sampling to identify ADA responses in patient circulation)		X			X			X[o]			5.4.2
sPD-L1 (serum)		X			X			X			5.5.2
Circulating soluble factors (plasma)		X	X		X						5.5.2
Tumor biopsy (newly acquired or archived <3 months old)	X[c]										5.5.1
Archival tumor sample ≥3 months old, if available	X										5.5.1
EGFR and ALK test	X[p]										3.3
Whole blood for SNP genotyping (pre-dose)		X									5.5.2
Whole blood for gene expression (PaxGene-RNA tubes)		X	X								5.5.2
Myeloid-derived suppressor cells (Cyto-Chex tube)		X									5.5.2
PBMCs		X	X								5.5.2

Table 2 Schedule of assessments for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy treatment and retreatment periods

	Screening	C1	C2	C3	C4	C5	C6	C7	C8 to C12	C13[a] etc	For details see Section
Week	**-4 to -1**	**0**	**4**	**8**	**12**	**16**	**20**	**24**	**28, 32, 36, 40, 44**	**48 etc**	
Day	**-28 to -1**	**1**	**29**	**57**	**85**	**113**	**141**	**169**	**197, 225, 253, 281, 309**	**337 etc**	
Window (days)	**NA**	**+3[b]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Tumor evaluation (CT or MRI) (RECIST 1.1)[q,r]	X	Every 6 weeks ± 1 week for the first 48 weeks relative to the date of randomization, and then every 8 weeks ± 1 week thereafter									5.1
PGx sample (optional DNA element for long-term storage/future use)	X										5.6
Health economics measurements											
Hospital resource use module (HOSPAD)[s]		To be completed at each hospitalization and unscheduled visit by site staff									8.5.7

a. Patients who continue on treatment will be assessed in the same manner
b. Every effort should be made to minimize the time between randomization and starting treatment. (ie. on the same day after randomization)
c. Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization. The collection of tumor biopsies at the time of progression prior to retreatment is mandated; the Investigator must consult with the Study Physician if such sampling is not feasible. The collection of additional biopsies upon progression is strongly encouraged.
d. Body weight is recorded along with vital signs.
e. Any clinically significant abnormalities detected require a confirmatory ECG.
f. If screening laboratory assessments are performed within 3 days prior to Day 1 (first infusion day), they do not need to be repeated at Day 1.
g. Free T_3 and free T_4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
h. For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
i. Pre-dose same day as infusion and within 1 hour of end of infusion.
j. To be administered in line with local practice for patients with non-squamous tumors who will receive pemetrexed if randomized to SoC group.
k. During the combination portion of treatment, tremelimumab will be administered first; the MEDI4736 infusion will start approximately 1 hour after the end of the tremelimumab infusion. If there are no clinically significant infusion reactions with the first cycle, then for all other cycles, the MEDI4736 can be given immediately after the tremelimumab infusion has finished.
l. Results for urea and electrolytes, full blood count, and liver function tests must be available before commencing an infusion (within 3 days).
m. In instances where the PRO assessments coincide with a clinic visit, the research nurse or study coordinator should verify that the questionnaire was completed before the visit. If not, the patient should complete the questionnaire prior to any other study procedures and before discussion of disease progression to avoid biasing the patient's responses to the questions.
n. PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.
o. MEDI4736 sample only at this time point.

p. For patients with unknown status of ALK and/or EGFR NSCLC. (If patients have squamous histology or are known to have a tumor with a KRAS mutation, then EGFR and ALK testing is not required.)
q. RECIST 1.1 assessments will be performed on CT/MRI images of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients at baseline and follow-up. Baseline assessments should be performed no more than 28 days before the date of randomization and, ideally, should be performed as close as possible to and prior to the start of study treatment. The radiological progression confirmatory scans should be performed no less than 4 weeks after the prior assessment of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their next scheduled visit.
r. Patients with confirmed PD can continue to receive MEDI4736 + tremelimumab or MEDI4736 at the discretion of the Investigator
s. HOSPAD module should be completed by site staff whenever the patient has attended or been admitted in to the hospital and at unscheduled visits. A reminder will be provided at each clinic visit.

Note: All assessments on treatment days are to be performed prior to infusion, unless otherwise indicated.
Note: For "retreatment", the same assessments should be done as in the first treatment period, with the exception of the PK, ADA, , SNP genotyping, and MDSCs assessments, which do not need to be collected a second time.

ADA Anti-drug antibody; AE Adverse event; ALK Anaplastic lymphoma kinase; APTT activated partial thromboplastin time; C Cycle; CT Computed tomography; ECG Electrocardiogram; ECOG Eastern Cooperative Oncology Group; EGFR Epidermal growth factor receptor; EORTC QLQ-C30 30-item core Eastern Cooperative Oncology Group Quality of Life Questionnaire; EORTC QLQ-LC13 13-item lung cancer EORTC QLQ; EQ-5D-5L EuroQol 5-Dimension, 5-Level health state utility index; HIV Human immunodeficiency virus; IM intramuscular; INR international normalized ratio; IP Investigational product; MDSC Myeloid-derived suppressor cell; MRI Magnetic resonance imaging; NSCLC Non-small-cell lung cancer; PBMC Peripheral blood mononuclear cell; PD Progressive disease; PGIC Patients' Global Impression of Change; PGx Pharmacogenetic research; PK Pharmacokinetic(s); PRO-CTCAE Patient-reported outcomes version of the Common Terminology Criteria for Adverse Events; RNA Ribonucleic acid; SAE Serious adverse event; SNP Single nucleotide polymorphism; sPD-L1 Soluble programmed cell death ligand 1; TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine; WHO World Health Organization.

Table 3 Schedule of assessments for Standard of Care therapy treatment period

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**+3[a]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Informed consent													
Informed consent: study procedures	X[b]												4.1, 10.4
Consent: genetic sample and analysis (optional)	X												3.3
Study procedures													
Physical exam (full)	X												5.2.2
Targeted physical exam (based on symptoms)		X	X	X	X	X	X	X	X	X	X	X	5.2.2
Vital signs[c]	X	X	X	X	X	X	X	X	X	X	X	X	5.2.4
ECG[d]	X	As clinically indicated											5.2.3
Concomitant medications	<---------- ------>												7.7
Demography, including baseline characteristics and tobacco use	X												4.1
Eligibility criteria	X												3.1, 3.2
Laboratory assessments													
Clinical chemistry[e,f]	X	X	X	X	X	X	X	X	X	X	X	X	Table 5
Hematology[e,f]	X	X	X	X	X	X	X	X	X	X	X	X	Table 6
APTT and INR	X	As clinically indicated											Table 6

Table 3 Schedule of assessments for Standard of Care therapy treatment period

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**+3[a]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
TSH, free T_3, and free T_4[g]	X					X				X			5.2.1
Urinalysis	X	As clinically indicated											Table 7
Hepatitis B and C and HIV	X												5.2.1
Pregnancy test[h]	X	As clinically indicated											5.2.1
Monitoring													
WHO/ECOG performance status	X	X	X	X	X	X	X	X	X	X	X	X	5.3.3
AE/SAE assessment	<----------												6.3.1
	---------->												
Pre-randomization medication													
Folic acid[i]	X	Continue in line with local practice											3.3
IM Vitamin B12[i]	X	Continue in line with local practice											3.3
SoC administration													
Platinum-based chemotherapy		X	Cycle every 3 weeks										7.2.1
PRO assessments[j]													
EORTC QLQ-C30, EQ-5D-5L	X	Every 4 weeks relative for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)											5.3.1.1, 5.3.1.5
EORTC QLQ-LC13, PRO-CTCAE[k]	X	Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)											5.3.1.2, 5.3.1.3
PGIC		Weeks 4, 8, 12, 24, and 48											5.3.1.4

Table 3 Schedule of assessments for Standard of Care therapy treatment period

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**+3[a]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Other laboratory assessments and assays													
sPD-L1 (serum)		X				X				X			5.5.2
Circulating soluble factors		X		X		X							5.5.2
Tumor biopsy (newly acquired or archived <3 months old)	X[b]												5.5.1
Archival tumor sample ≥3 months old, if available	X												5.5.1
EGFR and ALK test	X[l]												3.3
Whole blood for SNP genotyping		X											5.5.2
Whole blood gene expression (PaxGene-RNA tubes)		X		X									5.5.2
Myeloid-derived suppressor cells (Cyto-Chex tube)		X											5.5.2
PBMCs		X		X									5.5.2
Tumor evaluation (CT or MRI) (RECIST 1.1)[m]	X	Every 6 weeks ± 1 week for the first 48 weeks relative to the date of randomization, and then every 8 weeks ± 1 week thereafter											5.1
PGx sample (optional DNA element for long-term storage/future use)	X												5.6

Table 3 Schedule of assessments for Standard of Care therapy treatment period

	Screening	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10 to C17	C18, C19, etc	For details see Section
Week	**-4 to -1**	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27, 30, 33, 36, 39, 42, 45, 48**	**51, 54, etc**	
Day	**-28 to -1**	**1**	**22**	**43**	**64**	**85**	**106**	**127**	**148**	**169**	**190, 211, 232, 253, 274, 295, 316, 337**	**358, 379, etc**	
Window (days)	**NA**	**+3[a]**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	
Health economics measurements													
Hospital resource use module (HOSPAD)[n]		To be completed at each hospitalization and unscheduled visit by site staff											8.5.7

a. Every effort should be made to minimize the time between randomization and starting treatment. (ie, on the same day after randomization)
b. Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization
c. Before every infusion or administration and as clinically indicated.
d. Any clinically significant abnormalities detected require a confirmatory ECG.
e. To be collected every 3 weeks prior to the start of infusion and as clinically indicated.
f. If screening laboratory assessments are performed within 3 days prior to Day 1 (first infusion day), they do not need to be repeated at Day 1.
g. Free T3 and free T4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
h. For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
i. To be administered in line with local practice for patients with non-squamous tumors who will receive pemetrexed if randomized to SoC group.
j. In instances where the PRO assessments coincide with a clinic visit, the research nurse or study coordinator should verify that the questionnaire was completed before the visit. If not, the patient should complete the questionnaire prior to any other study procedures and before discussion of disease progression to avoid biasing the patient's responses to the questions.
k. PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.
l. For patients with unknown status of ALK and/or EGFR NSCLC. (If patients have squamous histology or is known to have a tumor with a KRAS mutation, then EGFR and ALK testing is not required.)
m. RECIST 1.1 assessments will be performed on CT/MRI images of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients at baseline and follow-up. Baseline assessments should be performed no more than 28 days before the date of randomization and, ideally, should be performed as close as possible to and prior to the start of study treatment. The radiological progression confirmatory scans should be performed no less than 4 weeks after the prior assessment of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their next scheduled visit.

HOSPAD module should be completed by site staff whenever the patient has attended or been admitted in to the hospital and at unscheduled visits. A reminder will be provided at each clinic visit.

Note: All assessments on treatment days are to be performed prior to infusion or administration, unless otherwise indicated.

ADA Anti-drug antibody; AE Adverse event; ALK Anaplastic lymphoma kinase; APTT activated partial thromboplastin time; C Cycle; CT Computed tomography; ECG Electrocardiogram; ECOG Eastern Cooperative Oncology Group; EGFR Epidermal growth factor receptor; EORTC QLQ-C30 30-item core Eastern Cooperative Oncology Group Quality of Life Questionnaire; EORTC QLQ-LC13 13-item lung cancer EORTC QLQ; EQ-5D-5L EuroQol 5-Dimension, 5-Level health state utility index; HIV Human immunodeficiency virus; IM intramuscular; INR international normalized ratio; IP Investigational product; MDSC Myeloid-derived suppressor cell; MRI Magnetic resonance imaging; NSCLC Non-small-cell lung cancer; PBMC Peripheral blood mononuclear cell; PD Progressive disease; PGIC Patients' Global Impression of Change; PGx Pharmacogenetic research; PK Pharmacokinetic(s); PRO-CTCAE Patient-reported outcomes version of the Common Terminology Criteria for Adverse Events; RNA Ribonucleic acid; SAE Serious adverse event; SNP Single nucleotide polymorphism; sPD-L1 Soluble programmed cell death ligand 1; TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine; WHO World Health Organization.

Table 4 Schedule of assessments for patients who have completed/discontinued treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or Standard of Care therapy

	Time since last dose of IP							
	Day (±3)	**Months (±1 week)**						**12 months and every 2 months (±2 weeks)**
Evaluation	**30**	**2**	**3**	**4**	**6**	**8**	**10**	
Physical examination (full)[a]	X							
Vital signs (temperature, respiratory rate, blood pressure, and pulse)	X							
Weight	X							
Pregnancy test[b]	X	As clinically indicated						
AE/SAE assessment[c]	X	X	X					
Concomitant medications	X	X	X					
WHO/ECOG performance status	At timepoints consistent with tumor assessments; at 30, 60, and 90 days; and then at initiation of subsequent anticancer therapy[d]							
Subsequent anticancer therapy[e]; and second progression assessment[f,g]	<---------- ----->							
Survival status[h]		X	X	X	X	X	X	X
Hematology	X	X	X					
Clinical chemistry	X	X	X					
TSH, free T_3, and free T_4[i]	X							
Pharmacokinetic assessment[j]			X					
Immunogenicity assessment (ADA sampling) to identify ADA responses in patient circulation[j]			X		X			
sPD-L1 concentration (to assess target engagement)[j]			X					
EORTC QLQ-C30[k], EQ-5D-5L[k]	Every 4 weeks relative for the first 8 weeks relative to the date of randomization, then every 8 weeks thereafter until second progression/death (whichever comes first)							

Table 4 Schedule of assessments for patients who have completed/discontinued treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or Standard of Care therapy

	Time since last dose of IP							
	Day (±3)	Months (±1 week)						12 months and every 2 months (±2 weeks)
Evaluation	**30**	**2**	**3**	**4**	**6**	**8**	**10**	
EORTC QLQ-LC13[k], PRO-CTCAE[k,l]	Every 2 weeks for the first 8 weeks relative to the date of randomization, then every 4 weeks thereafter until second progression/death (whichever comes first)							
Hospital resource use module (HOSPAD)	X							
Tumor assessment (CT or MRI)[m]	Every 6 weeks ± 1 week for the first 48 weeks (relative to the date of randomization), and then every 8 weeks ± 1 week thereafter until confirmed objective disease progression/death (whichever comes first). Additional scans to be completed per standard practice post progression.							

a. Physical exams are described in Section 5.2.2.
b. For women of childbearing potential only. A urine or serum pregnancy test is acceptable.
c. The AE/SAE follow-up for SoC is only 30 days
d. WHO/ECOG performance status should also be collected at other site visits that the patient attends, if appropriate site staff are available to collect such information. In addition, WHO performance status should be provided when information on subsequent anticancer therapy is provided, where possible.
e. Details of any treatment for NSCLC (including surgery) post the last dose of study treatment must be recorded in the eCRF.
f. PFS2 assessment will be performed by the Investigator and defined according to local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression, or death.
g. For patients who discontinue their assigned IP following confirmed progression, available readings of CT/MRI from local practice will be collected from the patients' medical charts while information on subsequent anticancer treatment and/or PFS2 is collected.
h. Patients may be contacted in the week following data cut-offs to confirm survival status. Details of any treatment for NSCLC (including surgery) post the last dose of study treatment must be recorded in the eCRF.
i. Free T3 and free T4 will only be measured if TSH is abnormal or if there is clinical suspicion of an AE related to the endocrine system.
j. For patients in the MEDI4736 + tremelimumab or MEDI4736 monotherapy groups only. The 3 month follow-up collections for MEDI4736 and Tremelimumab PK and ADA are relative to respective last dose and the 6-month follow-up collections for MEDI4736 and Trememlimumab ADA are relative to respective last dose.
k. Patients will complete PROs using handheld devices at home.
l. PRO-CTCAE will only be administered in those countries where a linguistically validated version exists.
m. Only for patients yet to progress, RECIST 1.1 assessments will be performed on CT/MRI images of the chest and abdomen (including liver and adrenal glands). Additional anatomy should be imaged based on signs and symptoms of individual patients. The radiological progression confirmatory scans should preferably be performed no less than 4 weeks after the prior assessment of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). If an unscheduled assessment was performed and the patient has not progressed, every attempt should be made to perform the subsequent assessments at their scheduled visits (relative to the date of randomization). All confirmatory scans should be recorded on the database.

ADA Anti-drug antibody; AE Adverse event; CT Computed tomography; ECOG Eastern Cooperative Oncology Group; eCRF electronic case repot form; EORTC QLQ-C30 Eastern Cooperative Oncology Group Quality of Life Questionnaire; EORTC QLQ-LC13 EORTC QLQ Lung Cancer 13; EQ-5D-5L EuroQol 5-Dimension, 5-Level health state utility index; IP Investigational product; MRI Magnetic resonance imaging; NSCLC Non-small-cell lung cancer; PD Progressive disease; PFS2 Time to second progression; PRO-CTCAE Patient-reported outcomes version of the Common Terminology Criteria for Adverse Events; RECIST Response Evaluation Criteria In Solid Tumors; SAE Serious adverse event; sPD-L1 Soluble programmed cell death ligand 1TSH Thyroid-stimulating hormone; T_3 Triiodothyronine; T_4 Thyroxine; WHO World Health Organization.

4.1 Enrollment/screening period

All screening and enrollment procedures will be performed according to the assessment schedule in Table 2 and Table 3. Demographic data and other characteristics will be recorded including date of birth or age, gender, smoking history, and race/ethnicity, according to local regulations. A standard medical and surgical history will be obtained.

Written informed consent and any locally required privacy act document authorization must be obtained prior to performing any protocol-specific procedures, including screening evaluations. All patients will be required to provide consent to supply a sample of their tumor (archived or newly acquired biopsy) for entry into this study. This consent is included in the main patient informed consent form (ICF).

Screening evaluations may be performed over more than 1 visit.

The timing of vital sign assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 2 and Table 3.

4.2 Treatment period

All procedures to be conducted during the treatment period will be performed according to the assessment schedule (see Table 2 and Table 3).

Whenever vital signs and blood draws are scheduled for the same nominal time, the assessments should occur in the following order: vital signs and then blood draws. The timing of the vital signs assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 2 and Table 3.

4.3 Follow-up period

All procedures to be conducted during the follow-up period will be performed according to the assessment schedule (see Table 4).

Whenever vital signs, ECGs, and blood draws are scheduled for the same nominal time, the assessments should occur in the following order: ECG, vital signs, and then blood draws. The timing of the first 2 assessments should be such that it allows the blood draw (eg, PK blood sample) to occur at the timepoints indicated in Table 4.

5. STUDY ASSESSMENTS

A Web-Based Data Capture (WBDC) system will be used for data collection and query handling. The Investigator will ensure that data are recorded on the eCRFs as specified in this study protocol and in accordance with the instructions provided.

The Investigator ensures the accuracy, completeness, and timeliness of the data recorded and the provision of answers to data queries according to the clinical study agreement (CSA). The Investigator will sign the completed eCRFs. A copy of the completed eCRFs will be archived at the study site.

5.1 Efficacy assessments

RECIST 1.1 criteria will be used to assess patient response to treatment by determining PFS, ORR, DoR, and APF12 using BICR assessments (primary), and using Investigator assessments for sensitivity analysis. The RECIST 1.1 guidelines for measurable, non-measurable, target, and non-target lesions and the objective tumor response criteria (CR, PR, SD, or PD) are presented in Appendix E. PFS2 defined by local standard clinical practice, and OS will also be evaluated.

The methods of assessment of tumor burden used at baseline are CT and MRI scans of the chest and abdomen (including liver and adrenal glands). Any other areas of disease involvement should be additionally imaged based on the signs and symptoms of individual patients.

The baseline assessment should be performed no more than 28 days before randomization and ideally as close as possible to the start of the assigned IP. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; Table 2, Table 3 and Table 4) then every 8 weeks (q8w) thereafter, until confirmed objective disease progression per RECIST 1.1 (irrespective of the reason for stopping treatment or subsequent therapy). 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.

For patients who discontinue IP (including SoC) due to toxicity in the absence of confirmed objective progression, objective tumor assessments per the scheduled assessments should be continued every 6 weeks for 48 weeks (relative to randomization) and then every 8 weeks until confirmed objective disease progression.

A confirmatory scan is required for all patients following the initial demonstration of PD. The confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD and preferably at the next scheduled visit in the absence of clinically significant deterioration. Treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC may continue between the initial assessment of progression and confirmation of progression. Progression would be considered confirmed per RECIST 1.1 criteria available in Appendix E using Investigator assessments.

If a patient discontinues treatment (and/or receives a subsequent anticancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

Categorization of objective tumor response assessment will be based on the RECIST 1.1 criteria of response: CR, PR, SD, and PD. Target lesion progression will be calculated in comparison to when the tumor burden was at a minimum (ie, smallest sum of diameters previously recorded on study). In the absence of progression, tumor response (CR or PR) and SD will be calculated in comparison to the baseline tumor measurements obtained before starting treatment.

Following confirmed progression, patients should continue to be followed up for survival every 1 to 2 months as outlined in the study plan (Table 4). Subsequent anticancer therapy information will be collected at the timepoints indicated in Table 4.

Patients in the MEDI4736 + tremelimumab who will receive retreatment must have a baseline tumor assessment within 28 days of restarting treatment and additional scans every 6 weeks for the first 48 weeks relative to the date of randomization, and then q8w thereafter until disease progression. All assessments in Table 2will be followed for patients who receive retreatment.

It is important to follow the assessment schedule as closely as possible. Please refer to the study plans (Table 2 and Table 3 [screening and the treatment period] and Table 4 for follow-up of patients who have completed or discontinued IP treatment]) and Appendix E.

5.1.1 Central reading of scans

A BICR of all radiological scans in accordance with RECIST 1.1 will be performed for the evaluation of the co-primary PFS endpoint (and all secondary endpoints determined from the tumor assessments).

All images will be collected centrally. Guidelines for imaging collection and storage will be provided in a separate document. Results of these independent reviews will not be communicated to investigators, and the management of patients will be based solely upon the results of the RECIST1.1 assessment conducted by the Investigator.

5.1.2 Survival assessments

Assessments for survival must be made every 2 months following treatment discontinuation. Survival information may be obtained via telephone contact with the patient, patient's family, or by contact with the patient's current physician. The details of first and subsequent therapies for cancer, after discontinuation of treatment, will be collected.

In addition, patients on treatment or in survival follow-up will be contacted following the data cut-off for the primary analysis of PFS and for each interim and final analysis of OS to provide complete survival data. These contacts should generally occur within 7 days of the data cut off.

5.2 Safety assessments

5.2.1 Laboratory safety assessments

Blood samples for determination of clinical chemistry and hematology will be taken at the times indicated in the assessment schedules and as clinically indicated (see Table 2, Table 3 and Table 4). Urine samples for analysis will be taken at screening.

Clinical laboratory safety tests, including serum pregnancy tests, will be performed in a licensed clinical laboratory according to local standard procedures. Sample tubes and sample sizes may vary depending on the laboratory method used and routine practice at the site. Urine pregnancy tests may be performed at the site using a licensed test (urine or serum pregnancy test). Abnormal clinically significant laboratory results should be repeated as soon as possible (preferably within 24 to 48 hours).

Additional safety samples may be collected if clinically indicated at the discretion of the Investigator. The date, time of collection, and results (values, units, and reference ranges) will be recorded on the appropriate eCRF.

The laboratory variables to be measured are presented in Table 5 (clinical chemistry), Table 6 (hematology), and Table 7 (urinalysis).

Pregnancy tests on either urine (human chorionic gonadotropin [hCG]) or blood (serum β-hCG) samples will be performed for pre-menopausal women of childbearing potential at screening and subsequent visits as specified in the assessment schedule (see Table 2, Table 3 and Table 4). Tests will be performed by the hospital's local laboratory. If results are positive, the patient must not start or continue treatment. In the event of a suspected pregnancy during the study, the test should be repeated.

Other safety tests to be performed at screening include assessment for HBV surface antigen, HCV antibodies, HIV antibodies, thyroid-stimulating hormone, free triiodothyronine (T_3), and free thyroxine (T_4).

Table 5 Clinical chemistry (serum or plasma)

Albumin	Glucose
Alkaline phosphatase	Lactate dehydrogenase
Alanine aminotransferase	Lipase[a]
Amylase[a]	Magnesium
Aspartate aminotransferase	Potassium
Bicarbonate	Sodium
Calcium	Total bilirubin[b]
Chloride	Total protein
Creatinine	Urea or blood urea nitrogen, depending on local practice
Gamma glutamyltransferase[c]	Uric acid

[a] In the event that amylase and lipase analyses cannot performed, 1 or the other will be performed in line with local practice.
[b] If total bilirubin is ≥2 × ULN (and evidence of Gilbert's syndrome) then fractionate into direct and indirect bilirubin.
[c] At screening and as clinically indicated.

Table 6 Hematology

Basophils	Mean corpuscular volume
Eosinophils	Monocytes
Hematocrit	Neutrophils
Hemoglobin	Platelet count
Lymphocytes	Red blood cell count
Mean corpuscular hemoglobin	Total white cell count
Mean corpuscular hemoglobin concentration	Activated partial thromboplastin time and international normalized ratio[a]

[a] Activated partial thromboplastin time and international normalized ratio are to be assessed at screening and as clinically indicated.

Table 7 Urinalysis

Bilirubin	Ketones
Blood	pH
Color and appearance	Protein
Glucose	Specific gravity

Note: Microscopy should be used as appropriate to investigate white blood cells and use the high power field for red blood cells.

If a patient shows an AST or ALT ≥3 × ULN together with total bilirubin ≥2 × ULN, refer to Appendix D for further instructions. These cases should be reported as SAEs if, after evaluation, they meet the criteria for a Hy's law case or if any of the individual liver test parameters fulfill any of the SAE criteria. All patients with an elevated AST, ALT, or

bilirubin value (the latter at ≥1.5 × ULN) at the time of the last dose of study treatment should have a further liver chemistry profile (AST, ALT, bilirubin, and alkaline phosphatase) performed 30 days (±3 days) after permanent discontinuation of study treatment.

Any clinically significant abnormal laboratory values should be repeated as clinically indicated and recorded on the eCRF. Situations in which laboratory safety results should be reported as AEs are described in Section 97

All patients with Grade 3 or 4 laboratory values at the time of completion or discontinuation from study treatment must have further tests performed until the laboratory values have returned to Grade 1 or 2, unless these values are not likely to improve because of the underlying disease.

5.2.2 Physical examination

Physical examinations will be performed according to the assessment schedules (see Table 2, Table 3 and Table 4). Full physical examinations will include assessments of the head, eyes, ears, nose, and throat and the respiratory, cardiovascular, GI, urogenital, musculoskeletal, neurological, dermatological, hematologic/lymphatic, and endocrine systems. Height will be measured at screening only. Targeted physical examinations are to be utilized by the Investigator on the basis of clinical observations and symptomatology. Situations in which physical examination results should be reported as AEs are described in Section 6.1.4.

5.2.3 Electrocardiograms

Resting 12-lead electrocardiograms (ECGs) will be recorded at screening and as clinically indicated throughout the study (see Table 2 and Table 3). ECGs should be obtained after the patient has been in a supine position for 5 minutes and recorded while the patient remains in that position.

At screening, a single ECG will be obtained on which QTcF must be <470 ms.

In case of clinically significant ECG abnormalities, including a QTcF value >470 ms, 2 additional 12-lead ECGs should be obtained over a brief period (eg, 30 minutes) to confirm the finding.

Situations in which ECG results should be reported as AEs are described in Section 6.1.4.

5.2.4 Vital signs

Vital signs (blood pressure [BP], pulse, temperature, and respiration rate) will be evaluated according to the assessment schedules (see Table 2, Table 3 and Table 4). Body weight is also recorded along with vital signs.

Supine or semi-supine BP will be measured using a semi-automatic BP recording device with an appropriate cuff size, after the patient has rested for at least 5 minutes. BP and pulse will

be collected from patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups at the following times (based on a 60-minute infusion):

For the first infusion, patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups will be monitored as follows:

- Prior to the beginning of the infusion (measured once from approximately 30 minutes before up to 0 minutes [ie, the beginning of the infusion])
- Approximately 30 minutes during the infusion (**halfway** through infusion)
- At the end of the infusion (approximately 60 minutes)
- A 1-hour observation period is recommended after the first infusion of MEDI4736 and tremelimumab. If no clinically significant infusion reactions are observed during or after the first cycle, subsequent infusion observation periods can be at the Investigator's discretion (suggested 30 minutes after each MEDI4736 and tremelimumab infusion).

On subsequent infusion days, patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups will be monitored at the start of the infusion and then per institution standard and as clinically indicated.

Patients in the SoC group will be monitored before every infusion or administration and as clinically indicated.

If the infusion takes longer than 60 minutes, then BP and pulse measurements should follow the principles as described above or be taken more frequently if clinically indicated. The date and time of collection and measurement will be recorded on the appropriate eCRF. Additional monitoring with assessment of vital signs is at the discretion of the Investigator per standard clinical practice or as clinically indicated.

Situations in which vital signs results should be reported as AEs are described in Section 6.1.4.

5.3 Other assessments

5.3.1 Patient-reported outcomes

"PRO" is an umbrella term referring to all outcomes and symptoms that are directly reported by the patient. PROs have become a significant endpoint when evaluating effectiveness of treatments in clinical trials. The following PROs will be administered in this study: EORTC QLQ-C30 (core questionnaire), QLQ-LC13 (lung cancer module), PRO-CTCAE, PGIC, and the EuroQol 5-Dimension, 5-Level health state utility index (EQ-5D-5L) (see Appendix F).

The PRO instruments will be completed by the patients using a handheld ePRO device. All assessments should be completed without assistance from anyone according to the assessment

schedules (see Table 2, Table 3 and Table 4). It takes approximately 30 to 45 minutes for patients to complete the questionnaires; therefore, the burden to the patient is moderate.

5.3.1.1 EORTC QLQ-C30

The EORTC QLQ-C30 v3 questionnaire is included for the purpose of assessing HRQoL and is a well-established measure of HRQoL/health status, and commonly used as an endpoint in cancer clinical trials. It assesses HRQoL/health status through 9 multi-item scales: 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea and vomiting), and a global health and QoL scale. Six single-item symptom measures are also included: dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties (see Appendix F). For each of the 15 domains, final scores are transformed such that they range from 0 to 100, where higher scores indicate greater functioning, greater HRQoL, or greater level of symptoms (Aaronson et al 1993).

5.3.1.2 EORTC QLQ-LC13

For patients with NSCLC, a disease-specific 13-item self-administered questionnaire for lung cancer was developed (QLQ-LC13; Appendix F) to be used in conjunction with the EORTC QLQ-C30 (Bergman et al 1994). It comprises both multi-item and single-item measures of lung cancer-associated symptoms (ie, coughing, hemoptysis, dyspnea, and pain) and side effects from conventional chemotherapy and radiotherapy (ie, hair loss, neuropathy, sore mouth, and dysphagia). Similar to the EORTC QLQ-C30, all questions except 1 have a 4-point scale: "Not at all," "A little," "Quite a bit," and "Very much." One question (#43 "Did you take any medicine for pain?") has a response option of "Yes" or "No." The scoring approach for the QLQ-LC13 is similar to the EORTC QLQ-C30.

5.3.1.3 PRO-CTCAE

The PRO-CTCAE is included to address tolerability from the patients' perspective. It was developed by the NCI. The PRO-CTCAE will only be administered in those countries where a linguistically validated version exists. It was developed in recognition that collecting side effect data directly from patients using PRO tools can improve the accuracy and efficiency of symptomatic AE data collection. This was based on findings from multiple studies demonstrating that physicians and nurses underestimate side effect onset, frequency, and severity in comparison with patient ratings (Antonia et al 2014b, Litwin et al 1998, Sprangers and Aaronson 1992). These side effects have been converted to patient terms (eg, the CTCAE term "myalgia" has been converted to "aching muscles"). For several symptoms, like fatigue and pain, additional questions are asked about the frequency, severity, and interference with usual activities. The items included in the PRO-CTCAE have undergone extensive qualitative review among experts and patients. These items have been extensively evaluated by cancer patients to be clear, comprehendible, and measure the symptom of interest. In this study, only items that are considered relevant for the trial, site of cancer, and cancer treatment are selected (see Appendix F).

5.3.1.4 Patients' Global Impression of Change

The PGIC item is included to assess how a patient perceives his/her overall change in health status since the start of study treatment. Patients will choose from response options from "Very Much Improved" to "Very Much Worse."

5.3.1.5 EQ-5D-5L

The EQ-5D is a standardized measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal (EuroQol Group 1990). Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care as well as in population health surveys. The questionnaire assesses 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 response options (no problems, slight problems, moderate problems, severe problems, and extreme problems) that reflect increasing levels of difficulty (EuroQol Group 2013).

Since 2009, the EuroQol group has been developing a more sensitive version of the EQ-5D (the EQ-5D-5L) which expands the range of responses to each dimension from 3 to 5 levels of increasing severity (Herdman et al 2011). Preliminary studies indicate that the 5L version improves upon the properties of the 3L measure in terms of reduced ceiling effect, increased reliability and an improved ability to differentiate between different levels of health (Janssen et al 2008a; Janssen et al 2008b; Pickard et al 2007).

The patient will be asked to indicate his/her current health state by selecting the most appropriate level in each of the 5 dimensions. The questionnaire also includes a visual analogue scale, where the patient will be asked to rate current health status on a scale of 0 to 100, with 0 being the worst imaginable health state (see Appendix F).

5.3.2 Administration of the patient-reported outcome questionnaires

Patients will complete the PRO assessments by using a handheld electronic device (ePRO).

Each center must allocate the responsibility for the administration of the PRO devices to a specific individual (eg, a research nurse or study coordinator) and, if possible, assign a back-up person to cover if that individual is absent. The PRO questionnaires must be completed per the schedule of assessments (see Table 2, Table 3 and Table 4). Patients will be instructed to bring their handheld devices to every clinic visit. In instances where the PRO assessments coincide with a clinic visit, the research nurse or study coordinator should verify that the questionnaire was completed before the visit. If not, the patient should complete the questionnaire prior to any other study procedures and before discussion of disease progression to avoid biasing the patient's responses to the questions.

The following best practice guidelines should be followed when collecting PRO data via an electronic device:

- The research nurse or appointed site staff must explain the value and relevance of participation to patients and inform them that these questions are being asked in order to find out from them directly how they feel. The research nurse or appointed site staff should also stress that the information is confidential. Therefore, if the patient has any medical problems, he/she should discuss them with the doctor or research nurse separately from the ePRO assessment.

- The research nurse or appointed site staff must train the patient on how to use the ePRO device using the materials and training provided by the ePRO vendor, and also provide guidance on whom to call if there are problems with the device.

- The research nurse or appointed site staff must remind patients that there are no right or wrong answers and avoid introducing bias by not clarifying items. The patient should not receive help from relatives, friends, or clinic staff to answer the PRO questionnaires.

- The research nurse or appointed site staff must monitor compliance; minimizing missing data is a key aspect of study success. Compliance must be checked at each study visit and should be checked more frequently to identify problems early. If compliance drops below 85%, a check-in call from the site to ask the patient if he/she has any difficulties is highly recommended.

5.3.3 WHO/ECOG performance status

WHO/ECOG performance status will be assessed at the times specified in the assessment schedules (see Table 2, Table 3 and Table 4) based on the following:

3. Fully active; able to carry out all usual activities without restrictions

4. Restricted in strenuous activity, but ambulatory and able to carry out light work or work of a sedentary nature (eg, light housework or office work)

5. Ambulatory and capable of self-care, but unable to carry out any work activities; up and about more than 50% of waking hours.

6. Capable of only limited self-care; confined to bed or chair more than 50% of waking hours

7. Completely disabled; unable to carry out any self-care and totally confined to bed or chair

Any significant changes from baseline or screening must be reported as an AE.

5.4 Pharmacokinetics

5.4.1 Collection of samples and determination of drug concentration

Blood samples for determination of MEDI4736 and tremelimumab concentration in serum will be obtained according to the assessment schedules (see Table 2 and Table 4).

Samples for determination of MEDI4736 and tremelimumab concentration in serum will be analyzed by a designated third party on behalf of AstraZeneca. Samples will be collected, labeled, stored, and shipped as detailed in the Laboratory Manual. Full details of the analytical method used will be described in a separate Bioanalytical Validation Report.

5.4.2 Collection of samples to measure for the presence of ADAs

The presence of ADA will be assessed in serum samples taken according to the assessment schedules (see Table 2, and Table 4).

Samples will be measured for the presence of ADAs and ADA-neutralizing antibodies for both IPs (MEDI4736 and tremelimumab) using validated assays. Tiered analysis will be performed to include screening, confirmatory, and titer assay components, and positive-negative cut points previously statistically determined from drug-naïve validation samples will be employed.

5.4.3 Storage and destruction of pharmacokinetic/ADA samples

PK and ADA samples will be disposed of a maximum of 10 years after the IPs are approved for marketing.

PK and ADA samples may be disposed of or destroyed and anonymized by pooling. Additional analyses may be conducted on the anonymized, pooled PK samples to further evaluate and validate the analytical method. Results from such analyses may be reported separately from the clinical study report (CSR).

Incurred sample reproducibility analysis, if any, will be performed alongside the bioanalysis of the test samples. The results from the evaluation will not be reported in the CSR but separately in a Bioanalytical Validation Report.

Any residual back-up PK samples may be used for future exploratory biomarker research (in this case, residual back-up PK samples will be shipped to AstraZeneca-designated Biobank; see details in the Laboratory Manual).

5.5 Biomarker analysis

The patient's consent to the use of donated biological samples is mandatory. Tissue samples will be obtained from all screened patients.

Pre-treatment tumor PD-L1 expression will be evaluated in all randomized patients. Data will be compared between groups to determine if baseline PD-L1 expression is prognostic and/or

predictive of outcomes associated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus SoC. Baseline tumor requirements are briefly described in Section 5.5.1.

Based on availability of tissue, additional exploratory biomarkers may be evaluated as described in Section 5.5.2. Also, descriptions of exploratory, peripheral measures are described in this section. Samples will be obtained according to the assessment schedules provided in Table 2, Table 3 and Table 4.

Details for collection, volumes, storage, and shipment of biologic samples are presented in a separate Laboratory Manual.

All samples collected for biomarker analyses will be stored at the study site, a reference laboratory, or at AstraZeneca-designated facilities and may be used for subsequent research relevant to evaluating biological and/or clinical response to immunotherapy.

The results may be pooled with biomarker data from other MEDI4736 and tremelimumab studies to evaluate biological responses across indications and to compare results in monotherapy versus combination settings.

5.5.1 Collection of patient samples for stratification by PD-L1

At screening, there are 2 mandatory options for provisions of tissue to be used for determination of eligibility. There is 1 subsequent mandatory provision of tissue at progression if retreatment is planned:

- MANDATORY: Provision of a recent tumor biopsy formalin fixed and embedded in paraffin. A freshly collected tumor biopsy is strongly preferred; however, if not clinically feasible, an archival sample taken less than 3 months prior to screening may be submitted.

Samples should be collected via a core needle (18 gauge or larger) or be collected as an excisional or incisional tumor biopsy sample.

When tissue is newly obtained for the purpose of entry into this study, 2 cores should be placed in formalin and processed to a single paraffin embedded block, as described in the Laboratory Manual.

The tumor specimen submitted to establish eligibility should be of sufficient quantity to allow for PD-L1 IHC analyses (see the Laboratory Manual). Newly acquired or archived specimens with limited tumor content and fine needle aspirates are inadequate for defining tumor PD-L1 status.

Tumor lesions used for fresh biopsies should not be the same lesions used as RECIST 1.1 target lesions, unless there are no other lesions suitable for biopsy. If a RECIST 1.1 target lesion is used for biopsy (preferably core needle biopsy), the lesion must be imaged after biopsy prior to randomization.

- MANDATORY: The collection of additional archived tumor tissue block greater than 3 months old (formalin-fixed paraffin-embedded) is mandated, where such samples exist in a quantity sufficient to allow for analysis. Tumor tissue block is preferred. If a tissue block is unavailable, unstained sections from the tissue block may be submitted. Please consult the laboratory manual for specific instructions and guidelines regarding sections.

- MANDATORY: The collection of tumor biopsies at the time of progression prior to retreatment is mandated. The Investigator can consult with the Study Physician if such sampling is not feasible, but retreatment is indicated.

- OPTIONAL: The collection of additional biopsies upon progression of patients in the MEDI4736 + tremelimumab and MEDI4736 monotherapy groups is strongly encouraged.

- Additional tumor biopsies collected as part of clinical care (eg, for mixed responses or upon PD) can be submitted for further analysis.

See the Laboratory Manual for further details of requirements including sample QC and shipping.

A brief description of exploratory tumor markers likely to be explored by IHC or RNA analysis is provided in Section 5.5.2.

The SP263 PD-L1 IHC assay will be used to determine PD-L1 IHC status in this study for the purposes of stratification and for the analysis of the original diagnostic sample.

To meet the requirement of Food and Drug Administration (FDA) approval of a companion diagnostic, sections of the tumor will be retained at or a approved laboratory for potential additional studies, as requested by the FDA, to support potential test approval.

5.5.2 Exploratory biomarkers

Blood and tumor samples for exploratory biomarker analyses will be obtained according to the schedules presented in Table 2, Table 3 and Table 4. Details for collection, volumes, storage, and shipment of biologic samples are presented in a separate Laboratory Manual.

Pharmacodynamic changes in biomarker measures will be monitored, when applicable. Baseline measures (and early, on-treatment changes) will be correlated with outcomes. Note that samples will be obtained from patients randomized to each treatment group. Comparisons will be made between baseline measures to determine if biomarkers (or combination of markers) are prognostic or predictive of outcomes associated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus SoC.

Additional sample collections and analyses may be completed at select study sites by site-specific amendment. All samples collected for such exploratory analyses will be stored at

site, a reference laboratory, or at AstraZeneca's facilities and may be used for subsequent research relevant to evaluating response to immunotherapy.

The exploratory biomarker plan is described by sample type below.

Whole blood for DNA/ SNP genotyping

Genomic DNA will be extracted from whole blood obtained pre-treatment from all patients. Genotyping of immunomodulatory genes such as, but not limited to, PD-1, PD-L1, CTLA-4, and human leukocyte antigen loci may be completed to determine if natural variation within such genes is associated with likelihood of clinical benefit (and/or with likelihood of drug-related AEs). Genes associated with NSCLC development, progression, or likelihood of response to chemotherapy may likewise be investigated. Genotyping will occur retrospectively, data will not be shared with patients, and results will not impact treatment decisions.

Genotypes also may be correlated with biomarker measures (eg, gene and/or protein expression) obtained from other sample types described in this exploratory biomarker section. A primary hypothesis is that different genotypes will be associated with different expression levels of factors within the PD-1 and CTLA-4 signaling pathways. Such variations in expression may affect the ability of an individual to mount an appropriate immune reaction to tumor and/or affect the likelihood of response to therapeutics targeting these pathways. Therefore, genotyping may provide easy-to-measure, baseline information regarding a patient's immune system, and a goal of this research is to understand how such genetic information may be used to predict pharmacodynamic responses to therapy.

Whole blood gene expression (PaxGene-RNA)

Whole blood samples will be obtained pre- or post-treatment from all patients as described in Table 2 or Table 3. Total RNA will be prepared for quantification of RNA and/or micro-ribonucleic acid (miRNA) expression using reverse transcription quantitative polymerase chain reaction (RT-QPCR), microarray, sequencing, or similar technology.

Focus is likely to be given to the expression of immunomodulatory genes previously found to be up-regulated in response to MEDI4736 and/or tremelimumab (data not shown). This battery of genes is likely to be similarly up-regulated in response to MEDI4736 + tremelimumab combination therapy (but not in response to SoC). Moreover, pre-treatment expression of such genes may indicate active immune responses that may be augmented by checkpoint inhibitor immunotherapies; correlations with outcome data will be completed on select candidate, predictive markers with the aim of identifying useful expression thresholds for identifying patients likely to receive benefit.

Myeloid-derived suppressor cells

Recent, collective findings suggest that a baseline measure of circulating MDSCs may be used as a prognostic tool in different disease settings and may specifically predict the likelihood of response to ipilimumab (anti-CTLA-4 therapy) (Kitano et al 2014, Meyer et al 2014). Flow

cytometry will be completed on all patients to quantify pre-treatment, circulating MDSC subtypes. Different MDSC count or percentage thresholds will be analyzed for their ability to predict clinical benefit from MEDI4736 + tremelimumab combination, MEDI4736 monotherapy, or SoC therapies.

Peripheral blood mononuclear cells

Whole blood samples will be collected for preparation of PBMCs and storage for potential downstream analyses. A variety of assays may be pursued, including but not limited to: immune cell composition/activation status analyses by flow cytometry, T cell functional assays (eg, Enzyme-Linked ImmunoSpot, receptor occupancy analyses to measure target engagement, tetramer analyses to monitor antigen-specific T cells, RNA/miRNA expression, and/or the assessment of the diversity and clonality of T cell receptor gene rearrangements using DNA.

Soluble factors - plasma

Plasma will be obtained pre- or post-treatment from all patients as described in Table 2, Table 3 and Table 4. The concentrations of a panel of cytokines and chemokines will be assessed. Focus is likely to be given to factors involved in T helper 1 cell-driven immune responses, including but not limited to IFN gamma, IL-18, chemokine (C-X-C motif) ligand (CXCL) 9, and CXCL10. Pharmacodynamic effects will be monitored to determine the specificity of response to immunotherapy (versus SoC). High pre-treatment expressions (concentrations) of such factors may indicate active immune responses, which may be augmented by checkpoint inhibitor immunotherapies; correlations with outcome data will be completed on select candidate, predictive markers with an aim of identifying useful expression thresholds for identifying patients likely to receive benefit, or alternatively, for identifying patients likely to suffer drug-related AEs.

Similarly, the concentrations of a battery of immune cell ligands or receptors may be assessed in plasma. Proteins of special interest include, but are not limited to, CTLA-4, PD-1, PD-L1, B7-1, B7-2, and IL-6R.

Focus may be given to circulating miRNAs currently thought to be putative, non-invasive prognostic biomarkers for cancer (Schwarzenbach et al 2014) and to subsets within which may be particularly useful for predicting responses to immunotherapies (Wang et al 2013). Additional, candidate miRNAs of interest may include those which regulate effectors within the PD-1 and CTLA-4 signaling pathways (eg, mIR-513, capable of PD-L1 down-regulation (Gong et al 2009).

Lastly, plasma may also be used for the detection/quantification of autoantibodies (against tumor-associated antigens). Seroconversion following treatment will be used as an indicator of overcoming tolerance. Pre-treatment seropositivity against specific antigens may provide predictive value, particularly when combined with data regarding the presence of antigen-specific T cells (Yuan et al 2011). Therefore, select, candidate autoantibody measures may be

evaluated for associations with clinical benefit and for directing PBMC-based research described in Section 5.5.1.

Tumor markers

Tissue obtained as part of screening procedures and for establishing PD-L1 expression status will be analyzed for additional markers by IHC. A primary goal is to measure cluster of differentiation (CD)8 and CD4/FoxP3 protein expression in an effort to enumerate cytotoxic versus regulatory T cells. Based on availability of tissue, a panel of additional, immune-relevant markers expressed on tumor-infiltrating lymphocytes or on tumor cells may be assessed. Markers of special interest include, but are not limited to, Ox40, GITR, PD-L2, Tim-3, CD137, and Lag 3.

Other tissue-based approaches may be pursued including RT-QPCR and in situ hybridization (eg, for detection of IFN gamma signaling genes such as CXCL9, CXCL10, and IFN gamma itself), and/or somatic mutation detection methodologies.

Management of biomarker data

The biomarker data will have unknown clinical significance. AstraZeneca will not provide biomarker research results to patients, their family members, any insurance company, an employer, clinical study Investigator, general physician, or any other third party, unless required to do so by law. The patient's samples will not be used for any purpose other than those described in the study protocol.

Individual patients will not be identified in any report or publication resulting from this work. The data and results of this research may be reviewed with collaborators and published, but neither the patient's name nor any other personal identifiers will appear in any publication or report.

5.5.3 Storage, re-use, and destruction of biological samples

Samples will be stored for a maximum of 15 years from the end of study, after which they will be destroyed. 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. The results of this biomarker research may be pooled with biomarker data from other studies involving MEDI4736 or tremelimumab to generate hypotheses to be tested in future research.

5.5.4 Labeling and shipment of biological samples

The Principal Investigator will ensure that samples are labeled and shipped in accordance with the Laboratory Manual and the Biological Substance, Category B, Regulations (materials containing or suspected to contain infectious substances that do not meet Category A criteria); see Appendix B "IATA 6.2 Guidance Document."

Any samples identified as Infectious Category A materials will not be shipped, and no further samples will be taken from the involved patients unless agreed upon with AstraZeneca and appropriate labeling, shipment, and containment provisions are approved.

5.5.5 Chain of custody of biological samples

A full chain of custody will be maintained for all samples throughout their life cycle.

The Principal Investigator at each center will keep full traceability of collected biological samples from the patients while in storage at the center until shipment or disposal (where appropriate) and will keep documentation of receipt of arrival.

The sample receiver will keep full traceability of the samples while in storage and during use until used or disposed of or until further shipment and will keep documentation of receipt of arrival.

AstraZeneca will keep oversight of the entire life cycle through internal procedures, monitoring of study sites, and auditing of external laboratory providers.

Samples retained for further use will be registered in the AstraZeneca-designated Biobank during the entire life cycle.

5.5.6 Withdrawal of informed consent for donated biological samples

If a patient withdraws consent to the use of donated biological samples, the samples will be disposed of or destroyed and the action documented. If samples have already been analyzed, AstraZeneca is not obliged to destroy the results of this research.

The Principal Investigator will:

- Ensure that AstraZeneca is immediately notified of the patients' withdrawal of informed consent to the use of donated samples
- Ensure that biological samples from that patient, if stored at the study site, are immediately identified, disposed of or destroyed and the action documented
- Ensure that the laboratory(ies) holding the samples is/are immediately informed about the withdrawn consent and that samples are disposed of or destroyed, the action is documented, and the signed document is returned to the study site
- Ensure that the patient and AstraZeneca are informed about the sample disposal

5.6 Pharmacogenetics

Refer to Appendix C for details of the genetic research (optional DNA component).

6. SAFETY REPORTING AND MEDICAL MANAGEMENT

The Principal Investigator is responsible for ensuring that all staff involved in the study are familiar with the content of this section.

6.1 Definition of adverse events

An AE is the development of an undesirable medical condition or the deterioration of a pre-existing medical condition during or following exposure to a pharmaceutical product, whether or not the condition is considered to be causally related to the product. An undesirable medical condition can be a symptom (eg, nausea or chest pain), sign (eg, tachycardia or enlarged liver), or the abnormal result of an investigation (eg, laboratory findings or ECG). In clinical studies, an AE can include an undesirable medical condition occurring at any time, including run-in or washout periods, even if no study treatment has been administered.

The term 'AE' is used to include both serious and non-serious AEs.

6.1.1 Causality collection

The Investigator will assess the causal relationship between the IPs and each AE and answer "yes" or "no" to the question "Do you consider that there is a reasonable possibility that the event may have been caused by the investigational product?"

For SAEs, causal relationship will also be assessed for other medications and study procedures. Note that, for SAEs that could be associated with any study procedure, the causal relationship is implied as "yes."

A guide to the interpretation of the causality question is found in Appendix A.

6.1.2 Relationship to protocol procedures

The Investigator is also required to provide an assessment of the relationship of SAEs to protocol procedures on the SAE report form. This includes both non-treatment–emergent (ie, SAEs that occur prior to the administration of MEDI4736 +/- tremelimumab and 30 days after last dose of SoC) and treatment-emergent SAEs. A protocol-related SAE may occur as a result of a procedure or intervention required during the study (eg, blood collection). The following guidelines should be used by Investigators to assess the relationship of SAEs to the protocol:

- Protocol related: The event occurred due to a procedure or intervention that was described in the protocol for which there is no alternative etiology present in the patient's medical record.
- Not protocol related: The event is related to an etiology other than the procedure or intervention that was described in the protocol. The alternative etiology must be documented in the study patient's medical record.

6.1.3 Adverse events based on signs and symptoms

All AEs spontaneously reported by the patient or reported in response to the open question from the study personnel: "Have you had any health problems since the previous visit/you were last asked?" or revealed by observation will be collected and recorded in the eCRF. When collecting AEs, the recording of diagnoses is preferred, when possible, to recording a list of signs and symptoms. However, if a diagnosis is known and there are other signs or symptoms that are not generally part of the diagnosis, the diagnosis and each sign or symptom will be recorded separately.

6.1.4 Adverse events based on examinations and tests

The results from protocol-mandated laboratory tests and vital signs measurements will be summarized in the CSR. Deterioration as compared to baseline in protocol-mandated laboratory values and vital signs should therefore only be reported as AEs if they fulfill any of the SAE criteria or are the reason for discontinuation of treatment with the IPs.

If deterioration in a laboratory value or vital sign is associated with clinical signs and symptoms, the sign or symptom will be reported as an AE and the associated laboratory result or vital sign will be considered as additional information. Whenever possible, the reporting Investigator should use the clinical rather than the laboratory term (eg, anemia versus low hemoglobin value). In the absence of clinical signs or symptoms, clinically relevant deteriorations in non-mandated parameters should be reported as AEs.

Deterioration of a laboratory value that is unequivocally due to disease progression should not be reported as an AE/SAE.

Any new or aggravated clinically relevant abnormal medical finding at a physical examination as compared with the baseline assessment will be reported as an AE.

6.1.5 Hy's law

Cases where a patient shows elevations in liver biochemistry may require further evaluation and occurrences of AST or ALT $\geq 3 \times$ ULN together with total bilirubin $\geq 2 \times$ ULN may need to be reported as SAEs. further instruction on cases of increases in liver biochemistry and evaluation of Hy's law. Further instruction on cases of increases in liver biochemistry and evaluation of Hy's law are shown in Appendix D.

6.1.6 Disease progression

Disease progression can be considered as a worsening of a patient's condition attributable to the disease for which the IP is being studied. It may be an increase in the severity of the disease under study and/or increases in the symptoms of the disease. The development of new or progression of existing metastasis to the primary cancer under study should be considered as disease progression and not an AE. Events that are unequivocally due to disease progression should not be reported as an AE during the study.

6.1.7 New cancers

The development of a new cancer should be regarded as an SAE. New primary cancers are those that are not the primary reason for the administration of the study treatment and have been identified after the patient's inclusion in this study.

6.1.8 Deaths

All deaths that occur during the study, or within the protocol-defined follow-up period after the administration of the last dose of study treatment, must be reported as follows:

- Death clearly resulting from disease progression should be reported to the Study Physician at the next monitoring visit and should be documented in the eCRF. It should not be reported as an SAE.

- Where death is not due (or not clearly due) to progression of the disease under study, the AE causing the death must be reported to the Study Physician as an SAE within 24 hours. The report should contain a comment regarding the co-involvement of PD, if appropriate, and should assign main and contributory causes of death.

- Deaths with an unknown cause should always be reported as an SAE. A post-mortem may be helpful in the assessment of the cause of death, and if performed, a copy of the post-mortem results should be forwarded to AstraZeneca Drug Safety or its representative within the usual timeframes.

6.1.9 Safety Data to be Collected Following the Final DCO of the Study

For patients continuing to receive durvalumab treatment after final DCO and database closure, it is recommended that the patients continue the scheduled site visits and investigators monitor the patient's safety laboratory results prior to and periodically during treatment with durvalumab in order to manage AEs in accordance with the durvalumab toxicity management guidelines (Appendix G). All data post the final DCO and database closure will be recorded in the patient notes but, with the exception of Serious Adverse Events, will not otherwise be reported for the purposes of this study.

All SAEs that occur in patients still receiving durvalumab treatment (or within the 90 days following the last dose of durvalumab treatment) post the final DCO and database closure must be reported as detailed in Section 6.4.

6.2 Definitions of serious adverse event

An SAE is an AE occurring during any study phase (ie, run-in, treatment, washout, or follow-up) that fulfills one or more of the following criteria:

- Results in death
- Is immediately life-threatening
- Requires in-patient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability or incapacity or substantial disruption of the ability to conduct normal life functions
- Is a congenital abnormality or birth defect
- Is an important medical event that may jeopardize the patient or may require medical intervention to prevent one of the outcomes listed above

For further guidance on the definition of an SAE, see Appendix A.

6.3 Recording of adverse events

6.3.1 Time period for collection of adverse events

AEs and SAEs will be collected from the time of the patient signing the informed consent form until the follow-up period is completed (90 days after the last dose of MEDI4736+/-tremelimumab and 30 days after last dose of SoC). AEs and SAEs collected prior to randomization will be reported as pre-randomization AEs and SAEs. If an event that starts post the defined safety follow up period noted above is considered to be due to a late onset toxicity to study drug then it should be reported as an AE or SAE as applicable.

6.3.2 Follow-up of unresolved adverse events

During the course of the study, all AEs and SAEs should be proactively followed up for each patient for as long as the event is ongoing. Every effort should be made to obtain a resolution for all events, even if the events continue after the patient has discontinued on study drug or the study has completed.

Any AEs that are unresolved at the patient's last visit in the study are followed up by the Investigator for as long as medically indicated, but without further recording in the eCRF. AstraZeneca retains the right to request additional information for any patient with ongoing AE(s)/SAE(s) at the end of the study, if judged necessary.

6.3.3 Variables

The following variables will be collected for each AE:

- AE (verbatim)
- The date when the AE started and stopped
- The maximum CTCAE grade reported
- Changes in CTCAE grade
- Whether the AE is serious or not
- Investigator causality rating against the IPs (yes or no)
- Action taken with regard to IPs
- Administration of treatment for the AE
- Whether the AE caused the patient's withdrawal from the study (yes or no)
- Outcome

In addition, the following variables will be collected for SAEs:

- Date the AE met criteria for SAE
- Date the Investigator became aware of the SAE
- Seriousness criteria fulfilled
- Date of hospitalization
- Date of discharge
- Probable cause of death
- Date of death
- Whether an autopsy was performed
- Causality assessment in relation to study procedure(s)
- Description of the AE

The grading scales found in the revised NCI CTCAE Version 4.03 will be utilized for all events with an assigned CTCAE grading. For those events without assigned CTCAE grades, the recommendation in the CTCAE criteria that converts mild, moderate, and severe events into CTCAE grades should be used. A copy of the CTCAE Version 4.03 can be downloaded from the Cancer Therapy Evaluation Program website (http://ctep.cancer.gov).

It is important to distinguish between serious and severe AEs. Severity is a measure of intensity, whereas seriousness is defined by the criteria in Section 6.2. An AE of severe intensity need not necessarily be considered serious. For example, nausea that persists for several hours may be considered severe nausea, but it is not an SAE unless it meets the criteria shown in Section 6.2. On the other hand, a stroke that results in only a limited degree of disability may be considered a mild stroke, but it would be an SAE if it satisfies the criteria shown in Section 6.2.

6.4 Reporting of serious adverse events

All SAEs have to be reported, whether or not they are considered causally related to the IPs or to any study procedure. All SAEs will be recorded in the eCRF.

If any SAE occurs in the course of the study, then the Investigator or other site personnel should inform the appropriate AstraZeneca representatives within 1 day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all the necessary information is provided to the AstraZeneca Patient Safety data entry site **within 1 calendar day** of initial receipt for fatal and life-threatening events **and within 5 calendar days** of initial receipt for all other SAEs.

For fatal or life-threatening AEs in which important or relevant information is missing, active follow-up is undertaken immediately. The Investigator or other site personnel will inform AstraZeneca representatives of any follow-up information on a previously reported SAE within 1 calendar day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

Once the Investigator or other site personnel indicates that an AE is serious in the WBDC system, an automated e-mail alert is sent to the designated AstraZeneca representative.

If the WBDC system is not available, then the Investigator or other study site personnel will report an SAE to the appropriate AstraZeneca representative by telephone.

The AstraZeneca representative will advise the Investigator or study site personnel how to proceed.

The reference documents for the definition of expectedness or listedness are the IBs for MEDI4736 and tremelimumab.

The Principal Investigator is responsible for ensuring that procedures and expertise are available to handle medical emergencies during the study. **A medical emergency usually constitutes an SAE and is to be reported as such.**

6.5 Overdose

Use of IP in doses in excess of that specified in the protocol is considered to be an overdose. There is currently no specific treatment in the event of overdose of IP, and possible symptoms of overdose are not established.

- An overdose with associated AEs will be recorded as the AE diagnosis or symptoms in the relevant AE modules of the eCRF and in the Overdose eCRF module.
- An overdose without associated symptoms will only be reported in the Overdose eCRF module.

If an overdose of an AstraZeneca study drug occurs in the course of the study, then the Investigator or other site personnel will inform appropriate AstraZeneca representatives immediately, or **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all relevant information is provided to the AstraZeneca Patient Safety data entry site.

For overdoses associated with an SAE, the standard reporting timelines apply, see Section 6.4. For other overdoses, reporting must occur within 30 days.

6.6 Pregnancy

All pregnancies and outcomes of pregnancy should be reported to AstraZeneca.

6.6.1 Maternal exposure

If a patient becomes pregnant during the course of the study, the IPs should be discontinued immediately.

Pregnancy itself is not regarded as an AE unless there is a suspicion that the IP under study may have interfered with the effectiveness of a contraceptive medication. Congenital abnormalities or birth defects and spontaneous miscarriages should be reported and handled as SAEs. Elective abortions without complications should not be handled as AEs. The outcome of all pregnancies (spontaneous miscarriage, elective termination, ectopic pregnancy, normal birth, or congenital abnormality) should be followed up and documented even if the patient was discontinued from the study.

If any pregnancy occurs in the course of the study, then the Investigator or other site personnel should inform the appropriate AstraZeneca representatives within 1 day, ie, immediately, but **no later than 24 hours** of when he or she becomes aware of it.

The designated AstraZeneca representative will work with the Investigator to ensure that all relevant information is provided to the AstraZeneca Patient Safety data entry site within 1 to 5 calendar days for SAEs (see Section 6.4) and within 30 days for all other pregnancies.

The same timelines apply when outcome information is available.

6.6.2 Paternal exposure

Male patients should refrain from fathering a child or donating sperm during the study and for 180 days after the last dose of MEDI4736 + tremelimumab combination therapy or 90 days after the last dose of MEDI4736 monotherapy. Please follow the local prescribing information relating to contraception and the time limit for such precautions for SoC agents.

Pregnancy of the patient's partner is not considered to be an AE. However, the outcome of all pregnancies (spontaneous miscarriage, elective termination, ectopic pregnancy, normal birth, or congenital abnormality) occurring from the date of the first dose until 90 days after the last dose should, if possible, be followed up and documented.

Where a report of pregnancy is received, prior to obtaining information about the pregnancy, the Investigator must obtain the consent of the patient's partner. Therefore, the local study team should adopt the generic ICF template in line with local procedures and submit it to the relevant Ethics Committees (EC)/Internal Review Boards (IRB) prior to use.

6.7 Management of investigational product-related toxicities

For guidance on the management of IP-related toxicities, please see Appendix G.

The following general guidance should be followed for management of toxicities.

- Treat each of the toxicities with maximum supportive care (including holding the agent suspected of causing the toxicity if required).
- If the symptoms promptly resolve with supportive care, consideration should be given to continuing the same dose of the assigned IP along with appropriate continuing supportive care. If medically appropriate, dose modifications are permitted (see Section 6.7.2). In addition, guidelines on dose modifications are provided in Appendix G.
- All dose modifications should be documented with clear reasoning and documentation of the approach taken.

In addition, guidelines on dose modification and toxicity management for immune-mediated reactions, infusion-related reactions, and non-immune-mediated reactions for MEDI4736 and tremelimumab will be provided to investigators (Appendix G). The most current version of the TMGs is also available through the following link: https://tmg.azirae.com. In addition a version of the current TMGs is maintained within the Site Master File. Please contact your clinical trial associate for information on how to gain access to this website.

There are certain circumstances in which MEDI4736 and tremelimumab should be permanently discontinued. Following the first dose of IP, subsequent administration of MEDI4736 and tremelimumab can be modified. All toxicities will be graded according to CTCAE Version 4.03. Dose reductions are not permitted. Dose modifications of MEDI4736 and tremelimumab may be required in the event of treatment-related toxicity. All toxicities will be graded according to NCI CTCAE Version 4.03. In case of doubt, the Investigator should consult with the Study Physician.

Acetaminophen and/or an antihistamine (eg, diphenhydramine) or equivalent medications per institutional standard may be administered prior to infusion at the discretion of the Investigator for primary prophylaxis against infusion-related reactions. In the event of Grade ≤2 infusion-related reaction, the infusion rate of IP may be decreased by 50% or interrupted until resolution of the event and re-initiated at 50% of the initial rate until completion of the infusion. In patients experiencing Grade ≤2 infusion-related reaction, subsequent infusions may be administered at 50% of the initial rate. If a patient experiences an infusion-related reaction, acetaminophen and/or an antihistamine (eg, diphenhydramine) and/or corticosteroid or equivalent medications per institutional standard may be administered prior to subsequent infusions at the discretion of the Investigator for secondary prophylaxis of infusion-related reactions. If the infusion-related reaction is Grade 3 or higher in severity, treatment with IP will be discontinued.

As with any antibody, allergic reactions to dose administration are possible. Appropriate drugs and medical equipment to treat acute anaphylactic reactions must be immediately available, and study personnel must be trained to recognize and treat anaphylaxis. The study site must have immediate access to emergency resuscitation teams and equipment in addition to the ability to admit patients to an intensive care unit if necessary.

6.7.1 MEDI4736 + tremelimumab and MEDI4736 monotherapy adverse events of special interest

An adverse event of special interest (AESI) is one of scientific and medical interest specific to understanding of the Investigational Product and may require close monitoring and rapid communication by the Investigator to the Sponsor. An AESI may be serious or non-serious. The rapid reporting of AESIs allows ongoing surveillance of these events in order to characterize and understand them in association with the use of this investigational product.

AESIs for MEDI4736 ± tremelimumab include, but are not limited to, events with a potential inflammatory or immune-mediated mechanism and that may require more frequent monitoring and/or interventions such as steroids, immunosuppressants, and/or hormone replacement therapy. These AESIs are being closely monitored in clinical studies with MEDI4736 monotherapy and combination therapy. An immune-mediated adverse event (imAE) is defined as an AE that is associated with drug exposure and is consistent with an immune-mediated mechanism of action and where there is no clear alternate aetiology. Serologic, immunologic, and histologic (biopsy) data, as appropriate, should be used to support an imAE diagnosis. Appropriate efforts should be made to rule out neoplastic, infectious, metabolic, toxin, or other etiologic causes of the imAE.

If the Investigator has any questions in regards to an AE being an imAE, the Investigator should promptly contact the Study Physician.

AESIs observed with MEDI4736 ± tremelimumab include:

- Diarrhea / Colitis and intestinal perforation
- Pneumonitis / ILD
- hepatitis / transaminase increases
- Endocrinopathies (ie, events of hypophysitis/hypopituitarism, adrenal insufficiency, hyper- and hypothyroidism and type I diabetes mellitus)
- Rash / Dermatitis
- Nephritis / Blood creatinine increases
- Pancreatitis / serum lipase and amylase increases
- Myocarditis
- Myositis / Polymyositis
- Neuropathy / neuromuscular toxicity (e.g. Guillain-Barré, and myasthenia gravis)
- Other inflammatory responses that are rare / less frequent with a potential immune-mediated aetiology include, but are not limited to, pericarditis, sarcoidosis, uveitis and other events involving the eye, skin, haematological and rheumatological events.

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

Further information on these risks (eg, presenting symptoms) can be found in the current version of the MEDI4736 IB. More specific guidelines for their evaluation and treatment are described in detail in Appendix G.

6.7.2 Immune-mediated adverse events

Based on the mechanism of action of MEDI4736 and tremelimumab leading to T-cell activation and proliferation, there is a possibility of observing imAEs during the conduct of this study. Potential imAEs may be similar to those seen with the use of ipilimumab, BMS-936558 (anti-PD-1 mAb), and BMS-936559 (anti-PD-L1 mAb) and may include immune-mediated enterocolitis, dermatitis, hepatitis (hepatotoxicity), pneumonitis, and endocrinopathies (Brahmer et al 2012, Hodi et al 2010, Topalian et al 2012). These AEs are inflammatory in nature and can affect any organ. With anti-PD-L1 and anti-CTLA-4 combination therapy, the occurrence of overlapping or increasing cumulative toxicities that include imAEs could potentially occur at higher frequencies than with either MEDI4736 or tremelimumab monotherapy. Patients should be monitored for signs and symptoms of imAEs. In the absence of an alternate etiology (eg, infection or PD), an immune-related etiology should be considered for signs or symptoms of colitis, dermatitis, pneumonitis, hepatitis, and

endocrinopathy. Dose modification guidelines are provided in Appendix G and it is recommended that:

1. Patients should be evaluated to identify any alternative etiology.

2. In the absence of a clear alternative etiology, all events of an inflammatory nature should be considered immune related.

3. Symptomatic and topical therapy should be considered for low-grade events.

4. Systemic corticosteroids should be considered for a persistent low-grade event or for a severe event.

5. More potent immunosuppressives should be considered for events not responding to systemic steroids (eg, infliximab or mycophenolate). If the Investigator has any questions in regards to an AE being an imAE, then the Investigator should immediately contact the Study Physician.

6.7.3 Standard of Care agents

IP-related toxicity management, including dose delays, reductions, and adjustments for patients in the SoC group should be performed as indicated in the local prescribing information for the relevant agent.

6.8 Study governance and oversight

The safety of all AstraZeneca clinical studies is closely monitored on an ongoing basis by AstraZeneca representatives in consultation with Patient Safety. Issues identified will be addressed; for instance, this could involve amendments to the study protocol and letters to Investigators.

An IDMC comprised of independent experts will be convened and will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review safety assessments and make recommendations to continue, amend, or stop the study based on safety findings. The committee will meet approximately every 6 months thereafter.

Full details of the IDMC procedures, processes, and interim analyses can be found in the statistical analysis plan and the IDMC Charter.

7. INVESTIGATIONAL PRODUCT AND OTHER TREATMENTS

7.1 Identity of investigational product(s)

AstraZeneca will supply MEDI4736 and tremelimumab, while the SoC treatments (paclitaxel + carboplatin, gemcitabine + cisplatin, gemcitabine + carboplatin, pemetrexed + cisplatin, pemetrexed + carboplatin, and pemetrexed maintenance) will be supplied locally (Table 8).

Table 8 List of investigational products for this study

Investigational product	Dosage form and strength	Manufacturer
MEDI4736	50 mg/mL, solution, IV	MedImmune
Tremelimumab	20 mg/mL, solution, IV	MedImmune
Standards of Care		
Paclitaxel[a]	IV (as sourced locally)	Sourced locally
Carboplatin[a]	IV (as sourced locally)	Sourced locally
Gemcitabine[a]	IV (as sourced locally)	Sourced locally
Cisplatin[a]	IV (as sourced locally)	Sourced locally
Pemetrexed[a]	IV (as sourced locally)	Sourced locally

[a] Under certain circumstances when local sourcing is not feasible, a Standard of Care treatment may be supplied centrally through AstraZeneca.
IV, intravenous.

7.1.1 MEDI4736

MEDI4736 will be supplied by AstraZeneca as a 500-mg vial solution for infusion after dilution. The solution contains 50 mg/mL MEDI4736, 26 mM histidine/histidine-hydrochloride, 275 mM trehalose dihydrate, and 0.02% (weight/volume) polysorbate 80; it has a pH of 6.0. The nominal fill volume is 10 mL. Investigational product vials are stored at 2°C to 8°C (36°F to 46°F) and must not be frozen.

Preparation of MEDI4736 doses for administration with an IV bag

The dose of MEDI4736 for administration must be prepared by the Investigator's or site's designated IP manager using aseptic technique.

Total time from needle puncture of the MEDI4736 vial to the start of administration should not exceed:

- 24 hours at 2°C to 8°C (36°F to 46°F)
- 4 hours at room temperature

A dose of 20 mg/kg will be administered using an IV bag containing 0.9% weight/volume (w/v) saline or 5% (w/v) dextrose, with a final MEDI4736 concentration ranging from 1 to 20 mg/mL, and delivered through an IV administration set with a 0.2 μm or 0.22 μm in-line filter.

Patient weight at baseline should be used for dosing calculations unless there is a ≥10% change in weight. Dosing day weight can be used for dosing calculations instead of baseline weight per institutional standard. A volume of 0.9% (w/v) saline or 5% (w/v) dextrose equal to the calculated volume of MEDI4736 to be added to the IV bag must be removed from the bag prior to addition of MEDI4736. The calculated volume of MEDI4736

is then added to the IV bag such that final concentration is within 1 mg/mL to 20 mg/mL. The bag is mixed by gentle inversion to ensure homogeneity of the dose in the bag.

Standard infusion time is 1 hour. However, if there are interruptions during infusion, the total allowed time should not exceed 8 hours at room temperature. In the event that either preparation time or infusion time exceeds the time limit, a new dose must be prepared using new vials. MEDI4736 does not contain preservatives, and any unused portion must be discarded.

No incompatibilities between MEDI4736 and polyvinylchloride or polyolefin have been observed.

Dose calculation

The volume of MEDI4736 (in mL) to add to the IV bag is calculated as follows:

20 mg/kg × patient weight (kg) ÷ MEDI4736 concentration (nominal: 50 mg/mL)

Example: For a patient weighing 80 kg, dosed at 20 mg/kg, 32 mL [20 mg/kg × 80 kg divided by 50 mg/mL] of MEDI4736 is to be diluted into an IV bag such that the final MEDI4736 concentration is within 1 to 20 mg/mL (100 – 1000 mL bag sizes). First, 32 mL of IV bag diluent is removed from the IV bag, and then 32 mL of MEDI4736 is added to the bag. The bag is mixed by gentle inversion to ensure homogeneity of the dose in the bag and the diluted MEDI4736 is administered as described above.

7.1.2 Tremelimumab

Tremelimumab will be supplied by AstraZeneca as a 400 mg vial solution for infusion after dilution. The solution contains 20 mg/mL tremelimumab, 20 mM histidine/histidine-hydrochloride, 222 mM trehalose dihydrate, 0.27 mM disodium edetate dihydrate and 0.02% (w/v) polysorbate 80; it has a pH of 5.5. The nominal fill volume is 20 mL. Investigational product vials are stored at 2°C to 8°C (36°F to 46°F) and must not be frozen.

Product preparation of tremelimumab for administration with an IV bag

The dose of tremelimumab for administration must be prepared by the Investigator's or site's designated IP manager using aseptic technique. Total time from needle puncture of the tremelimumab vial to start of administration should not exceed

- 24 hours at 2°C to 8°C (36°F to 46°F) or
- 4 hours at room temperature

A dose of 1 mg/kg will be administered using an IV bag containing 0.9% (w/v) saline, with a final tremelimumab concentration ranging from 0.15 mg/mL to 10 mg/mL, and delivered through an IV administration set with a 0.2 μm or 0.22 μm in-line filter.

Patient weight at baseline should be used for dosing calculations unless there is a ≥10% change in weight. Dosing day weight can be used for dosing calculations instead of baseline weight per institutional standard. The volume of 0.9% (w/v) saline equal to the volume of tremelimumab to be added to the IV bag must be removed from the bag prior to the addition of tremelimumab. The volume of tremelimumab is then added to the IV bag such that final concentration is within 0.15 mg/mL to 10 mg/mL. The bag is then mixed by gentle inversions to ensure homogeneity of the dose in the bag.

Standard infusion time is 1 hour, however if there are interruptions during infusion, the total allowed time should not exceed 8 hours at room temperature. In the event that either preparation time or infusion time exceeds the time limit, a new dose must be prepared using new vials. Tremelimumab does not contain preservatives, and any unused portion must be discarded.

No incompatibilities between tremelimumab and polyvinylchloride or polyolefin have been observed. However, polycarbonate syringes and administration set containing cellulose-based filters should not be used with tremelimumab.

Dose calculation of tremelimumab

The volume of tremelimumab (in mL) to add to the IV bag is calculated as follows:

In combination with MEDI4736: 1 mg/kg × patient weight (kg) ÷ tremelimumab concentration (nominal: 20 mg/mL)

Example: For a patient weighing 80 kg, dosed at 1 mg/kg, 4 mL [1 mg/kg × 80 kg divided by 20 mg/mL] of tremelimumab is to be diluted into 0.9% (w/v) saline bag such that the final tremelimumab concentration is within 0.15 to 10 mg/mL (100 – 500 mL bag sizes). First, 4 mL of IV bag diluent is removed from the IV bag, and then 4 mL of tremelimumab is added to the bag. The bag is mixed by gentle inversion to ensure homogeneity of the dose in the bag and the diluted tremelimumab is administered as described above..

7.1.3 Standard of Care treatment

Each SoC agent will be sourced as commercially available material/locally sourced, prescribed according to local regulations, and will be administered according to prescribing information or treatment guidance in general use by the Investigating site. Under certain circumstances when local sourcing is not feasible, an SoC will be supplied centrally by AstraZeneca. This will be labeled with local language translated text in accordance with regulatory guidelines.

7.2 Dose and treatment regimens

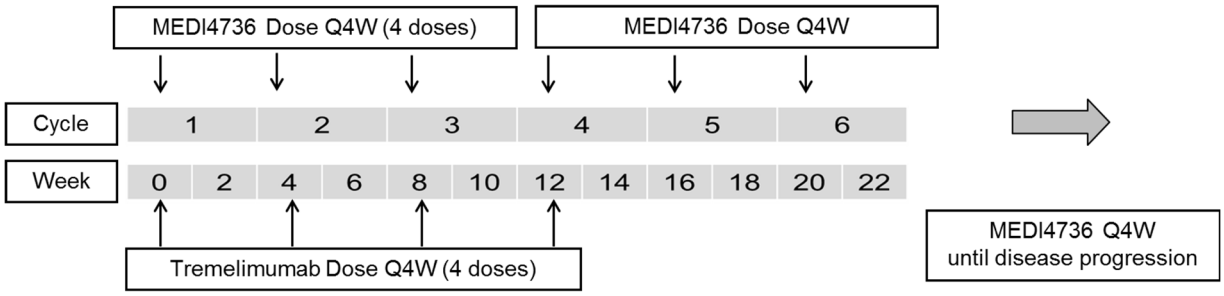
Patients will be randomized in a 1:1:1 ratio to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC.

7.2.1 Treatment regimens

MEDI4736 + tremelimumab combination therapy

Patients in the MEDI4736 + tremelimumab combination therapy group will receive 20 mg/kg MEDI4736 via IV infusion q4w for up to 4 doses/cycles and 1 mg/kg tremelimumab via IV infusion q4w for up to 4 doses/cycles, and then continue 20 mg/kg MEDI4736 q4w starting on Week 16 (see Figure 3). Dosing outside the window should be discussed with the Study Physician. Tremelimumab will be administered first. MEDI4736 infusion will start approximately 1 hour after the end of tremelimumab infusion. The duration will be approximately 1 hour for an infusion. A 1-hour observation period is recommended after the first infusion of MEDI4736 and tremelimumab. If no clinically significant infusion reactions are observed during or after the first cycle, subsequent infusion observation periods can be at the Investigator's discretion (suggested 30 minutes after each MEDI4736 and tremelimumab infusion).

Figure 3 MEDI4736 + tremelimumab combination therapy dosing scheme

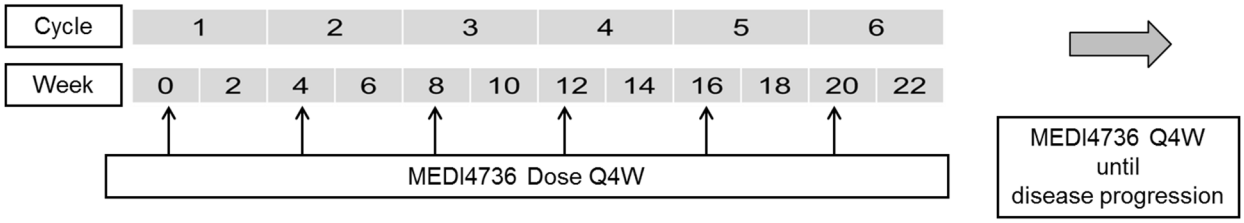


Q4W every 4 weeks.

MEDI4736 monotherapy

Patients in the MEDI4736 monotherapy treatment group will receive 20 mg/kg MEDI4736 via IV infusion q4w (see Figure 4).

Figure 4 MEDI4736 monotherapy dosing scheme



Q4W every 4 weeks.

Standard of Care treatment

Patients in the SoC group will receive 1 of the following treatments until documented PD (unconfirmed and confirmed) , initiation of alternative anticancer therapy, unacceptable toxicity, withdrawal of consent to continued treatment, or other reasons to discontinue treatment criterion occur:

- Paclitaxel + carboplatin: Paclitaxel 200 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + cisplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + cisplatin 75 or 80 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Gemcitabine + carboplatin (squamous patients only): Gemcitabine 1000 or 1250 mg/m^2 via IV infusion on Days 1 and 8 of each 21-day cycle + carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD
- Pemetrexed + cisplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and cisplatin 75 mg/m^2 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.
- Pemetrexed + carboplatin (non-squamous patients only): Pemetrexed 500 mg/m^2 and carboplatin AUC 5 or 6 via IV infusion on Day 1 of each 21-day cycle, for 4 to 6 cycles or until documented PD. Non-squamous patients who have not progressed after 4 cycles will be eligible for pemetrexed maintenance therapy.

For all SoC therapies, a particular treatment (paclitaxel, gemcitabine, cisplatin, carboplatin, or pemetrexed) will not be used in patients who have experienced recurrence or progression of disease within 6 months of prior multimodal therapy using that particular treatment.

A confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

7.2.2 Duration of treatment and criteria for retreatment

Duration of treatment

Unless specific treatment discontinuation criteria are met, patients in all groups will continue therapy until disease progression.

Progression during treatment

At the Investigator's discretion, patients in all groups may continue receiving therapy in the setting of unconfirmed PD until PD is confirmed. According to RECIST 1.1 modified for confirmation of progression, a confirmatory scan is required following an overall time-point assessment of progression (PD), preferably at the next scheduled visit and no earlier than 4 weeks after the previous assessment of PD.

Patients in all groups, excluding the SoC arm, with PD according to RECIST 1.1 (unconfirmed and confirmed) who, in the Investigator's opinion, continue to receive benefit from their assigned treatment and who meet the criteria for treatment in the setting of PD may continue to receive their assigned treatment for as long as they are gaining clinical benefit.

However, patients in the immunotherapy group(s) will not be permitted to continue immunotherapy if progression occurs in a target lesion that has previously shown a confirmed response. Confirmed response is CR or PR, as defined by RECIST 1.1.

Patients in the MEDI4736 + tremelimumab group may restart treatment with the combination therapy if they complete the 4 dosing cycles with MEDI4736 + tremelimumab (with clinical benefit per Investigator's judgment), but subsequently have PD during treatment with MEDI4736 alone and if they meet eligibility criteria for retreatment.

For all groups, excluding the SoC arm, treatment through progression and retreatment in the MEDI4736 + tremelimumab combination therapy group are at the Investigator's discretion, and the Investigator should ensure that patients do not have any significant, unacceptable, or irreversible toxicities that indicate that continuing treatment will not provide further benefit. The Investigator should ensure that patients still meet all of the inclusion criteria and none of the exclusion criteria for this study and that these patients meet the following specific criteria for treatment in the setting of PD:

- Written informed consent for retreatment in the setting of PD. This consent document will specify that treatment beyond initial evidence of PD is not the SoC

and that alternative treatment options, either locally licensed treatments or other clinical trials, are available for this patient population.

- Absence of clinical symptoms or signs indicating clinically significant disease progression and no decline in WHO performance status to >1

- Absence of rapid disease progression or threat to vital organs or critical anatomical sites (eg, central nervous system metastasis, respiratory failure due to tumor compression, or spinal cord compression) requiring urgent alternative medical intervention

Post final Data Cut Off (DCO)

Patients who continue to receive benefit from their assigned treatment at the final DCO and database closure may continue to receive their assigned treatment for as long as they and their physician feel they are gaining clinical benefit. For patients continuing to receive durvalumab treatment following the final DCO and database closure, it is recommended that the patients continue the scheduled site visits and investigators monitor the patient's safety laboratory results prior to and periodically during treatment with durvalumab in order to manage AEs in accordance with the durvalumab toxicity management guidelines (please see Section 6.1.9).

In the event that a roll-over or safety extension study is available at the time of the final DCO and database closure, patients currently receiving treatment with durvalumab may be transitioned to such a study, and the current study would reach its end. The roll-over or safety extension study would ensure treatment continuation with visits assessment per its protocol. Any patient that would be proposed to move to such study would be given a new Informed Consent.

7.3 Labeling

Labels will be prepared in accordance with Good Manufacturing Practice (GMP) and local regulatory guidelines. The labels will fulfill GMP Annex 13 requirements for labelling. Label text will be translated into the local language.

Labels will be provided as either a single panel label or as multi-language booklet labels.

7.4 Storage

All study drugs should be kept in a secure place under appropriate storage conditions. The IP label on the pack/bottle/carton specifies the appropriate storage. Storage is also described in the IB.

7.5 Compliance

The administration of all study drugs (including IP) should be recorded in the appropriate sections of the eCRF.

Treatment compliance will be assured by reconciliation of site drug accountability logs.

7.6 Accountability

The study drug provided for this study will be used only as directed in the study protocol. The study personnel will account for all study drugs

Drug accountability should be performed until the patient stops study treatment completely. Study site personnel will account for all study drugs received at the site, for all unused study drugs, and for appropriate destruction of study drugs. Certificates of delivery, destruction, and return should be signed.

7.7 Concomitant and other treatments

The Investigator must be informed as soon as possible about any medication taken from the time of screening until the end of the clinical treatment phase of the study including the 3 month follow up period following the last dose of study drug. Any concomitant medication(s), including herbal preparations, taken during this time will be recorded in the eCRF.

Patients must be instructed not to take any medications, including over-the-counter products, without first consulting with the Investigator.

Restricted, prohibited, and permitted concomitant medications are described in the following tables. Refer also to the Dosing Modification and Toxicity Management Guidelines in Appendix G. For agents in the SoC arm, please refer to the local prescribing information with regards to warnings, precautions, and contraindications.

Prohibited medication/class of drug:	**Usage:**
Additional investigational anticancer therapy concurrent with those under investigation in this study	Should not be given whilst the patient is on IP treatment (including SoC)
mAbs against CTLA-4, PD-1, or PD-L1	Should not be given whilst the patient is on IP treatment (including SoC) through 90 days after the last dose of IP.
Any concurrent chemotherapy, local therapy (except palliative radiotherapy for non-target lesions, eg, radiotherapy, surgery, radiofrequency ablation), biologic therapy, or hormonal therapy for cancer treatment	Should not be given whilst the patient is on IP treatment (including SoC). (Concurrent use of hormones for non-cancer-related conditions [eg, insulin for diabetes and hormone replacement therapy] is acceptable.)

Prohibited medication/class of drug:	Usage:
Immunosuppressive medications, including, but not limited to: systemic corticosteroids at doses exceeding 10 mg/day of prednisone or its equivalent, methotrexate, azathioprine, and tumor necrosis factor -α blockers	Should not be given whilst the patient is on IP treatment (including SoC). (Use of immunosuppressive medications for the management of IP-related AEs or in patients with contrast allergies is acceptable. In addition, use of inhaled, topical, and intranasal corticosteroids is permitted. Systemic corticosteroids may be used for prevention of chemotherapy-related toxicities (nausea/vomiting prevention and prophylaxis) A temporary period of steroids will be allowed if clinically indicated and considered to be essential for the management of the patient (eg, for chronic obstructive pulmonary disease, radiation, nausea, etc).
Drugs with laxative properties	Should be used with caution for 90 days after the last dose of tremelimumab during the study
Herbal or natural remedies (all herbal or natural remedies, rather than those for constipation only)	Should be avoided for 90 days after the last dose of MEDI4736 monotherapy or combination treatment (MEDI4736+ tremelimumab)
Sunitinib	Should not be given within 3 months of a dose of tremelimumab (acute renal failure has been reported with combination therapy of tremelimumab and sunitinib)
Live attenuated vaccines	Should not be given through 30 days after the last dose of IP (including SoC) during the study
EGFR TKIs	Should not be given concomitantly whilst the patient is on study treatment. In addition they should be used with caution in the 90 days after the last dose of durvalumab. (Increased incidences of pneumonitis (with third generation EGFR TKIs) and increased incidence of transaminase increases (with 1st generation EGFR TKIs) has been reported when durvalumab has been given concomitantly).

Rescue/supportive medication/class of drug:	Usage:
Concomitant medications or treatments (eg, acetaminophen or diphenhydramine) deemed necessary by the Investigator to provide adequate prophylactic or supportive care, except for those medications identified as "prohibited" as listed above	To be administered as prescribed by the Investigator

Rescue/supportive medication/class of drug:	Usage:
Best supportive care (including antibiotics, nutritional support, growth factor support, correction of metabolic disorders, optimal symptom control, and pain management [including palliative radiotherapy, etc])	Should be used when necessary for all patients
Inactivated viruses, such as those in the influenza vaccine	Permitted

7.7.1 Other concomitant treatment

Medications other than those described in Section 7.7 that are considered necessary for the patient's safety and well-being may be given at the discretion of the Investigator and should be recorded in the appropriate sections of the eCRF.

7.8 Post study access to study treatment

After the final analysis, AstraZeneca will continue to supply open-label drug to patients receiving MEDI4736 + tremelimumab or MEDI4736 monotherapy (see Section 7.2.2).

8. STATISTICAL ANALYSES BY ASTRAZENECA

8.1 Statistical considerations

All statistical analyses will be performed by AstraZeneca or its representatives.

A comprehensive statistical analysis plan (SAP) will be prepared and finalized within 3 months of the first randomized patient and any subsequent amendments will be documented, with final amendments completed prior to reporting of the data. The primary aim of the study is to compare the efficacy and safety of MEDI4736 in combination with tremelimumab and MEDI4736 monotherapy to SoC.

8.2 Sample size estimate

The study will plan to enroll approximately 1850 patients in order to randomize 1092 eligible patients 1:1:1 to MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC. The 1092 patients will comprise approximately 480 patients who have PD-L1 positive$_{25\%}$ tumors.

The study is sized to characterize the OS benefit of MEDI4736 in combination with tremelimumab and MEDI4736 monotherapy versus SoC in patients with EGFR and ALK wild-type advanced or metastatic NSCLC in patients with PD-L1 positive$_{25\%}$ tumors and PFS benefit of MEDI4736 in combination with tremelimumab versus SoC in patients with PD-L1 positive$_{25\%}$ tumors. The sizing for PFS and OS assumes a 3-month delay in separation of the PFS curves between each group, hence the use of average hazard ratios (HRs).

Two interim analysis of OS will be performed; the first at the time of the primary PFS analysis and the second when 80% of the target OS events have occurred. The alpha will be split between the 3 OS analyses using the Lan and DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, with the boundaries for the treatment comparison derived based upon the exact number of OS events at the time of analysis.

The primary PFS analysis for superiority will be performed when both of the following conditions have been met:

- Approximately 231 BICR PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (72% maturity) AND
- Approximately 44 weeks follow up from last patient randomized to the study

The final (primary) OS analysis for superiority will be performed when the following conditions have been met:

- Approximately 225 OS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity) AND

- Approximately 225 OS events have occurred across the MEDI4736 monotherapy and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity)

MEDI4736 + tremelimumab versus SoC (PFS in PD-L1 positive$_{25\%}$ population)

If PFS at 12 months was 34.1% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.59, with 231 PFS events(72% maturity), the trial will have 88% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.69. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (OS in PD-L1 positive$_{25\%}$ population)

If OS at 18 months was 53.2% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average OS HR is 0.62, an estimated 225 death events (70% maturity) are expected to have occurred at 33 months from "first patient in." With at minimum 225 deaths, the study will have 86% power to demonstrate statistical significance at the 2-sided alpha level of 1.32% (with overall alpha for OS 1.5%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.72.

MEDI4736 monotherapy versus SoC (OS in PD-L1 positive$_{25\%}$ population)

If OS at 18 months was 53.2% with MEDI4736 monotherapy and 36% with SoC (with a 12.9 month median OS and assuming the true average OS HR is 0.62, an estimated 225 death events (70% maturity) are expected to have occurred at 33 months from "first patient in." With at minimum 225 deaths, the study will have 90% power to demonstrate statistical significance at the 2-sided alpha level of 2.58% (with overall alpha for OS 3%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.74.

MEDI4736 + tremelimumab versus SoC (OS in PD-L1 positive$_{1\%}$ population)

The overall alpha level for this comparison will be 1.5% (if only the OS comparison of comination versus SoC in PD-L1 positive$_{25\%}$ population is significant) or 3% (if only the OS comparison of monotherapay versus SoC in PD-L1 positive$_{25\%}$ population is significant) or 4.5% (if both these hypotheses are significant) (See Multiple Testing Procedure in Section 8.5).

With approximately 236 patients with PD-L1 positive$_{1\%}$ tumors per treatment arm (*i.e.,* 65% prevalence rate), if OS at 18 months was 51.1% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS), and assuming the true average OS HR is 0.66, an

estimated 337 death events (71% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum of 337 deaths, the study will have at least 90% power to demonstrate statistical significance at the 2-sided alpha level of 1.32%, (with overall alpha for OS 1.5%,), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an at least average HR of 0.76.

MEDI4736 + tremelimumab versus SoC (OS, all patients)

Similary as the comparison of OS between MEDI4736 + tremelimumab versus SoC arm in PD-L1 positive$_{1\%}$ population, the overall alpha level for this comparison will be 1.5% (if only the OS comparisons of comination versus SoC in PD-L1 positive$_{25\%}$ population and in PD-L1 positive$_{1\%}$ population are significant) or 3% (if only the OS comparison of monotherapay versus SoC in PD-L1 positive$_{25\%}$ population and the OS comparison of comination versus SoC in PD-L1 positive$_{1\%}$ population are significant) or 4.5% (if all these hypotheses are significant) (See Multiple Testing Procedure in Section 8.5).

If OS at 18 months was 49% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS) and assuming the true average OS HR is 0.70, an estimated 528 death events (73% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum of 528 deaths, the study will have at least 90% power to demonstrate statistical significance at the 2-sided alpha level of 1.32% (with overall alpha for OS 1.5%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.81.

MEDI4736 monotherapy versus SoC (PFS in PD-L1 positive$_{25\%}$ population)

If PFS at 12 months was 34.1% with MEDI4736 monotherapy and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.59, with 231 PFS events (72% maturity), the trial will have 88% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 monotherapy versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.69. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PFS in PD-L1 positive$_{1\%}$ population)

With approximately 236 patients with PD-L1 positive$_{1\%}$ tumors per treatment arm (*i.e.,* 65% prevalence rate), if PFS at 12 months was 30.2% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.64, with 348 PFS events (74% maturity), the trial will have 90% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.74. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PFS, all patients)

If PFS at 12 months was 26% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.71, with 551 PFS events (76% maturity),, the trial will have 89% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.79. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

8.3 Definitions of analysis sets

Definitions of the analysis sets for each outcome variable are provided in Table 9.

Table 9 Summary of outcome variables and analysis populations

Outcome variable	Population
Efficacy data	
PFS and OS	PD-L1 positive$_{25\%}$ analysis set
PFS and OS	PD-L1 positive$_{1\%}$ analysis set
PFS and OS	Full analysis set (ITT population)
ORR, DoR, APF12, PFS2, PROs, and symptom endpoints	PD-L1 positive$_{25\%}$ analysis set
ORR, DoR, APF12, PFS2, PROs, and symptom endpoints	PD-L1 positive$_{1\%}$ analysis set
PFS, OS, ORR, , DoR, APF12, PFS2	PD-L1-low/negative analysis set
ORR, DoR, APF12, PFS2, PROs, and symptom endpoints	Full analysis set (ITT population)
PK data	PK analysis Set

Outcome variable	Population
Safety Data	
Exposure	Safety analysis Set
AEs	Safety analysis Set
Laboratory measurements	Safety analysis Set
Vital signs	Safety analysis Set
ECGs	Safety analysis Set

8.3.1 Full analysis set

The full analysis set (FAS) will include all randomized patients. The full analysis set will be used for all efficacy analyses (including PROs). Treatment groups will be compared on the basis of randomized study treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the analysis in the treatment group to which they were randomized.

8.3.2 PD-L1 positive$_{25\%}$ analysis set

The PD-L1 positive$_{25\%}$ analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive$_{25\%}$ as defined by the SP263 PD-L1 IHC assay (ie, ≥25% PD-L1–membrane expression in tumoral tissue).

8.3.3 PD-L1 positive$_{1\%}$ analysis set

The PD-L1 positive$_{1\%}$ analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive$_{1\%}$ as defined by the SP263 PD-L1 IHC assay (ie, ≥1% PD-L1–membrane expression in tumoral tissue). This analysis set will be used for for supportive analyses of efficacy endpoints.

8.3.4 PD-L1- low/negative analysis set

The PD-L1-low/negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 low/negative as defined by the SP263 PD-L1 IHC assay (ie, <25% PD-L1–membrane expression in tumoral tissue).

8.3.5 PD-L1- negative analysis set

The PD-L1-negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 negative as defined by the SP263 PD-L1 IHC assay (ie, <1% PD-L1–membrane expression in tumoral tissue).

8.3.6 Safety analysis set

The safety analysis set (SAS) will consist of all patients who received at least 1 dose of study treatment. Safety data will not be formally analyzed but summarized using the SAS, according to the treatment received, that is, erroneously treated patients (e.g., those randomized to treatment A but actually given treatment B) will be summarized according to the treatment they actually received.

8.3.7 Pharmacokinetic analysis set

All patients who received at least 1 dose of IP per the protocol for whom any post-dose data are available and who do not violate or deviate from the protocol 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.

8.4 Outcome measures for analyses

8.4.1 Calculation or derivation of efficacy variables

The analysis of the co-primary endpoint, PFS, and the analyses of the secondary endpoints, ORR, DoR, and APF12, will be based on BICR tumor assessments according to RECIST 1.1. OS will be evaluated as a co-primary endpoint from all-cause mortality. Additionally, PFS2 will be defined by local clinical practice.

A sensitivity analysis of PFS will be performed using the Investigator tumor assessments. In addition, PFS and ORR by irRECIST 1.1 criteria using BICR assessments will also be performed for exploratory purposes.

8.4.2 RECIST 1.1-based endpoints

Blinded Independent Central Review of RECIST 1.1-based assessments

The BICR will be performed on all radiological scans of all patients. All images will be collected centrally. Prior radiotherapy reports will also be provided to the BICR to allow the selection of appropriate target lesions. The imaging scans will be reviewed by 2 independent radiologists using RECIST 1.1 and will be adjudicated, if required. For each patient, the BICR will define the overall visit response data (CR, PR, SD, PD, or not evaluable [NE]). If a patient has had a tumor assessment that cannot be evaluated, then the patient will be assigned a visit response of NE (unless there is evidence of progression, in which case the response will be assigned as PD). Endpoints (of PFS) will be derived from the scan dates.

PFS and ORR by irRECIST 1.1 criteria using BICR assessments will also be performed for exploratory purposes.

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

Investigator RECIST 1.1-based assessments

All RECIST 1.1 assessments, 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.

At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD, or PD depending on the status of their disease compared with baseline and previous assessments. Baseline will be assessed within the 28 days prior to randomization. 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).

The definitions of CR, PR, SD, and PD are given in Appendix E.

8.4.2.1 Co-Primary endpoints

OS and PFS are the co-primary endpoints.

Progression-free survival

PFS (per RECIST 1.1 using BICR assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression. 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. If the patient has no evaluable visits or does not have baseline data, they will be censored at 0 days unless they die within 2 visits of baseline.

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:

- The date of progression will be determined based on the earliest of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or of either reviewer where both select PD as a time point response and there is no adjudication for central review (BICR).

- For investigational assessments, the date of progression will be determined based on the earliest of the RECIST 1.1 assessment/scan dates of the component that indicates 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 target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the target lesions, and similarly for non-target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the non-target lesions.

PFS based on RECIST 1.1 BICR data modified for confirmation of progression will be performed for exploratory purposes using the algorithm described above for the RECIST 1.1 BICR assessments, but following a modification whereby any objective disease progression must be confirmed by the next scheduled scan. The confirmatory scan must be no sooner than 4 weeks after the initial suspected progression and preferably at the next scheduled visit in the absence of clinically significant deterioration. If disease progression is confirmed (or disease progression occurs and no further scans are recorded) then the date of progression will be when it was originally observed. Patients with a single disease progression and no further tumor assessment scans will be treated as PD in the analysis. In the absence of significant clinical deterioration, the investigational site is advised to continue the patient on their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy until progression has been confirmed. If progression is not confirmed, the patient should continue their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy treatment and on-treatment assessments. In the SoC arm, treatment through PD, until PD is confirmed, is at the Investigator's discretion; however, a confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

PFS by irRECIST 1.1 criteria using BICR assessments will also be reported.

Overall survival

OS is defined as the time from the date of randomization until death due to any cause. 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.

Note: Survival calls will be made following the date of data cut-off for the analysis (these contacts should generally occur within 7 days of the data cut off). If patients are confirmed to be alive or if the death date is after the data cut-off date, these patients will be censored at the date of data cut-off. Death dates may be found by checking publicly available death registries.

8.4.2.2 Secondary endpoints

Objective response rate

ORR (per RECIST 1.1 using BICR assessments) is defined as the number (%) of patients with at least 1 visit response of CR or PR. 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 will also be obtained using the algorithm described above for the RECIST 1.1 site Investigator tumor data.

ORR by irRECIST 1.1 criteria using BICR assessments will also be reported.

Duration of response

DoR (per RECIST 1.1 using BICR assessments) will be defined as the time from the date of first documented response until the first date of documented progression or death in the absence of disease progression. 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 CR or PR. 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.

Time from randomization to second progression (PFS2)

PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death. The date of second progression will be recorded by the Investigator in the eCRF and defined according to local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression, or death. Second progression status will be reviewed (every 6 weeks for the first 48 weeks relative to the date of randomization and then every 8 weeks thereafter) following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression, that is, censored at the latest of the PFS or PFS2 assessment date if the patient has not had a second progression or death.

Proportion of patients alive and progression free at 12 months

The APF12 will be defined as the Kaplan-Meier estimate of PFS (per RECIST 1.1 as assessed using BICR assessments) at 12 months.

Best objective response

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST 1.1 assessment, described in Appendix D. It is the best response a patient has had during their time in the study up until RECIST progression or the last evaluable assessment in the absence of RECIST.1.1 progression, as determined by BICR.

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

BoR will be determined programmatically based on RECIST 1.1 using all BICR 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 1.1assessments, if the death occurs ≤90 days (ie, 2*(6 weeks ±3 days)) after randomization, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST 1.1assessments, if the death occurs >90 days (ie, 2*(6 weeks ±3 days)) after the date of randomization then BoR will be assigned to the NE category.

Progression events that have been censored due to them being >90 days after the last evaluable assessment will not contribute to the BoR derivation.

8.4.3 Calculation or derivation of safety variables

8.4.3.1 Adverse events

Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of Medical Dictionary for Regulatory Activities [MedDRA] preferred terms and CTCAE grade) will be listed individually by patient.

Any AE occurring before treatment with IP will be included in the data listings but will not be included in the summary tables of AEs. Any AE occurring within 90 days of discontinuation of the last dose of MEDI4736 + tremelimumab combination therapy or MEDI4736 monotherapy, or within 30 days of the last dose of SoC) may be included in the AE summaries, but the majority of those summaries will omit those AEs observed after a patient has received further therapy for cancer. Further details will be provided in the SAP. Any events in this period that occur after a patient has received further therapy for cancer (following discontinuation of IP) will be flagged in the data listings.

A separate data listing of AEs occurring more than 90 days after discontinuation of MEDI4736 + tremelimumab combination therapy or MEDI4736 monotherapy, or more than 30 days after discontinuation of SoC will be produced. These events will not be included in AE summaries.

8.4.3.2 Other significant adverse events

During the evaluation of the AE data, an AstraZeneca medically qualified expert will review the list of AEs that were not reported as SAEs and AEs leading to discontinuation. Based on the expert's judgment, significant AEs of particular clinical importance may, after consultation with the Global Patient Safety Physician, be considered other significant adverse events (OAEs) and reported as such in the CSR. A similar review of laboratory/vital signs/ECG data will be performed for identification of OAEs. Examples of these are marked hematological and other laboratory abnormalities, and certain events that lead to intervention (other than those already classified as serious) or significant additional treatment.

8.4.3.3 Safety assessments

For the change from baseline summaries for vital signs, laboratory data, ECGs, and physical examination, the baseline value will be the latest result obtained prior to the start of study treatment.

The QTcF will be derived during creation of the reporting database using the reported ECG values (RR and QT).

QTcF = QT/RR^ (1/3) where RR is in seconds

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

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

The denominator used in laboratory summaries will only include evaluable patients, in other words those who had sufficient data to have the possibility of an abnormality.

For example:

- If a 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 CTCAE criterion does not consider changes from baseline, to be evaluable the patient need only have 1 post dose-value recorded.

The denominator in vital signs data should include only those patients with recorded data.

8.4.4 Calculation or derivation of patient-reported outcome variables

PRO questionnaires, a secondary endpoint of interest will be assessed using the EORTC QLQ-C30 with the QLQ-LC13 module (HRQoL and lung cancer specific symptoms), PRO-CTCAE, PGIC, and EQ-5D-5L. All items/questionnaires will be scored according to published scoring guidelines or the developer's guidelines, if published guidelines are not available. All PRO analyses will be based on the Full Analysis Set (FAS; ITT population), and the PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ analysis sets, unless otherwise stated. The clinical meaningfulness threshold of the PRO analyses described below will be provided in the SAP.

8.4.4.1 EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea/vomiting), 5 individual items (dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a global measure of health status. The EORTC QLQ-C30 will be scored according to the EORTC QLQ-C30 scoring manual (Fayers et al 2001). An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, each of the functional scales, and the global health status scale in the EORTC QLQ-C30 according to the EORTC QLQ-C30 Scoring Manual. Higher scores on the global health status and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity.

Changes in score compared with baseline will be evaluated. For each subscale, if <50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded. If there is evidence that the missing data are systematic, missing values will be handled to ensure that any possible bias is minimized.

Time to HRQoL/function deterioration

For the following HRQoL items of the EORTC QLQ-C30, time to deterioration will be analyzed:

The Global Health Status/ QoL scale consisting of items 29 and 30 of the EORTC QLQ-C30. Item 29: How would you rate your overall health during the past week? Item 30: How would you rate your overall quality of life during the past week? Patients are asked to rate their overall health and overall quality of life on a scale from 1 (very poor) to 7 (excellent).

Time to deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful deterioration that is confirmed at a subsequent visit or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to HRQoL/function deterioration. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the HRQoL/function change could be evaluated.

Patients whose HRQoL (as measured by EORTC QLQ-C30) has not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the HRQoL/function could be evaluated. Also, if HRQoL deteriorates after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated. If a patient has no evaluable visits or does not have baseline data they will be censored at 0 days.

Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement in that symptom from baseline.

HRQoL/function improvement rate

The HRQoL/function improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement in that scale from baseline.

8.4.4.2 Lung cancer module (EORTC QLQ-LC13)

The QLQ-LC13 is a lung cancer specific module from the EORTC for lung cancer comprising 13 questions to assess lung cancer symptoms (cough, hemoptysis, dyspnea, and sute-specific pain), treatment-related side effects (sore mouth, dysphagia, peripheral neuropathy, and alopecia), and pain medication. Changes in score compared with baseline will be evaluated.

Time to symptom deterioration

For each of the following key symptom scales/items in the QLQ-LC13 and QLQ-C30, time to deterioration will be analyzed:

- Dyspnoea (multi-item scale based on three questions: "Were you short of breath when you rested; walked; climbed stairs") LC13,
- Cough: one item ("How much did you cough?"), LC13
- Pain: three individual items ("Have you had pain in your chest; your arm or shoulder; other parts of your body?"). LC13
- Appetite Loss (" Have you lacked appetite?") C30
- Fatigue ("Have you felt weak?", "Did you need to rest?" "Were you tired?") C30

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales/single items of the EORTC QLQ-C30.

Time to symptom deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful symptom deterioration that is confirmed at a subsequent visit or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to symptom deterioration. Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the symptom change could be evaluated.

Patients whose symptoms (as measured by QLQ-LC13) have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the symptom could be evaluated. Also, if symptoms deteriorate after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where the symptom could be evaluated. If a patient has no evaluable visits or does not have baseline data they will be censored at 0 days.

Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement in that symptom from baseline.

8.4.4.3 Calculation or derivation of healthy state utility (EQ-5D-5L)

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied (Oemar and Janssen 2013). In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

8.4.5 Calculation or derivation of pharmacokinetic variables

8.4.5.1 Population pharmacokinetics and exposure-response/safety analysis

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. The results of such an analysis will be reported in a separate report. The PK, pharmacodynamic, 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.

8.4.5.2 Pharmacokinetic non-compartmental analysis

The PK analyses will be performed at AstraZeneca. The actual sampling times will be used in the PK calculations. PK concentration data and summary statistics will be tabulated. Individual and mean blood concentration-time profiles will be generated. PK parameters will be determined using standard non-compartmental methods. The following PK parameters will be determined after the first and steady-state doses: peak and trough concentration (as data allow). Samples below the lower limit of quantification will be treated as missing in the analyses.

8.4.5.3 Immunogenicity analysis

Immunogenicity results will be analyzed descriptively by summarizing the number and percentage of patients who develop detectable ADAs against MEDI4736 and tremelimumab. The immunogenicity titer and presence of neutralizing ADAs will be reported for samples confirmed positive for the presence of ADAs. The effect of immunogenicity on PK, pharmacodynamics, efficacy, and safety will be evaluated, if the data allow.

8.4.6 Calculation or derivation of biomarker variables

Biomarker status, as defined in the secondary objectives, will be assessed according to pre-specified criteria that will be detailed in the SAP.

8.5 Methods for statistical analyses

The formal statistical analysis will be performed to test the main hypotheses:

- H0: No difference between MEDI4736 + tremelimumab and SoC
- H1: Difference between MEDI4736 + tremelimumab and SoC

The co-primary endpoints are OS and PFS in patients with PD-L1 positive$_{25\%}$ tumors (with PFS using BICR assessments per RECIST 1.1) . The study has been sized to characterize the OS and PFS benefits of MEDI4736 + tremelimumab versus SoC in patients with PD-L1 positive$_{25\%}$ and OS benefit of MEDI4736 monotherapy versus SoC in patients with PD-L1 positive$_{25\%}$ tumors.

The primary PFS analysis for superiority will be performed when both of the following conditions have been met:

- Approximately 231 BICR PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (72% maturity) AND
- Minimum 42 weeks follow up from last patient randomized to the study

The final (primary) OS analysis for superiority will be performed when the following conditions have been met:

- Approximately 225 OS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity) AND
- Approximately 225 OS events have occurred across the MEDI4736 monotherapy and SoC treatment groups in patients with PD-L1 positive$_{25\%}$ tumors (70% maturity)

Descriptive statistics will be used for all variables, as appropriate, and will be presented by treatment group. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by frequency counts and percentages for each category. Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment group.

Baseline will be the last assessment of the variable under consideration prior to the intake of the first dose of IP, except for efficacy variables. For efficacy variables, baseline is defined as the last visit prior to randomization.

All data collected will be listed. Efficacy and PRO data will be summarized and analyzed based on the FAS, the PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ analysis sets. PK data will be summarized and analyzed based on the PK Analysis Set. Safety data will be summarized on the Safety Analysis Set.

All outputs will be summarized by treatment group for all randomized patients (ITT) and for all randomized patients in the PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ analysis sets, and where required, for all randomized patients within the PD-L1-low/negative analysis set.

Results of all statistical analysis will be presented using appropriately sized confidence intervals (CIs) and 2-sided p-values, unless otherwise stated.

The following table (Table 10) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint.

Table 10 Pre-planned statistical and sensitivity analyses to be conducted

Endpoints analyzed	Notes
Progression-free survival	Stratified log-rank tests for: Co-Primary analyses using BICR RECIST 1.1 assessments: – MEDI4736 + tremelimumab versus SoC for PD-L1 positive$_{25\%}$ population (stratified only for histology) Secondary analysis using BICR RECIST 1.1 assessments: – MEDI4736 monotherapy versus SoC for PD-L1positive$_{25\%}$ population (stratified only for histology) – MEDI4736 + tremelimumab versus SoC for PD-L1 positive$_{1\%}$ population MEDI4736 + tremelimumab versus SoC (ITT population) – - – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1 positive$_{25\%}$, population (stratified only for histology) Sensitivity analyses using Investigator assessments (RECIST 1.1) Sensitivity analysis based on RECIST 1.1 modified for confirmation of progression using BICR assessments Exploratory analysis using BICR data for irRECIST 1.1

Endpoints analyzed	Notes
Overall survival	Stratified log-rank tests for: Co-primary analysis – : MEDI4736 + tremelimumab versus SoC for PD-L1 positive$_{25\%}$ population (stratified only for histology) – MEDI4736 monotherapy versus SoC for PD-L1 positive$_{25\%}$ population (stratified only for histology) Secondary analysis: – MEDI4736 + tremelimumab versus SoC for PD-L1 positive$_{1\%}$ population MEDI4736 + tremelimumab versus SoC ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1 positive$_{25\%}$, population (stratified only for histology)
Objective response rate	Logistic regression for: Secondary analysis for the ITT, PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$, populations using BICR RECIST 1.1 assessments Sensitivity analysis for the ITT, PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$ populationa using Investigator RECIST 1.1 assessments Exploratory analysis using BICR data for irRECIST 1.1
Duration of response	Analysis methods as described by Ellis et al 2008 for: Secondary analysis using BICR assessments (RECIST 1.1)
Proportion of patients alive and progression free at 12 months	Hazard ratio using the Kaplan Meier estimates of progression free survival at 12 months (following method described by Klein et al 2007)
Time from randomization to second progression	Stratified log-rank test
Time to deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Stratified log-rank test

Multiple testing strategy

In order to strongly control the type I error at 5% 2-sided, a multiple testing procedure (MTP) with gatekeeping strategy will be used across the co-primary endpoints (PFS, OS), analysis populations (ITT, PD-L1 positive$_{25\%}$ and PD-L1 positive$_{1\%}$, and treatment regimens (MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC). If the higher level hypothesis in the MTP is rejected for superiority, the following hypothesis will then be tested as shown in Figure 5.

Hypotheses will be tested using a multiple testing procedure with an alpha-exhaustive recycling strategy (Burman et al 2009). With this approach, hypotheses will be tested in a pre-defined order as outlined in Figure 5. According to alpha (test mass) splitting and alpha recycling, the test mass that becomes available after each rejected hypothesis is recycled to secondary hypotheses not yet rejected. Since OS is tested at multiple timepoints (ie, 2 interim analyses and final analysis), the OS tests that for the same comparison/population (ie, shown in 1 box in the MTP) will be considered as 1 test family. As long as one test in the family can be rejected, the family is rejected thus the assigned total alpha to the family can be recycled to next MTP level. This testing procedure stops when the entire test mass is allocated to non-rejected hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses. Figure 5 shows the multiple testing framework.

The details on the alpha-exhaustive recycling procedure will be provided in the Statistical Analysis Plan.

The co-primary endpoint OS is tested at 2 interim and a final timepoint. The alpha level allocated to OS will be controlled at the interim and primary time points by using the Lan DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, where the alpha level applied at the interim depends upon the proportion of information available. The first OS interim analysis for superiority will occur at the primary PFS analysis, when it is expected that approximately 68% of the target death events may occur.

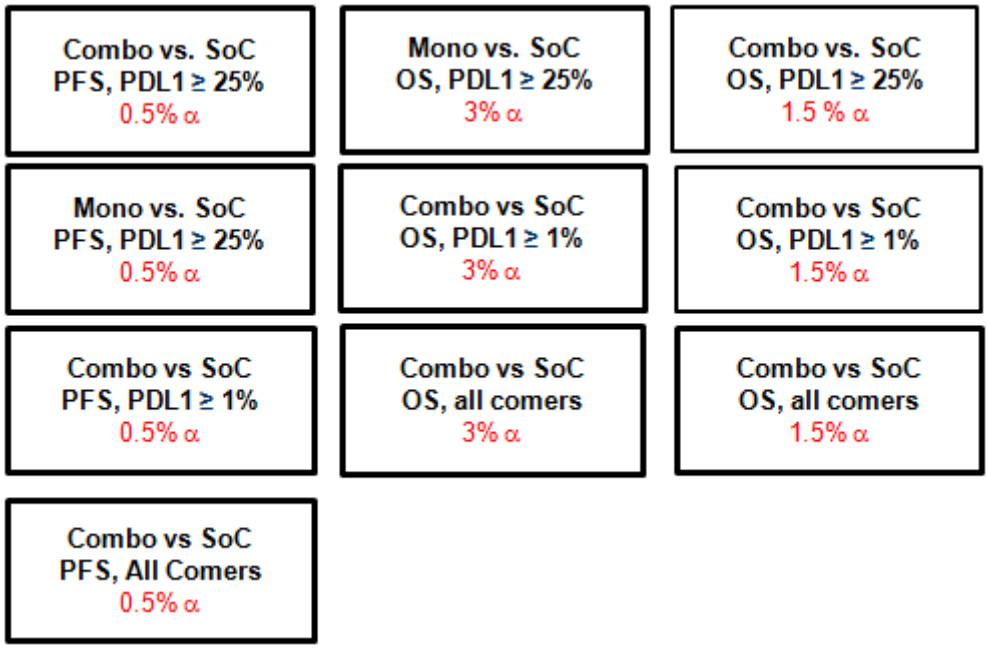
The second OS interim will subsequently be performed at approximately 80% of the target death events, with the primary OS analysis performed when 225 deaths have accumulated. If exactly 68% and 80% of the target events in the PD-L1 positive$_{25\%}$ patients are available at the time of the first and second interims analyses, respectively (ie, 152/225 and 180/225 deaths have occurred), with overall 2-sided alpha levels of 0.015 and 0.03 respectively for the comparisons of MEDi4736 + tremelimumab versus SoC MEDi4736 monotherapy versus SoC , the 2-sided alpha level to be applied for the interim and final analyses would be 0.0023,0.0049 and 0.0132 for the comparison of MEDi4736 + tremelimumab versus SoC and 0.0062,0.0113 and 0.0258 for the comparison of MEDi4736 monotherapy versus SoC.

If the interim or final analyses indicate superiority in OS for either monotherapy or combination therapy in PD-L1 positive$_{25\%}$, then subsequent analyses of secondary OS endpoints will be performed in accordance with the hierarchical testing strategy. A separate Lan DeMets (O'Brien Fleming) spending function will be used to determine the alpha levels at the interim and final analyses for testing the PD-L1 positive$_{1\%}$ and the all-comers hypotheses.

If the interim results do not meet the criterion of stopping for superiority for a given hypothesis, then follow-up will continue until the final target number of OS events for that comparison has been observed, following which the hypothesis will be re-tested. If the hypothesis is then rejected, subsequent testing will continue hierarchically. The above testing

procedure will ensure strong control of the family-wise error rate (Glimm et al, 2010). Additional details of the multiple testing procedure will be provided in the statsistical analysis plan.

Figure 5 Multiple testing procedures for controlling the type 1 error rate



Combo MEDI4736 + tremelimumab combination therapy; Mono MEDI4736 monotherapy; SoC Standard of care.

8.5.1 Analysis of the co-primary endpoints

8.5.1.1 Progression-free survival

The co-primary PFS analyses will be based on the programmatically derived RECIST 1.1 using the BICR tumor assessments. The co-primary analysis performed in the PD-L1 positive$_{25\%}$, population uses a stratified log-rank test adjusting for histology (squamous versus non-squamous) only. The effect of MEDI4736 + tremelimumab versus SoC treatment will be estimated by the HR together with its corresponding 99.5% CI and p-value.

The HR and its CI can be estimated from the Cox proportional hazards model (Cox 1972).

All of the secondary analyses will be performed using the same methodology as for the primary analyses described above, except in cases where the secondary analyses are

performed on the FAS or PD-L1 positive$_{1\%}$ population, in which case the stratification will also adjust for PD-L1 status (≥25%, <25%).

Kaplan-Meier plots of PFS will be presented by treatment group, and by treatment group and PD-L1 tumor status subgroup, where appropriate. Summaries of the number and percentage of patients experiencing a PFS event and the type of event (RECIST 1.1 or death) will be provided along with median PFS for each treatment.

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time-points. The midpoint between the time of progression and the previous evaluable RECIST assessment will be analyzed using a log-rank test. For patients whose death was treated as PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust even in highly asymmetric assessment schedules (Sun and Chen 2010).

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the PFS analysis is reversed.

Ascertainment bias will be assessed by analyzing the site Investigator data. The stratified log rank test will be repeated on the programmatically derived PFS using the site Investigator data based upon RECIST 1.1. The HR and CI will be presented.

If there is an important discrepancy between the primary analysis using the BICR assessments and this sensitivity analysis using Investigator assessments, then the proportion of patients with site but no central confirmation of progression will be summarized; such patients have the potential to introduce bias in the central review due to informative censoring. An approach that imputes an event at the next visit in the central review analysis may help inform the most likely HR value (Fehrenbacher et al 2016, Fleischer et al 2011), but only if an important discrepancy exists.

An exploratory analysis of PFS using BICR assessment based on RECIST 1.1 modified for confirmation of progression as well as PFS based on BICR assessments according to irRECIST 1.1 criteria will be performed. The stratified log-rank test used for the primary analysis of PFS will be repeated.

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1) between MEDI4736 + tremelimumab versus SoC in the following subgroups PD-L1 positive$_{25\%}$, and PD-L1 positive$_{1\%}$ analysis sets and the of the FAS as deemed appropriate(but not limited to):

- Sex (male versus female)

- Age at randomization (<65 versus ≥65 years of age)
- PD-L1 status (≥25%, <25%)
- Histology (squamous versus non-squamous)
- Smoking (smoker versus non-smoker [never smoker])
- Race (Asian versus non-Asian)
- PD-L1 using cutpoints of 1%, 10% and 50% tumour expression (<1% versus ≥1%, <10% versus ≥10% and <50% versus ≥50%)

Other baseline variables may also be assessed if there is clinical justification or an imbalance is observed between the treatment groups. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors.

No adjustment to the significance level for testing of the subgroup and sensitivity analyses will be made since all these analyses will be considered supportive of the analysis of PFS.

Unless there is a marked difference between the results of the statistical analyses of the PFS from the BICR data and that of the site Investigator tumor data, these subgroup analyses will only be performed on the PFS endpoint using the BICR data.

Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. An initial model will be constructed, containing treatment and the stratification factors alone, to ensure that any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

Interactions between treatment and stratification factor will also be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

Additionally, for each subgroup, the HR (for the treatment comparisons of interest) and 95% CI will be calculated from a single model that contains treatment and subgroup factor. These will be presented on a forest plot including the HR and 95% CI.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events in a subgroup), the relationship between that subgroup and PFS will not be formally analyzed. In this case, only descriptive summaries will be provided.

8.5.1.2 Overall survival

Co-primary OS in the PD-L1 positive$_{25\%}$ population will be analyzed using a stratified log-rank tests, using the same methodology as described for the co-primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC and MEDI4736 monotherapy versus SoC will be estimated by the HR together with its corresponding two-sided CI (98.5% for

MEDI4736 + tremelimumab versus SoC and 97% for MEDI4736 monotherapy versus SoC, adjusted for two interim analyses) and p-value. Kaplan-Meier plots will be presented by treatment group. 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 for each treatment.

The boundaries (ie, adjusted alpha levels) for the treatment comparison at the interim and final analyses for OS will be derived based upon the exact number of OS events using the Lan and DeMets approach that approximates the O'Brien Fleming spending function (see Section 8.28.2).

Subgroup analyses will be conducted, as deemed appropriate, in the PD-L1 positive$_{25\%}$ analysis set comparing OS between MEDI4736 + tremelimumab versus SoC and MEDI4736 monotherapy versus SoC and in the PD-L1 positive$_{1\%}$ analysis set comparing OS between MEDI4736 + tremelimumab versus SoC using the same subgroups as specified for PFS. Similar subgroup analyses will be performed in the FAS for the comparison between MEDI4736 + tremelimumab versus SoC as deemed appropriate.

Impact of switching (crossover outside of this study) to immunotherapies (or other potentially active investigational agents) on OS analyses

Exploratory analyses of OS adjusting for the impact of subsequent immunotherapy or other investigational treatment may be performed, if a sufficient proportion of patients switch. Methods such as Rank Preserving Structural Failure Time (Robins and Tsiatis 1991), Inverse Probability of Censoring Weighting (Robins 1993), and other methods in development will be explored. The decision to adjust and the final choice of methods will be based on a blinded review of the data and the plausibility of the underlying assumptions. Baseline and time-dependent characteristics will be explored, and summaries of baseline characteristics will be summarized by treatment group, splitting between those that have and have not switched at the time of the analyses. Further detail will be provided in the Payer Analysis Plan.

8.5.2 Objective response rate

The ORR will be based on the programmatically derived RECIST 1.1 using the BICR tumor data. The ORR will be compared between MEDI4736 + tremelimumab versus SoC using logistic regression models adjusting for the same factors as the primary endpoint (PD-L1 tumor expression and histology, as appropriate). The results of the analysis will be presented in terms of an odds ratio together with its associated profile likelihood CI and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model). This analysis will be performed in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations.. The analysis of the PD-L1 subgroup patients will be performed using a logistic regression model adjusting for only histology.

This analysis of ORR will be repeated using the results of the programmatically derived ORR using the site Investigator tumor data based upon RECIST 1.1 as a sensitivity analysis to confirm the results of the primary analysis derived from the eCRFs.

ORR by irRECIST 1.1 criteria using BICR assessments will also be reported in the ITT population.

Summaries will be produced that present the number and percentage of patients with a tumor response (CR/PR). Overall visit response data will be listed for all patients. For each treatment group, best overall response (BoR) will be summarized by n (%) for each category (CR, PR, SD, PD, and NE). No formal statistical analyses are planned for BoR.

8.5.3 Duration of response

In order to analyze the DoR between MEDI4736 + tremelimumab and SoC, the expected duration of response (EDoR) will be derived for each treatment group (Ellis et al 2008) using the BICR tumor data. The EDoR is the product of the proportion of patients responding to treatment and the mean DoR in responding patients and provides an estimate based on all randomized patients. Treatments will be compared by calculating the ratio of EDoRs, using an appropriate probability distribution (to be specified in the SAP) for DoR in responding patients. Additionally, descriptive data will be provided for the DoR in responding patients, including the associated Kaplan-Meier curves (without any formal comparison of treatment groups or p-value attached). This analysis will be performed in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations.

8.5.4 Proportion of patients alive and progression free at 12 months

The APF12 will be summarized (using the Kaplan-Meier curve) and presented by treatment group. APF12 will be compared between MEDI4736 + tremelimumab and SoC by using the Kaplan-Meier estimator of PFS at 12 months for each treatment to obtain the HR. The HR and CI will be presented using the following approach (Klein et al 2007).

- The *HR(group1:group2)* is estimated as $\frac{\ln \hat{S}_1(t)}{\ln \hat{S}_2(t)}$
- The variance for *ln(HR)* is estimated as $\frac{\hat{\sigma}_1(t)^2}{\ln^2 S_1(t)} + \frac{\hat{\sigma}_2(t)^2}{\ln^2 S_2(t)}$

where $\hat{\sigma}_i(t)^2 = \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$ is the variance for ln{S(t) derived from greenwood's formula for the variance of *S(t)* and can be estimated from standard software packages, where d_i and n_i refer to the number of events and patients at risk for each risk set.

The *ln(HR)* and its variance in each strata will be estimated and combined by weighting inversely proportionately according to each within stratum variance (Whitehead and Whitehead 1991).

This analysis will be performed in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations.

8.5.5 Time from randomization to second progression

PFS2 is defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death. PFS2 in the ITT population will be analyzed using a stratified log-rank tests, using the same methodology as described for the primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

For supportive purposes, the time to the start of subsequent therapy will be analyzed using the same methodology and model. The HR for the treatment effect together with its 95% CI will be presented. In addition, a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment group and the time between progression and starting subsequent therapy will be assessed. No multiplicity adjustment will be applied as these are viewed as supportive endpoints.

A summary table of first subsequent therapies by treatment group will be provided, as well as response to first subsequent therapy by treatment group.

This analysis will be performed in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations.

8.5.6 Patient reported outcomes

8.5.6.1 EORTC QLQ-C30

Time to deterioration in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 3 symptom scales and the 5 individual symptom items will be produced. Similarly, a summary of QoL/function improvement rate for each of the 5 function scales (physical, role, emotional, cognitive, and social) and global health status/QoL will be produced.

Summaries of original and change from baseline values of each symptom scale/item, the global HRQoL score and each functional domain will be reported by visit for each treatment

group. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration as defined in Section 8.4.4.1 will also be produced for each treatment group.

8.5.6.2 EORTC QLQ-LC13

Time to deterioration in the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 6 individual symptom items will be produced.

Summaries of original and change from baseline values of each symptom (dyspnea, cough, hemoptysis, chest pain, arm/shoulder pain, other pain) and each treatment-related side effect (sore mouth, dysphagia, peripheral neuropathy and alopecia) will be reported by visit for each treatment group. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal symptom item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration as defined in Section 8.4.4.2) will also be produced for each treatment group.

For PRO symptoms and HRQoL endpoints, the overall type I error (5% 2-sided) will be controlled across the 5 primary PRO measures of cough, dyspnea, and pain as assessed by the EORTC QLQ-LC13 and fatigue and appetite loss as assessed by the EORTC QLQ-C30 using the Bonferroni-Holm procedure (Holm 1979).

The physical functioning and overall health status domains of the EORTC QLQ-C30 are furthermore pre-specified endpoints of interest.

8.5.6.3 PRO-CTCAE

PRO-CTCAE data will be presented using summaries and descriptive statistics based on the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations. Further details will be provided in the SAP.

8.5.6.4 Patients' Global Impression of Change

PGIC data will be presented using summaries and descriptive statistics based on the PD-L1 positive$_{25\%}$, PD-L1 positive$_{1\%}$ and ITT populations. Further details will be provided in the SAP.

8.5.6.5 EQ-5D-5L

Descriptive statistics, graphs, and listings will be reported for health state utility values and the visual analogue scale by visit, as well as the change in these scores from baseline. To support future economic evaluations, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post-progression, and will be outlined in the payer analysis plan.

8.5.7 Healthcare resource use

An exploratory health economic analysis of hospital episodes including type of contact (hospitalization, outpatient, or day case), reason, length of stay by ward type (including intensive care unit), procedures, and tests may be undertaken to examine the impact of disease and treatment on resource use to primarily support the economic evaluation of MEDI4736 + tremelimumab in comparison to SoC, and will be outlined in the payer analysis plan. This would include providing descriptive statistics as appropriate, including means, median, and ranges.

8.5.8 Safety data

Safety and tolerability data will be presented by treatment group using the safety population.

Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of MedDRA preferred terms and CTCAE grade) will be listed individually by patient. The number of patients experiencing each AE will be summarized by treatment group and CTCAE grade. Additionally, data presentations of the rate of AEs per person-years at risk may be produced. Any safety summaries examining retreatment with MEDI4736 + tremelimumab will be produced separately, if required.

Other safety data will be assessed in terms of physical examination, clinical chemistry, hematology, vital signs, and ECGs. Exposure to MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC will be summarized. Time on study, MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC, dose delays/interruptions and dose reductions will also be summarized. At the end of the study, appropriate summaries of all safety data will be produced, as defined in the SAP.

8.5.9 Pharmacokinetic data

PK concentration data will be listed for each patient and each dosing day, and a summary provided for all evaluable patients in the PK analysis population.

8.5.10 Immunogenicity analysis

Immunogenicity results will be listed by patient and a summary will be provided of the number and percentage of patients who develop detectable anti-MEDI4736 and anti-tremelimumab antibodies. The immunogenicity titer and neutralizing ADA data will be listed for samples confirmed positive for the presence of anti-MEDI4736 and anti-tremelimumab antibodies.

The effect of immunogenicity on PK, pharmacodynamics, efficacy and safety will be evaluated if data allow.

8.5.11 Pharmacokinetic/Pharmacodynamic relationships

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modeling approach.

8.5.12 Biomarker data

The relationship of PD-L1 expression and if appropriate, other exploratory biomarkers to clinical outcomes (including but not restricted to) PFS and ORR may be presented.

PD-L1 expression determined by IHC 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.

8.5.13 Interim analysis

Interim monitoring for safety will be conducted by the IDMC. Details of the plan and communication process will be provided in the statistical analysis plan and the IDMC charter.

In addition, two OS interim analyses will be performed for superiority; the first one at the time of the primary PFS analysis and the second one when approximately 80% of the final number of deaths has been reached. These analyses will be performed by an IDMC.

The Lan DeMets spending function that approximates an O'Brien Fleming approach will be used to account for multiplicity introduced by including the 2 interim analyses for superiority (Lan and DeMets 1983).

The criterion for superiority in the OS co-primary endpoint is a statistically significant improvement in OS at the interim analysis. The alpha level that is to be spent at the interim and final analyses for the OS analyses in the PD-L1 positive$_{25\%}$, population will be calculated using the Lan DeMets spending function separately. If exactly 68% and 80% of the target events in the PD-L1 positive$_{25\%}$ patients are available at the time of the first and second interims analyses, respectively (ie, 152/225 and 180/225 deaths have occurred), with overall 2-sided alpha levels of 0.015 and 0.03 respectively for the comparisons of MEDi4736 + tremelimumab versus SoC MEDi4736 monotherapy versus SoC , the 2-sided alpha level to be applied for the interim and final analyses would be 0.0023,0.0049 and 0.0132 for the comparison of MEDi4736 + tremelimumab versus SoC and 0.0062,0.0113 and 0.0258 for the comparison of MEDi4736 monotherapy versus SoC. .

If the interim analyses indicate superiority in the PD-L1 positive$_{25\%}$ population, then subsequent analyses of the further secondary endpoints will be performed in accordance with the hierarchical multiple testing strategy.

If the interim results do not meet the criterion of stopping for superiority in the PD-L1 positive$_{25\%}$ population, then follow-up will continue until at least 225 deaths have occurred in the PD-L1 positive$_{25\%}$ population. OS will be retested in the PD-L1 positive$_{25\%}$ population at the final analysis. Similarly if the criteria for statistical significance at the interim is met for PD-L1 positive$_{25\%}$ but not PD-L1 positive$_{1\%}$, then follow-up will continue until 337 events have been reached in this population after which the OS will be retested.

9. STUDY AND DATA MANAGEMENT BY ASTRAZENECA

9.1 Training of study site personnel

Before the first patient is enrolled in the study, an AstraZeneca representative will review and discuss the requirements of the clinical study protocol and related documents with the investigational staff and train them in any study-specific procedures and IVRS/IWRS, WBDC, and any electronic PRO systems to be utilized.

The Principal Investigator will ensure that appropriate training relevant to the study is given to all of these staff and that any new information relevant to the performance of this study is forwarded to the staff involved.

The Principal Investigator will maintain a record of all individuals involved in the study (medical, nursing, and other staff).

9.2 Monitoring of the study

During the study, an AstraZeneca representative will have regular contacts with the study site, including visits to

- Provide information and support to the Investigator(s)
- Confirm that facilities remain acceptable
- Confirm that the investigational team is adhering to the protocol, data are being accurately and timely recorded in the eCRFs, biological samples are handled in accordance with the Laboratory Manual, and study drug accountability checks are being performed
- Perform source data verification (a comparison of the data in the eCRFs with the patient's medical records at the hospital or practice and other records relevant to the study), including verification of informed consent of participating patients. This will require direct access to all original records for each patient (eg, clinic charts).
- Ensure that withdrawal of informed consent for the use of the patient's biological samples is reported, biological samples are identified and disposed of or destroyed accordingly, and the action is documented and reported to the patient.

The AstraZeneca representative will be available between visits if the Investigators or other staff at the centers need information and advice about the study conduct.

9.2.1 Source data

Refer to the CSA for the location of source data.

9.2.2 Direct access to source data in Japan

The Head of the study site and the Principal Investigator/Investigator will cooperate for monitoring and audit by AstraZeneca and accept inspection by the IRB or regulatory authorities. All study documents such as raw data will be open for direct access to source data at the request of the monitor and the auditor of AstraZeneca, the IRB, or regulatory authorities.

The monitor will verify data from the eCRFs against source data before the Principal Investigator signs the eCRFs to ensure accuracy and completeness of documentation and ensure that the Principal Investigator has submitted the eCRFs to AstraZeneca.

9.2.3 Study agreements

The Principal Investigator at each center should comply with all the terms, conditions, and obligations of the CSA for this study. In the event of any inconsistency between this clinical study protocol and the CSA, the terms of clinical study protocol shall prevail with respect to the conduct of the study and the treatment of patients. In all other respects not relating to study conduct or treatment of patients, the terms of the CSA shall prevail.

Agreements between AstraZeneca and the Principal Investigator should be in place before any study-related procedures can take place or before any patients are enrolled.

9.2.4 Archiving of study documents

The Investigator will follow the principles outlined in the CSA.

9.3 Study timetable and end of study

The end of the study is defined as the "last visit of the last patient undergoing the study." The Investigator will be notified by AstraZeneca when recruitment is complete.

The study is expected to start in and end by .

The study may be terminated at individual centers if the study procedures are not being performed according to Good Clinical Practice (GCP) or if recruitment is slow. AstraZeneca may also terminate the entire study prematurely if concerns for safety arise within this study or in any other study involving MEDI4736.

9.4 Data management by AstraZeneca or delegate

Data management will be performed by a chosen vendor according to the Data Management Plan. AEs and medical/surgical history will be classified according to the terminology of the latest version of MedDRA. Medications will be classified according to the AstraZeneca Drug Dictionary. Classification coding will be performed by the chosen vendor.

The data collected through third party sources will be obtained and reconciled against study data.

Data queries will be raised for inconsistent, impossible, or missing data. All entries to the study database will be available in an audit trail.

The data will be validated as defined in the Data Management Plan. Quality control procedures will be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly. The Data Management Plan will also clarify the roles and responsibilities of the various functions and personnel involved in the data management process.

When all data have been coded, validated, signed, and locked, a clean file will be declared. Any treatment-revealing data may be added thereafter, and the final database will be locked.

Serious adverse event reconciliation

SAE reconciliation reports are produced and reconciled with the Patient Safety database and/or the investigational site.

Data management of genotype data

Any genotype data generated in this study will be stored in the AstraZeneca genotyping database or other appropriate secure system within AstraZeneca and/or a third party contracted to work with AstraZeneca to analyze samples. The results from this genetic research may be reported in the CSR for the main study or in a separate report, as appropriate.

Some or all of the clinical datasets from the main study may be merged with the genetic data in a suitable secure environment separate from the clinical database.

Data associated with human biological samples

Data associated with human biological samples will be transferred from laboratories internal or external to AstraZeneca.

10. ETHICAL AND REGULATORY REQUIREMENTS

10.1 Ethical conduct of the study

The study will be performed in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation (ICH)/GCP, applicable regulatory requirements, and the AstraZeneca policy on Bioethics and Human Biological Samples. The applicable regulatory requirements in Japan are "Good Clinical Practice for Trials on Drugs" (Ministry of Health, Labor, and Welfare [MHLW] Ordinance No. 28, 27 March 1997, partially revised by MHLW Ordinance and their related notifications).

10.2 Patient data protection

The ICF will incorporate wording that complies with relevant data protection and privacy legislation. In some cases, such wording will be in a separate accompanying document.

AstraZeneca will not provide individual genotype results to patients, their family members, their general physician, any insurance company, any employer, or any other third party, unless required to do so by law**.**

Precautions are taken to preserve confidentiality and prevent genetic data from being linked to the identity of the patient. In exceptional circumstances, however, certain individuals might see both the genetic data and the personal identifiers of a patient. For example, in the case of a medical emergency, an AstraZeneca Physician or an Investigator might know a patient's identity and might also have access to his or her genetic data. Also, regulatory authorities may require access to the relevant files. Even so, the patient's medical information and the genetic files would remain physically separate.

10.3 Ethics and regulatory review

An EC/IRB should approve the final study protocol, including the final version of the ICF and any other written information and/or materials to be provided to the patients. The Investigator will ensure the distribution of these documents to the applicable EC/IRB and to the study site staff.

The opinion of the EC/IRB should be given in writing. The Investigator should submit the written approval to AstraZeneca before enrollment of any patient into the study.

The EC/IRB should approve all advertising used to recruit patients for the study.

AstraZeneca should approve any modifications to the ICF that are needed to meet local requirements.

If required by local regulations, the protocol should be re-approved by the EC/IRB annually.

Before enrollment of any patient into the study, the final study protocol, including the final version of the ICF, should be approved by the national regulatory authority or a notification to the national regulatory authority should be approved, according to local regulations.

AstraZeneca will handle the distribution of these documents to the national regulatory authorities.

AstraZeneca will provide regulatory authorities, ECs/IRBs, and Principal Investigators safety updates or reports according to local requirements.

Each Principal Investigator is responsible for providing the EC/IRB with reports of any serious and unexpected adverse drug reactions from any other study conducted with the IP. AstraZeneca will provide this information to the Principal Investigator so that he/she can meet these reporting requirements.

10.4 Informed consent

The Principal Investigator(s) at each center will:

- Ensure that each patient is given full and adequate oral and written information about the nature, purpose, possible risk, and benefit of the study
- Ensure that each patient is notified that he or she is free to discontinue from the study at any time
- Ensure that each patient is given the opportunity to ask questions and allowed time to consider the information provided
- Ensure that each patient provides a signed and dated informed consent before conducting any procedure specifically for the study
- Ensure that the original, signed ICF(s) is/are stored in the Investigator's Study File
- Ensure that a copy of the signed ICF is given to the patient
- Ensure that any incentives for patients who participate in the study as well as any provisions for patients harmed as a consequence of study participation are described in the ICF that is approved by an EC/IRB.

For sites in Japan only

If any new information on the study medication becomes available that may influence the decision of the patient to continue the study, the Investigator should inform the patient of such information immediately, record this in a written form, and confirm with the patient if he or she wishes to continue the participation in the study. In addition, if the Investigator deems it necessary to revise the ICF, he or she should revise it immediately (refer to Section 10.5).

The Investigator should re-explain to the patients using the updated ICF even if the patients have already been informed of the new information verbally. Written informed consent to continue participation in the study should be provided separately.

10.5 Changes to the protocol and informed consent form

For sites outside Japan

Study procedures will not be changed without the mutual agreement of the Principal Investigator and AstraZeneca.

If there are any substantial changes to the study protocol, then these changes will be documented in a study protocol amendment and, where required, in a new version of the study protocol (revised clinical study protocol).

The amendment is to be approved by the relevant EC/IRB and, if applicable, the national regulatory authority, before implementation. Local requirements are to be followed for revised protocols.

AstraZeneca will distribute any subsequent amendments and new versions of the protocol to each Principal Investigator. For distribution to EC/IRB, see Section 10.3.

If a protocol amendment requires a change to a center's ICF, AstraZeneca and the center's EC/IRB are to approve the revised ICF before the revised form is used.

If required by local regulations, any administrative change will be communicated to or approved by each EC/IRB.

10.6 Audits and inspections

Authorized representatives of AstraZeneca, a regulatory authority, or an EC/IRB may perform audits or inspections at the center, including source data verification. The purpose of an audit or inspection is to systematically and independently examine all study-related activities and documents, to determine whether these activities were conducted, and to determine if data were recorded, analyzed, and accurately reported according to the protocol, GCPs, ICH guidelines, and any applicable regulatory requirements. The Investigator will contact AstraZeneca immediately if contacted by a regulatory agency about an inspection at the center.

11. LIST OF REFERENCES

Aaronson et al 1993
Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85(5):365-76.

Antonia et al 2014a
Antonia S, Goldberg S, Balmanoukian A, Narwal R, Robbins P, D'Angelo G, et al. A Phase 1 open-label study to evaluate the safety and tolerability of MEDI4736, an anti-programmed cell death-ligand 1 (PD-L1) antibody, in combination with tremelimumab in patients with advanced non-small cell lung cancer (NSCLC). Poster presented at European Society of Medical Oncology (ESMO) Meeting; 2014 Sep 26-30; Madrid, Spain.

Antonia et al 2014b
Antonia S, Ou SI, Khleif SN, Brahmer J, Blake-Haskins A, Robbins PB, et al. Clinical activity and safety of MEDI4736, an anti-programmed cell death ligand-1 (PD-L1) antibody in patients with NSCLC. Poster presented at the European Society for Medical Oncology (ESMO) Meeting; 2014 Sep 26-30; Madrid, Spain.

Bergman et al 1994
Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M; EORTC Study Group on Quality of Life. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. Eur J Cancer 1994;30A(5):635-42.

Berry et al 1991
Berry G, Kitchin RM, Mock PA. A comparison of 2 simple hazard ratio estimators based on the logrank test. Stat Med 1991;10(5):749-55.

Bonomi 2010
Bonomi PD. Implications of key trials in advanced nonsmall cell lung cancer. Cancer 2010;116(5):1155-64.

Borghaei et al 2015
Borghaei H, Paz-Ares L, Horn L, Spigel D, Steins M, Ready N, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N Engl J Med 2015;373:1627-39.

Brahmer et al 2012
Brahmer JR, Tykodi SS, Chow LQM, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. N Engl J Med 2012;366(26):2455-65.

Brahmer et al 2014
Brahmer JR, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA et al. Nivolumab (anti-PD-1, BMS-936558, ONO-4538) in patients (pts) with advanced non-small-cell lung cancer (NSCLC): Survival and clinical activity by subgroup analysis. Journal of Clinical Oncology, 2014 ASCO Annual Meeting Abstracts; 32(15) (2014 May 20 Supplement): 8112.

Brahmer et al 2015
Brahmer JR, Reckamp K, Baas P, Crino L, Eberhardt W, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med. 2015;373:123-35.

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009;28:739-61.

Ciuleanu et al 2009
Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomized, double-blind, phase 3 study. Lancet 2009;374(9699):1432-40.

Cox 1972
Cox DR. Regression models and life-tables. J Royal Stat Society 1972;Series B 34(2):187-220.

D'Addario et al 2010
D'Addario G, Fruh M, Reck M, Baumann P, Klepetko W, Felip E, et al. Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2010;21(Suppl 5):116-9.

Drake et al 2013
Drake CG, McDermott DF, Sznol M, Choueiri TK, Kluger HM, Powderly JD et al. Survival, safety, and response duration results of nivolumab (Anti-PD-1; BMS-936558; ONO-4538) in a phase I trial in patients with previously treated metastatic renal cell carcinoma (mRCC): Long-term patient follow-up. J Clin Oncol 31, 2013 (suppl; abstr 4514).

Dunn et al 2004
Dunn GP, Old LJ, Schreiber RD. The three Es of cancer immunoediting. Annu Rev Immunol 2004;22:329-60.

Ellis et al 2008
Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008;29(4):456-65.

EuroQol Group 1990
EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208.

EuroQol Group 2013
EuroQol Group. EQ-5D-5L user guide: basic information on how to use the EQ-5D-5L instrument, version 2.0, October 2013. Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/UserGuide_EQ-5D-5L_v2.0_October_2013.pdf. Accessed 21 November 2013.

Fairman et al 2014
Fairman D, Narwal R, Liang M, Robbins PB, Schneider A, Chavez C, et al. Pharmacokinetics of MEDI4736, a fully human anti-PDL1 monoclonal antibody, in patients with advanced solid tumours. Journal of Clinical Oncology, 2014 ASCO Annual Meeting Abstracts;32(5s): (suppl; abstr 2602).

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer: 2001.

FDA Guidance 2011
Food and Drug Administration. Guidance for Industry: Clinical trial endpoints for the approval of non-small cell lung cancer drugs and biologics. June 2011. Available from: URL: http://www.fda.gov/downloads/Drugs/Guidances/UCM259421.pdf. Accessed 21 May 2014.

Fehrenbacher et al 2016
Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, Phase 2, randomised, controlled trial. Lancet 2016; published online 09 March 2016.

Fleischer et al 2011
Fleischer F, Gaschler-Markefski B, Bluhmki E. How is retrospective independent review influenced by investigator-introduced informative censoring: a quantitative approach. Stat Med 2011;30(29):3373-86.

Forde et al 2014
Forde PM, Kelly RJ, Brahmer JR. New strategies in lung cancer: translating immunotherapy into clinical practice. Clin Cancer Res 2014;20(5):1067-73.

Gail and Simon 1985
Gail M, Simon R. Tests for qualitative interactions between treatment effects and patient subsets. Biometrics 1985;41(2):361–72.

Glimm et al 2010
Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Stat Med 2010;29:219-28.

GLOBOCAN 2012
GLOBOCAN. Lung cancer, estimated incidence, mortality and prevalence worldwide in 2012. International Agency for Research on Cancer, World Health Organization; Lyon, 2012. Available at http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx. Accessed on 16 April 2015.

Gong et al 2009
Gong AY, Zhou R, Hu G, Li X, Splinter PL, O'Hara SP, et al. MicroRNA-513 regulates B7-H1 translation and is involved in IFN-gamma-induced B7-H1 expression in cholangiocytes. J Immunol 2009;182(3):1325-33.

Hellmann et al 2016
Hellmann MD, Rizvi NA, Goldman JW, Gettinger SN, Borghaei H, Brahmer JR, et al. Nivolumab plus ipilimumab as first-line treatment for advanced non-small-cell lung cancer (CheckMate 012): results of an open-label, phase 1, multicohort study. Lancet Oncol 2016; [Epub ahead of print].

Herbst et al 2013
Herbst RS, Gordon MS, Fine GD, Sosman JA, Soria JC, Hamid O, et al. A study of MPDL3280A, an engineered PD-L1 antibody in patients with locally advanced or metastatic tumours [abstract]. J Clin Oncol 2013;31(Suppl 15):Abstract 3000.

Herdman et al 2011
Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727-36.

Hirano et al 2005
Hirano F, Kaneko K, Tamura H, Dong H, Wang S, Ichikawa M, et al. Blockade of B7-H1 and PD-1 by monoclonal antibodies potentiates cancer therapeutic immunity. Cancer Res 2005;65(3):1089-96.

Hodi et al 2010
Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanan JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma [published erratum appears in N Engl J Med 2010;363(13):1290]. N Engl J Med 2010;363(8):711-23.

Hodi et al 2014
Hodi FS, Sznol M, Kluger HM, McDermott DF, Carvajal RD, Lawrence DP et al. Long-term survival of ipilimumab-naive patients (pts) with advanced melanoma (MEL) treated with nivolumab (anti-PD-1, BMS-936558, ONO-4538) in a phase I trial. J Clin Oncol, 2014 ASCO Annual Meeting Abstracts; 32(15) (2014 May 20 Supplement): 9002.

Holm 1979
Holm S. A simple sequentially rejective multiple test procedure. Scand J Statistics 1979;6:65-70.

Howlander et al 2014
Howlander N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, et al. SEER cancer statistics review, 1975-2011. Bethesda (MD): National Cancer Institute,. Available from: URL:http://seer.cancer.gov/csr/1975_2011/ based on November 2013 SEER data submission (posted to the SEER web site, April 2014).

IASLC Staging Manual in Thoracic Oncology
Goldstraw P, ed. Staging manual in thoracic oncology. 7th ed. IASLC 2010. (Available on request)

Iwai et al 2002
Iwai Y, Ishida M, Tanaka Y, Okazaki T, Honjo T, Minato N. Involvement of PD-L1 on tumour cells in the escape from host immune system and tumour immunotherapy by PD-L1 blockade. Proc Natl Acad Sci USA 2002;99(19):12293-7.

Janssen et al 2008a
Janssen MF, Birnie E, Haagsma JA, Bonsel GJ. Comparing the standard EQ-5D three-level system with a five-level version. Value Health 2008;11(2):275-84.

Janssen et al 2008b
Janssen MF, Birnie E, Bonsel GJ. Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. Qual Life Res 2008;17(3):463-73.

Kirkwood et al 2010
Kirkwood JM, Lorigan P, Hersey P, Hauschild A, Robert C, McDermott D, et al. Phase II trial of tremelimumab (CP-675,206) in patients with advanced refractory or relapsed melanoma. Clin Cancer Res 2010;16(3):1042-8.

Kitano et al 2014
Kitano S, Postow MA, ZieglerCG, Kuk D, Panageas KS, Cortez C, et al. Computational algorithm-driven evaluation of monocytic myeloid-derived suppressor cell frequency for prediction of clinical outcomes. Cancer Immunol Res 2014;2(8):812-21.

Klein et al 2007
Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007;26(24):4505-19.

Korn et al 2008
Korn EL, Liu PY, Lee SJ, Chapman JA, Niedzwiecki D, Suman VJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. J Clin Oncol 2008;26(4):527-34.

Lan and DeMets 1983
Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983; 70 (3):659-663.

Litwin et al 1998
Litwin MS, Lubeck DP, Henning JM, Carroll PR. Differences in urologist and patient assessments of health related quality of life in men with prostate cancer: results of the CaPSURE database. J Urol 1998;159:1988-92.

Meyer et al 2014
Meyer C, Cagnon L, Costa-Nunes Cm, Baumgaaertner P, Montandon N, Leyvraz L, et al. Frequencies of circulating MDSC correlate with clinical outcome of melanoma patients treated with ipilimumab. Cancer Immunol Immunother 2014;63(3):247-57.

Mu et al 2011
Mu CY, Huang JA, Chen Y, Chen C, Zhang XG. High expression of PD-L1 in lung cancer may contribute to poor prognosis and tumour cells immune escape through suppressing tumour infiltrating dendritic cells maturation. Med Oncol 2011;28(3):682-8.

NCCN 2014
National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology (NCCN Guidelines®). Non-Small Cell Lung Cancer. Version 4.2014. www.nccn.org.

Oemar and Janssen 2013
Oemar M, Oppe M. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument V5.0 (October 2013). Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

Okudaira et al 2009
Okudaira K, Hokari R, Tsuzuki Y, Okada Y, Komoto S, Watanabe C, et al. Blockade of B7-H1 or B7-DC induces an anti-tumour effect in a mouse pancreatic cancer model. Int J Oncol 2009;35(4):741-9.

Osoba et al 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16(1):139-44.

Pardoll 2012
Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12(4):252-64.

Paz-Ares et al 2013
Paz-Ares LG, de Marinis F, Dediu M, Thomas M, Pujol JL, Bidoli P, et al. PARAMOUNT: Final overall survival results of the phase III study of maintenance pemetrexed versus placebo

immediately after induction treatment with pemetrexed plus cisplatin for advanced nonsquamous non-small-cell lung cancer. J Clin Oncol 2013;31(23):2895-902.

Pazdur 2008
Pazdur R. Endpoints for assessing drug activity in clinical trials. Oncologist 2008;13(Suppl 2):19-21.

Peggs et al 2009
Peggs KS, Quezada SA, Allison JP. Cancer immunotherapy: co-stimulatory agonists and coinhibitory antagonists. Clin Exp Immunol 2009;157:9-19.

Pickard et al 2007
Pickard AS, De Leon MC, Kohlmann T, Cella D, Rosenbloom S. Psychometric comparison of the standard EQ-5D to a 5 level version in cancer patients. Med Care 2007;45(3):259-63.

Pisters and LeChevalier 2005
Pisters KM, LeChevalier T. Adjuvant chemotherapy in completely resected non-small-cell lung cancer. J Clin Oncol 2005;23(14):3270-8.

Reck et al 2014
Reck M, Popat S, Reinmuth N, De Ruysscher D, Kerr KM, Peters S; ESMO Guidelines Working Group. Metastatic non-small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014;25(Suppl 3):iii27-39.

Reck et al 2016
Reck M, Rodriquez-Abreu, D, Robinson AG, Hui R, Csözi T, Fülöp A, et al. Pembrolizumab versus chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer. New Engl J Med 2016; 375(19):1823-33.

Ribas et al 2013
Ribas A, Kefford R, Marshall MA, Punt CJA, Haanen JB, Marmol M, et al. Phase III randomized clinical trial comparing tremelimumab with standard-of-care chemotherapy in patients with advanced melanoma. J Clin Oncol 2013;31:616-22.

Robins 1993
Robins JM. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. Proceedings of the Biopharmaceutical Section, American Statistical Association 1993; 24-33

Robins and Tsiatis 1991
Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun Stat Theory Methods 1991; 20(8):2609-31.

Sandler et al 2006
Sandler A, Gray R, Perry MC, Brahmer J, Schiller JH, Dowlati A, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med 2006;355(24):2524-50.

Scagliotti et al 2008
Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008;26(21)3543-51.

Schadendorf et al 2013
Schadendorf D, Hodi FS, Robert C et al. Pooled analysis of long-term survival data from phase II and phase III trials of ipilimumab in metastatic or locally advanced, unresectable melanoma. Presented at: European Cancer Congress 2013 (ECCO-ESMO-ESTRO); September 27-October 1, 2013; Amsterdam, The Netherlands. Abstract 24.

Schiller et al 2002
Schiller JH, Harringdon D, Belani CP, Langer C, Sandler A, Krook J, et al. Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. N Engl J Med 2002;10(346):92-8.

Schwarzenbach et al 2014
Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical evidence of circulating cell-free microRNAs in cancer. Nat Rev Clin Oncol 2014;11(3):145-56.

Selke and Siegmund 1983
Selke T, Siegmund D. Sequential analysis of the proportional hazards model. Biometrika 1983;70:315-26.

Sprangers and Aaronson 1992
Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol 1992;45:743-60.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method with the conventional approach for interval-censored data analysis. Stat Biopharm Res 2010;2(1):97-108.

Tarhini and Kirkwood 2008
Tarhini AA, Kirkwood JM. Tremelimumab (CP-675,206): a fully human anticytotoxic T lymphocyte-associated antigen 4 monoclonal antibody for treatment of patients with advanced cancers. Expert Opin Biol Ther 2008;8(10):1583-93.

Topalian et al 2012
Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N Engl J Med 2012;366(26):2443-54.

Topalian et al 2014
Topalian SL, Sznol M, McDermott DF, Kluger HM, Carvajal RD, Sharfman WH et al. Survival, durable tumor remission, and long-term safety in patients with advanced melanoma receiving nivolumab. J Clin Oncol, 2014 Apr 1; 32(10):1020-30.

Wang et al 2013
Wang Z, Han J, Cui Y, Fan K, Zhou X. Circulating microRNA-21 as noninvasive predictive biomarker for response in cancer immunotherapy. Med Hypotheses 2013;81(1):41-3.

Weber et al 2012
Weber J, Kähler KC, Hauschild A. Management of immune-related adverse events and kinetics of response with ipilimumab. J Clin Oncol 2012;30(21):2691-7.

Whitehead and Whitehead 1991
Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.

Wolchok et al 2013
Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, et al. Nivolumab plus ipilimumab in advanced melanoma. N Engl J Med 2013;369(2):122-133

Yuan et al 2011
Yuan J, Adamow M, Ginsberg BA, Rasalan TS, Ritter E, Gallardo HF, et al. Integrated NY-ESO-1 antibody and CD8+ T-cell responses correlated with clinical benefit in advanced melanoma patients treated with ipilimumab. Proc Natl Acad Sci USA 2011;108(40):16723-8.

Zhang et al 2008
Zhang C, Wu S, Xue X, Li M, Qin X, Li W, et al. Anti-tumour immunotherapy by blockade of the PD-1/PD-L1 pathway with recombinant human PD-1-IgV. Cytotherapy 2008;10(7):711-9.

Appendix A - Additional Safety Information

FURTHER GUIDANCE ON THE DEFINITION OF A SERIOUS ADVERSE EVENT (SAE)

Life threatening

'Life-threatening' means that the patient was at immediate risk of death from the AE as it occurred or it is suspected that use or continued use of the product would result in the patient's death. 'Life-threatening' does not mean that had an AE occurred in a more severe form it might have caused death (eg, hepatitis that resolved without hepatic failure).

Hospitalization

Outpatient treatment in an emergency room is not in itself a serious AE, although the reasons for it may be (eg, bronchospasm, laryngeal edema). Hospital admissions and/or surgical operations planned before or during a study are not considered AEs if the illness or disease existed before the patient was enrolled in the study, provided that it did not deteriorate in an unexpected way during the study.

Important medical event or medical intervention

Medical and scientific judgement should be exercised in deciding whether a case is serious in situations where important medical events may not be immediately life threatening or result in death, hospitalization, disability, or incapacity but may jeopardize the patient or may require medical intervention to prevent 1 or more outcomes listed in the definition of serious. These should usually be considered as serious.

Simply stopping the suspect drug does not mean that it is an important medical event; medical judgement must be used.

Examples of such events are:

- Angioedema not severe enough to require intubation but requiring iv hydrocortisone treatment
- Hepatotoxicity caused by paracetamol (acetaminophen) overdose requiring treatment with N-acetylcysteine
- Intensive treatment in an emergency room or at home for allergic bronchospasm
- Blood dyscrasias (eg, neutropenia or anemia requiring blood transfusion, etc) or convulsions that do not result in hospitalization
- Development of drug dependency or drug abuse

A GUIDE TO INTERPRETING THE CAUSALITY QUESTION

When making an assessment of causality consider the following factors when deciding if there is a 'reasonable possibility' that an AE may have been caused by the drug.

- Time Course. Exposure to suspect drug. Has the patient actually received the suspect drug? Did the AE occur in a reasonable temporal relationship to the administration of the suspect drug?

- Consistency with known drug profile. Was the AE consistent with the previous knowledge of the suspect drug (pharmacology and toxicology) or drugs of the same pharmacological class? OR could the AE be anticipated from its pharmacological properties?

- Dechallenge experience. Did the AE resolve or improve on stopping or reducing the dose of the suspect drug?

- No alternative cause. The AE cannot be reasonably explained by another etiology such as the underlying disease, other drugs, or other host or environmental factors.

- Rechallenge experience. Did the AE reoccur if the suspected drug was reintroduced after having been stopped? AstraZeneca would not normally recommend or support a rechallenge.

- Laboratory tests. A specific laboratory investigation (if performed) has confirmed the relationship?

In difficult cases, other factors could be considered such as:

- Is this a recognized feature of overdose of the drug?

- Is there a known mechanism?

Causality of 'related' is made if following a review of the relevant data, there is evidence for a 'reasonable possibility' of a causal relationship for the individual case. The expression 'reasonable possibility' of a causal relationship is meant to convey, in general, that there are facts (evidence) or arguments to suggest a causal relationship.

The causality assessment is performed based on the available data including enough information to make an informed judgment. With limited or insufficient information in the case, it is likely that the event(s) will be assessed as 'not related'.

Causal relationship in cases where the disease under study has deteriorated due to lack of effect should be classified as no reasonable possibility.

MEDI4736 AND TREMELIMUMAB

There is no information to date on drug-drug interactions with MEDI4736 or tremelimumab either pre-clinically or in patients. As MEDI4736 and tremelimumab are monoclonal antibodies and therefore proteins, they will be degraded to small peptides and amino acids and will be eliminated by renal and reticuloendothelial clearance. It is therefore not expected that MEDI4736 or tremelimumab will induce or inhibit the major drug metabolizing cytochrome P450 pathways. As a result, there are no expected pharmacokinetic drug-drug interactions.

No formal drug-drug interaction studies have been conducted with tremelimumab. However, in renal cell carcinoma studies, acute renal failure has been reported with the combination of tremelimumab and sunitinib.

The mechanism of action of MEDI4736 involves binding to PD-L1, and the mechanism of action of tremelimumab involves binding to CTLA-4; therefore, significant pharmacodynamic drug interactions with the commonly administered concomitant medications are not expected. Despite this, appropriate clinical monitoring in all of the planned clinical studies will be conducted to evaluate any potential drug-drug interactions.

Appendix B - IATA 6.2 Guidance Document

INTERNATIONAL AIRLINE TRANSPORTATION ASSOCIATION (IATA) 6.2 GUIDANCE DOCUMENT

LABELLING AND SHIPMENT OF BIOHAZARD SAMPLES

International Airline Transportation Association (IATA) classifies biohazardous agents into 3 categories **(http://www.iata.org/whatwedo/cargo/dangerous_goods/infectious_substances.htm)**. For transport purposes, the classification of infectious substances according to risk groups was removed from the Dangerous Goods Regulations (DGR) in the 46th edition (2005). Infectious substances are now classified either as Category A, Category B, or Exempt. There is no direct relationship between Risk Groups and Categories A and B.

Category A Infectious Substances are infectious substances in a form that, when exposure to it occurs, is capable of causing permanent disability, life-threatening, or fatal disease in otherwise healthy humans or animals. Category A pathogens are, eg, Ebola, Lassa fever virus:

- Are to be packed and shipped in accordance with IATA Instruction 602.

Category B Infectious Substances are infectious substances that do not meet the criteria for inclusion in Category A. Category B pathogens are, eg, Hepatitis A, B, C, D, and E viruses, Human immunodeficiency virus (HIV) types 1 and 2. They are assigned the following UN number and proper shipping name:

- UN 3373 – Biological Substance, Category B
- Are to be packed in accordance with UN3373 and IATA 650

Exempt - all other materials with minimal risk of containing pathogens

- Clinical trial samples will fall into Category B or exempt under IATA regulations
- Clinical trial samples will routinely be packed and transported at ambient temperature in IATA 650 compliant packaging **(http://www.iata.org/whatwedo/cargo/dangerous_goods/infectious_substances.htm)**
- **Biological samples transported in dry ice require additional dangerous goods specification for the dry-ice content**
- IATA compliant courier and packaging materials should be used for packing and transportation and packing should be done by an IATA certified person, as applicable

- Samples routinely transported by road or rail are subject to local regulations which require that they are also packed and transported in a safe and appropriate way to contain any risk of infection or contamination by using approved couriers and packaging / containment materials at all times. The IATA 650 biological sample containment standards are encouraged wherever possible when road or rail transport is used.

Appendix C - Pharmacogenetics Research

BACKGROUND AND RATIONALE

AstraZeneca intends to perform genetic research in the MEDI4736 clinical development program to explore how genetic variations may affect the clinical parameters associated with this drug combination. Collection of DNA samples from populations with well-described clinical characteristics may lead to improvements in the design and interpretation of clinical trials and, possibly, to genetically guided treatment strategies.

Future research may suggest other genes or gene categories as candidates for influencing not only response to MEDI4736, but also susceptibility to NSCLC. Thus, this genetic research may involve study of additional un-named genes or gene categories, but only as related to NSCLC and MEDI4736 treatment.

GENETIC RESEARCH OBJECTIVES

The objective of this research is to collect and store DNA, derived from a blood sample, for future exploratory research into genes/genetic variations that may influence response, ie, distribution, safety, tolerability, and efficacy of MEDI4736, and/or susceptibility to NSCLC.

GENETIC RESEARCH PLAN AND PROCEDURES

Selection of genetic research population

Study selection record

All enrolled patients who take part in the main study will be asked to participate in this genetic research. Participation is voluntary, and if a patient declines to participate, there will be no penalty or loss of benefit. The patient will not be excluded from any aspect of the main study.

Inclusion criteria

For inclusion in this genetic research, patients must fulfil all of the inclusion criteria described in the main body of the Clinical Study Protocol **and**:

- Provide informed consent for the genetic sampling and analyses.

Exclusion criteria

Exclusion from this genetic research may be for any of the exclusion criteria specified in the main study or any of the following:

- Previous allogeneic bone marrow transplant
- Non-leukocyte depleted whole blood transfusion in 120 days of genetic sample collection

Discontinuation of patients from this genetic research

Specific reasons for discontinuing a patient from this genetic research are:

Withdrawal of consent for genetic research: patients may withdraw from this genetic research at any time, independent of any decision concerning participation in other aspects of the main study. Voluntary discontinuation will not prejudice further treatment. Procedures for discontinuation are outlined in Section 3.9 of the main Clinical Study Protocol.

Collection of samples for genetic research

Blood samples will ideally be collected during the screening/baseline period. If for any reason the sample is not drawn during the screening/baseline period, it should be taken as soon as possible, but not later than the last study visit. Although genotype is a stable parameter, early sample collection is preferred to avoid introducing bias through excluding patients who may withdraw due to an adverse event, as these patients would be important to include in any genetic analysis. Only 1 sample should be collected per patient for genetics during the study.

If the patient agrees to participate, an 9-mL blood sample will be collected into a tube containing reagents that coagulate blood and stabilize blood cell DNA and gently inverted a minimum of 5 times to mix thoroughly. Tubes will be identified with the protocol study number, center number, enrollment code, and date of sample collection. No personal identifiers (patient name, initials, or date of birth) will be placed on the tube or accompanying documentation. A record of the date of the patient consent to the host genetic research and the date of the blood sample collection will be recorded.

AstraZeneca/MedImmune, or its designee, will act as the central laboratory for sample logistics. This will include the supply of site material and all transport arrangements.

A single blood sample will be stored frozen (-20°C or below) at the site and sent to the central laboratory. The central laboratory will then send the samples to AstraZeneca/MedImmune, or its designee laboratory, for DNA extraction. Samples must remain frozen at all times. Further details on the processing of the samples are outlined in the Laboratory Manual for Investigators.

Coding and storage of DNA samples

The processes adopted for the coding and storage of samples for genetic analysis are important to maintain patient confidentiality. Samples will be stored for a maximum of 15 years from the date of the last patient's last visit, after which they will be destroyed. DNA is a finite resource that may be used up during analyses. Samples will be stored and used until no further analyses are possible or the maximum storage time has been reached.

For all samples irrespective of the type of coding used the DNA will be extracted from the blood sample. The DNA sample will be assigned a unique number replacing the information on the sample tube. Thereafter, the DNA sample will be identifiable by the unique DNA number only. The DNA number is used to identify the sample and corresponding data at the AstraZeneca/MedImmune genetics laboratories, or at the designated contract laboratory. No personal details identifying the individual will be available to any person (AstraZeneca/MedImmune employee or contract laboratory staff working with the DNA).

The samples and data for genetic analysis in this study will be single coded. The link between the patient enrollment/randomization code and the DNA number will be maintained and stored in a secure environment, with restricted access WITHIN the Clinical Genotyping Group Laboratory Information Management System (LIMS) at AstraZeneca. The link will be used to identify the relevant DNA samples for analysis, facilitate correlation of genotypic results with clinical data, allow regulatory audit and to trace samples for destruction in the case of withdrawal of consent when the patient has requested disposal/destruction of collected samples not yet analyzed.

ETHICAL AND REGULATORY REQUIREMENTS

The principles for ethical and regulatory requirements for the study, including this genetics research component, are outlined in Section 10 of the main Clinical Study Protocol.

The study will be performed in accordance with ethical principles that have their origin in the Declaration of Helsinki and are consistent with International Conference on Harmonisation/Good Clinical Practice, applicable regulatory requirements, and the AstraZeneca policy on Bioethics.

Informed consent

The portion of this study evaluating genetic alterations in blood samples is optional, and the patient may participate in other components of the main study without participating in this specific genetic analysis. To participate in this genetic component of the study, the patient must sign and date both the consent form for the main study and the genetic component of the study. Copies of both signed and dated consent forms must be given to the patient and the original filed at the study center. The principal investigator(s) is responsible for ensuring that consent is given freely and that the patient understands that they may freely discontinue from the genetic aspect of the study at any time.

Subject data protection

AstraZeneca/MedImmune will not provide individual genotype results to patients, any insurance company, any employer, their family members, their general physician, or any other third party, unless required to do so by law.

Extra precautions are taken to preserve confidentiality and prevent genetic data being linked to the identity of the patient. Individual patients will not be identified in any report or publication resulting from this work. The data and results of this study may be reviewed with collaborators and published, but neither the patient's name nor any other personal identifiers will appear in any publication or report. In exceptional circumstances, however, certain individuals might see both the genetic data and the personal identifiers of a patient. For example, in the case of a medical emergency, an AstraZeneca/MedImmune physician or an investigator might know a patient's identity and also have access to his or her genetic data. Regulatory authorities may also require access to the relevant files, though the patient's medical information and the genetic files would remain physically separate.

DATA MANAGEMENT

Any genotype data generated in this study will be stored in the AstraZeneca genotyping LIMS database, or other appropriate secure system within AstraZeneca and/or third party contracted to work with AstraZeneca/MedImmune to analyze the samples.

The results from this genetic research will be reported separately from the clinical study report for the main study.

Some or all of the clinical datasets from the main study may be merged with the genetic data in a suitable secure environment separate from the clinical database.

STATISTICAL METHODS AND DETERMINATION OF SAMPLE SIZE

The number of patients that will agree to participate in the genetic research is unknown. It is therefore not possible to establish whether sufficient data will be collected to allow a formal statistical evaluation or whether only descriptive statistics will be generated. A statistical analysis plan will be prepared where appropriate.

LIST OF REFERENCES

None

Appendix D - Hy's Law

ACTIONS REQUIRED IN CASES OF COMBINED INCREASE OF AMINOTRANSFERASE AND TOTAL BILIRUBIN - HY'S LAW

INTRODUCTION

This Appendix describes the process to be followed in order to identify and appropriately report cases of Hy's Law. It is not intended to be a comprehensive guide to the management of elevated liver biochemistries. Specific guidance on the managing liver abnormalities can be found in Section 6.7 of the protocol.

During the course of the study the Investigator will remain vigilant for increases in liver biochemistry. The investigator is responsible for determining whether a patient meets potential Hy's Law (PHL) criteria at any point during the study.

The Investigator participates, together with AstraZeneca clinical project representatives, in review and assessment of cases meeting PHL criteria to agree whether Hy's Law (HL) criteria are met. HL criteria are met if there is no alternative explanation for the elevations in liver biochemistry other than Drug Induced Liver Injury (DILI) caused by the Investigational Medicinal Product (IMP).

The Investigator is responsible for recording data pertaining to PHL/HL cases and for reporting Adverse Events (AE) and Serious Adverse Events (SAE) according to the outcome of the review and assessment in line with standard safety reporting processes.

DEFINITIONS

Potential Hy's Law (PHL)

Aspartate Aminotransferase (AST) or Alanine Aminotransferase (ALT) ≥3 × Upper Limit of Normal (ULN) **together with** Total Bilirubin (TBL) ≥2 × ULN at any point during the study following the start of study medication irrespective of an increase in Alkaline Phosphatase (ALP).

Hy's Law (HL)

AST or ALT ≥3 × ULN **together with** TBL ≥2 × ULN, where no other reason, other than the IMP, can be found to explain the combination of increases, eg, elevated ALP indicating cholestasis, viral hepatitis, another drug.

For PHL and HL to be met the elevation in transaminases must precede or be coincident with (ie, on the same day) the elevation in TBL, but there is no specified timeframe within which the elevations in transaminases and TBL must occur.

IDENTIFICATION OF POTENTIAL HY'S LAW CASES

In order to identify cases of PHL it is important to perform a comprehensive review of laboratory data for any patient who meets any of the following identification criteria in isolation or in combination:

- ALT ≥3 × ULN
- AST ≥3 × ULN
- TBL ≥2 × ULN

The Investigator will remain vigilant for any local laboratory reports where the identification criteria are met, where this is the case the Investigator will:

- Notify the AstraZeneca representative
- Request a repeat of the test (new blood draw) by the laboratory
- Complete the appropriate unscheduled laboratory CRF module(s) with the original local laboratory test result

When the identification criteria are met from local laboratory results the Investigator will without delay:

- Determine whether the patient meets PHL criteria (see Section 0 of this Appendix for definition) by reviewing laboratory reports from all previous visits Review each new laboratory report and if the identification criteria are met will:
  - Determine whether the patient meets PHL criteria (see Section 0 of this Appendix for definition) by reviewing laboratory reports from all previous visits
  - Promptly enter the laboratory data into the laboratory CRF

FOLLOW-UP

Potential Hy's Law Criteria not met

If the patient does not meet PHL criteria the Investigator will:

- Perform follow-up on subsequent laboratory results according to the guidance provided in the Clinical Study Protocol.

Potential Hy's Law Criteria met

If the patient does meet PHL criteria the Investigator will:

- Determine whether PHL criteria were met at any study visit prior to starting study treatment in the presence of liver metastases (Actions required when potential Hy's law criteria are met before and after starting study treatment.)
- Notify the AstraZeneca representative who will then inform the central Study Team

The Study Physician contacts the Investigator, to provide guidance, discuss, and agree an approach for the study patients' follow-up and the continuous review of data. Subsequent to this contact the Investigator will:

- Monitor the patient until liver biochemistry parameters and appropriate clinical symptoms and signs return to normal or baseline levels, or as long as medically indicated
- Investigate the etiology of the event and perform diagnostic investigations as discussed with the Study Physician
- Complete the 3 Liver CRF Modules as information becomes available
- If at any time (in consultation with the Study Physician) the PHL case meets serious criteria, report it as an SAE using standard reporting procedures

REVIEW AND ASSESSMENT OF POTENTIAL HY'S LAW CASES

The instructions in this section should be followed for all cases where PHL criteria are met.

No later than 3 weeks after the biochemistry abnormality was initially detected, the Study Physician contacts the Investigator in order to review available data and agree on whether there is an alternative explanation for meeting PHL criteria other than DILI caused by the IMP. The AstraZeneca Medical Science Director and Global Safety Physician will also be involved in this review together with other subject matter experts as appropriate.

According to the outcome of the review and assessment, the Investigator will follow the instructions below.

If there **is** an agreed alternative explanation for the ALT or AST and TBL elevations, a determination of whether the alternative explanation is an AE will be made and subsequently whether the AE meets the criteria for a SAE:

- If the alternative explanation is **not** an AE, record the alternative explanation on the appropriate CRF

- If the alternative explanation is an AE/SAE, record the AE /SAE in the CRF accordingly and follow the AZ standard processes

If it is agreed that there is **no** explanation that would explain the ALT or AST and TBL elevations other than the IMP:

- Report an SAE (report term 'Hy's Law') according to AstraZeneca standard processes.
  - The 'Medically Important' serious criterion should be used if no other serious criteria apply
  - As there is no alternative explanation for the HL case, a causality assessment of 'related' should be assigned.

If, there is an unavoidable delay, of over 3 weeks, in obtaining the information necessary to assess whether or not the case meets the criteria for HL, then it is assumed that there is no alternative explanation until such time as an informed decision can be made:

- Report an SAE (report term 'Potential Hy's Law') applying serious criteria and causality assessment as per above
- Continue follow-up and review according to agreed plan. Once the necessary supplementary information is obtained, repeat the review and assessment to determine whether HL criteria are met. Update the SAE report according to the outcome of the review

ACTIONS REQUIRED WHEN POTENTIAL HY'S LAW CRITERIA ARE MET BEFORE AND AFTER STARTING STUDY TREATMENT

This section is applicable to patients who meet PHL criteria on study treatment (including the 30-day follow-up period post discontinuation of study treatment) having previously met PHL criteria at a study visit prior to starting study treatment.

At the first on study treatment occurrence of PHL criteria being met the Investigator will:

- Determine if there has been a significant change in the patients' condition[#] compared with the last visit where PHL criteria were met[#]
  - If there is no significant change no action is required
  - If there is a significant change notify the AstraZeneca representative, who will inform the central Study Team, then follow the subsequent process described in Section 0 of this Appendix

[#] A 'significant' change in the patient's condition refers to a subsequent clinically relevant change in any of the individual liver biochemistry parameters (ALT, AST, or total bilirubin) in isolation or in combination, or a clinically relevant change in associated symptoms such as fatigue, vomiting, rash, right upper quadrant pain, jaundice, or eosinophilia. The determination of whether there has been a significant change will be at the discretion of the Investigator; this may be in consultation with the Study Physician if there is any uncertainty.

ACTIONS REQUIRED FOR REPEAT EPISODES OF POTENTIAL HY'S LAW

This section is applicable when a patient meets PHL criteria on study treatment (including the 30-day follow-up period) and has already met PHL criteria at a previous on study treatment visit.

The requirement to conduct follow-up, review, and assessment of a repeat occurrence(s) of PHL is based on the nature of the alternative cause identified for the previous occurrence.

The Investigator should determine the cause for the previous occurrence of PHL criteria being met and answer the following question:

- Was the alternative cause for the previous occurrence of PHL criteria being met chronic or progressing malignant disease or did the patient meet PHL criteria prior to starting study treatment and at their first on study treatment visit as described in Section 6?

If No: follow the process described in Section 0 of this Appendix

If Yes:

Determine if there has been a significant change in the patient's condition[#] compared with when PHL criteria were previously met

- If there is no significant change no action is required
- If there is a significant change follow the process described in Section 0 of this Appendix

[#] A 'significant' change in the patient's condition refers to a subsequent clinically relevant change in any of the individual liver biochemistry parameters (ALT, AST, or total bilirubin) in isolation or in combination, or a clinically relevant change in associated symptoms, such as fatigue, vomiting, rash, right upper quadrant pain, jaundice, or eosinophilia. The determination of whether there has been a significant change will be at the discretion of the Investigator; this may be in consultation with the Study Physician if there is any uncertainty.

REFERENCES

FDA Guidance for Industry (issued July 2009) 'Drug-induced liver injury: Premarketing clinical evaluation'
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf

Appendix E - Guidelines for Evaluation of Objective Tumor Response

GUIDELINES FOR EVALUATION OF OBJECTIVE TUMOR RESPONSE USING RECIST 1.1 CRITERIA (RESPONSE EVALUATION CRITERIA IN SOLID TUMORS)

INTRODUCTION

This appendix details the implementation of Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 guidelines (Eisenhauer et al 2009) for the D419AC00001 study with regards to Investigator assessment of tumor burden including protocol-specific requirements for this study.

DEFINITION OF MEASURABLE, NON-MEASURABLE, TARGET AND NON-TARGET LESIONS

Only patients with measurable disease at baseline should be included in the study. Measurable disease is defined by the presence of at least one measurable lesion which has not been previously irradiated.

Tumor lesions selected for screening biopsy must not be used as target lesions, unless there are no other lesions suitable for biopsy.

Measurable:

A lesion, not previously irradiated or biopsied per the protocol prior to randomisation, that can be accurately measured at baseline as ≥10 mm in the longest diameter (except lymph nodes which must have short axis ≥15 mm) with computed tomography (CT) or magnetic resonance imaging (MRI) and which is suitable for accurate repeated measurements.

Non-measurable:

- All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with ≥10 mm to <15 mm short axis at baseline[1]).
- Truly non-measurable lesions include the following: bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung and, abdominal masses/abdominal

[1] Nodes with <10 mm short axis are considered non-pathological and should not be recorded or followed as non-target lesions (NTLs).

organomegaly identified by physical examination that is not measurable by CT or MRI.

- Tumour lesions selected for screening biopsy
- Previously irradiated lesions[2]
- Skin lesions assessed by clinical examination
- Brain metastasis

Special cases:

- Lytic bone lesions or mixed lytic–blastic lesions, with identifiable soft tissue components, can be considered measurable if the soft tissue component meets the definition of measurability. Blastic lesions are considered non-measurable.
- Cystic metastases can be considered measurable lesions if they meet the criteria for measurability from a radiological point of view, but if non-cystic lesions are present in the same patient, these should be selected as target lesions (TLs).

Target lesions:

A maximum of 5 measurable lesions (with a maximum of 2 lesions per organ), representative of all lesions involved suitable for accurate repeated measurement, should be identified as TLs at baseline.

Non-target lesions:

All other lesions (or sites of disease) not recorded as TL should be identified as NTL at baseline.

METHODS OF ASSESSMENT

The same method of assessment and the same technique should be used to characterize each identified and recorded lesion at baseline and during follow-up visits.

A summary of the methods to be used for RECIST assessment is provided in Table 11, and those excluded from tumor assessments for this study are highlighted with the rationale provided.

[2] Localised post-radiation changes which affect lesion sizes may occur. Therefore, lesions that have been previously irradiated will not be considered measurable and must be selected as NTL at baseline and followed up as part of the NTL assessment.

Table 11 Summary of methods of assessment

Target lesions	Non-target lesions	New lesions
CT (preferred)	CT (preferred)	CT (preferred)
MRI	MRI	MRI
	Clinical examination	Clinical examination
	X-ray, Chest X-ray	X-ray, Chest X-ray
		Ultrasound
		Bone scan
		FDG-PET

CT Computed tomography; FDG-PET 18-Fluoro-deoxyglucose positron emission tomography; MRI Magnetic resonance imaging.

CT and MRI

CT and MRI are generally considered to be the best currently available and reproducible methods to measure TL selected for response assessment and to assess NTL and identification of any new lesions.

In the D419AC00001 study, it is recommended that CT examinations of the chest and abdomen (including liver and adrenal glands) will be used to assess tumour burden at baseline and follow-up visits. CT examination with intravenous contrast media administration is the preferred method. MRI should be used where CT is not feasible or it is medically contra-indicated. For brain lesion assessment, MRI is the preferred method.

Clinical examination

In the D419AC00001 study, clinical examination will not be used for assessment of TL. Clinically detected lesions can be selected as TLs if they are assessed by CT or MRI scans. Clinical examination can be used to assess NTL and to identify the presence of new lesions.

X-ray

Chest X-ray

In the D419AC00001 study, chest X-ray assessment will not be used for assessment of TL as they will be assessed by CT examination or MRI examination. Chest X-ray can, however, be used to assess NTL and to identify the presence of new lesions.

Plain X-ray

In the D419AC00001 study plain X-ray may be used as a method of assessment for bone NTL and to identify the presence of new bone lesions.

Ultrasound

In the D419AC00001 study, ultrasound examination will not be used for assessment of TL and NTL as it is not a reproducible method, does not provide an accurate assessment of tumor

size, and it is subjective and operator dependent. Ultrasound examination can, however, be used to identify the presence of new lesions. If new clinical symptoms occur and an ultrasound is performed then new lesions should be confirmed by CT or MRI examination.

Endoscopy and laparoscopy

In the D419AC00001 study, endoscopy and laparoscopy will not be used for tumor assessments as they are not validated in the context of tumor assessment.

Tumor markers

In the D419AC00001 study, tumor markers will not be used for tumor response assessments as per RECIST 1.1.

Cytology and histology

In the D419AC00001 study histology will not be used as part of the tumor response assessment as per RECIST 1.1.

Cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment is required when the measurable tumour has met criteria for response or stable disease. In such circumstances, the cytology is necessary to differentiate between response/stable disease (an effusion may be a side effect of the treatment) and progressive disease (if the neoplastic origin of the fluid is confirmed). Where cytology findings are not available, any effusion that significantly worsens (from trace to large) or appearance of clinically significant effusion (requiring change in drug therapy) during the study treatment will be considered to be progression of NTL, or disease progression due to new lesions.

Isotopic bone scan

Bone lesions identified on an isotopic bone scan at baseline and confirmed by CT, MRI, or X-ray at baseline should be recorded as NTL and followed by the same method as per baseline assessment.

In the D419AC00001 study, isotopic bone scans may be used as a method of assessment to identify the presence of new bone lesions at follow-up visits. New lesions will be recorded where a positive hot-spot that was not present on the baseline bone scan assessment is identified on a bone scan performed at any time during the study. The Investigator should consider the positive hot-spot to be a significant new site of malignant disease and represent true disease progression in order to record the new lesion. Confirmation by CT, MRI and X-ray is recommended where bone scan findings are equivocal.

FDG-PET scan

^{18}F-Fluoro-deoxyglucose positron emission tomography (FDG-PET) scans may be used as a method for identifying new lesions, according with the following algorithm: New lesions will

be recorded where there is positive ^{18}F-Fluoro-deoxyglucose uptake[3] not present on baseline FDG-PET scan or in a location corresponding to a new lesion on CT/MRI at the same follow-up visit. If there is no baseline FDG-PET scan available, and no evidence of new lesions on CT/MRI scans then follow-up CT/MRI assessments should be continued, scheduled as per protocol or clinical indicated, in order to confirm new lesions.

TUMOR RESPONSE EVALUATION

Schedule of evaluation

The methods of assessment of tumor burden used at baseline CT/MRI scans of the chest and abdomen (including liver and adrenal glands) must be used at each subsequent follow-up assessment. Additional imaging may be performed based on the signs and symptoms of the patient, eg, new lesions at follow up.

Baseline assessments should be performed no more than 28 days before start of study treatment, and ideally should be performed as close as possible to the start of investigational product. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; see Section 3.1 of the Clinical Study Protocol), then every 8 weeks thereafter until confirmed objective disease progression as defined by RECIST 1.1 (irrespective of the reason for stopping treatment/or subsequent therapy). 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.

For patients who discontinue study drug due to toxicity in the absence of confirmed objective progression, objective tumor assessments should be continued every 6 weeks for 48 weeks (relative to the date of randomization) then every 8 weeks until confirmed objective disease progression.

Disease progression requires confirmation; the confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD in the absence of clinically significant deterioration.

If progression is not confirmed then the patient should continue on study treatment and on treatment assessments.

If an unscheduled assessment was 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] A positive FDG-PET scan lesion should be reported only when an uptake greater than twice that of the surrounding tissue is observed.

Additional assessments will be performed post confirmed objective disease progression for patients remaining on IMT treatment, re-treatment, or until subsequent cancer therapy according to the clinical study protocol.

Target lesions

Documentation of target lesions

A maximum of 5 measurable lesions, with a maximum of 2 lesions per organ (including lymph nodes collectively considered as a single organ), representative of all lesions involved should be identified as TL at baseline. Target lesions should be selected on the basis of their size (longest diameter for non-nodal lesions or short axis for nodal lesions), but in addition should be those that lend themselves to reproducible repeated measurements. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion, which can be measured reproducibly, should be selected.

The site and location of each TL should be documented as well as the longest diameter for non-nodal lesions (or short axis for lymph nodes). All measurements should be recorded in millimeters. At baseline the sum of the diameters for all TL will be calculated and reported as the baseline sum of diameters. At follow-up visits the sum of diameters for all TL will be calculated and reported as the follow-up sum of diameters.

Special cases:

- For TL measurable in 2 or 3 dimensions, always report the longest diameter. For pathological lymph nodes measurable in 2 or 3 dimensions, always report the short axis.
- If the CT/MRI slice thickness used is >5 mm, the minimum size of measurable disease at baseline should be twice the slice thickness of the baseline scan.
- If a lesion has completely disappeared, the longest diameter should be recorded as 0 mm.
- If a TL splits into two or more parts, then record the sum of the diameters of those parts.
- If two or more TLs merge then the sum of the diameters of the combined lesion should be recorded for one of the lesions and 0 mm recorded for the other lesion(s).
- If a TL is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned. If an accurate measure can be given, this should be recorded, even if it is below 5 mm.
- If a TL cannot be measured accurately due to it being too large, provide an estimate of the size of the lesion.

- When a TL has had any intervention eg, radiotherapy, embolization, surgery, during the study, the size of the TL should still be provided where possible.

Evaluation of target lesions

This section provides the definitions of the criteria used to determine objective tumor visit response for TL (see Table 12).

Table 12 Evaluation of target lesions

Complete Response (CR)	Disappearance of all TLs since baseline. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10 mm.
Partial Response (PR)	At least a 30% decrease in the sum of the diameters of TL, taking as reference the baseline sum of diameters
Stable Disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Progression of disease (PD)	At least a 20% increase in the sum of diameters of TLs, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm.
Not Evaluable (NE)	Only relevant if any of the TLs were not assessed or not evaluable or had a lesion intervention at this visit. Note: if the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a TL response

CR Complete response; PR Partial response; PD Progression of disease; NE Not evaluable; SD Stable disease; TL Target lesion.

Non-target lesions

Evaluation of non-target lesions

All other lesions (or sites of disease) not recorded as TL should be identified as NTL at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits. At each visit an overall assessment of the NTL response should be recorded by the Investigator. This section provides the definitions of the criteria used to determine and record overall response for NTL at the investigational site at each visit (see Table 13).

Table 13 Evaluation of non-target lesions

Complete response (CR)	Disappearance of all NTLs since baseline. All lymph nodes must be non-pathological in size (<10 mm short axis).
Non CR/Non PD	Persistence of one or more NTL.
Progression (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.
Not evaluable (NE)	Only relevant when one or some of the NTLs were not assessed and, in the Investigator's opinion, they are not able to provide an evaluable overall NTL assessment at this visit. Note: for patients without TLs at baseline, this is relevant if any of the NTLs were not assessed at this visit and the progression criteria have not been met.

CR Complete response; PR Partial response; PD Progression of disease; NE Not evaluable; NTL Non-target lesion; TL Target lesion.

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

New lesions

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: ie, not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumor.

If a new lesion is equivocal, for example because of its small size, the treatment and tumor assessments should be continued until the new lesion has been confirmed. If repeat scans confirm there is a new lesion, then the progression date should be declared using the date of the initial scan.

Symptomatic deterioration

Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping study therapy.

Patients with 'symptomatic deterioration' 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.

Evaluation of overall visit response

The overall visit response will be derived using the algorithm shown in Table 14.

Table 14 Overall visit response

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	NA	No	CR
CR	Non CR/Non PD	No	PR
CR	NE	No	PR
PR	Non PD or NE	No	PR
SD	Non PD or NE	No	SD
NE	Non PD or NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR Complete response, PR Partial response, SD Stable disease, PD Progression of disease, NE Not evaluable, NA Not applicable (only relevant if there were no non-target lesions at baseline).

CONFIRMATION OF PROGRESSION

Disease progression requires confirmation; the confirmatory scan should occur no earlier than 4 weeks after the initial assessment of progression of disease (PD) in the absence of clinical deterioration.

Progression would be considered confirmed if the following criteria are met:

- ≥20% increase in the sum diameters of TLs compared with the nadir at 2 consecutive visits, with an absolute increase of at least 5mm in sum of diameters compared to nadir,

- *and/or* significant progression (worsening) of NTLs and/or pre-existing new lesions at the confirmatory scan time-point compared with the first time point where progression of NTLs or new lesions identified

- *and/or* additional new unequivocal lesions at the confirmatory scan time-point

In the absence of significant clinical deterioration, the Investigator should continue study treatment until progression is confirmed.

If progression is not confirmed, then the patient should continue on study treatment and on treatment assessments.

If a patient discontinues treatment (and/or receives a subsequent cancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

CENTRAL REVIEW

The Contract Research Organization appointed by AstraZeneca to perform the independent central review for this study will provide specification for radiological imaging protocols in standard acquisition guidelines documentation. The management of patients will be based solely upon the results of the RECIST 1.1 and confirmation of radiologic progression assessments conducted by the Investigator.

REFERENCES

Eisenhauer et al 2009
Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumors: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228-47.

Appendix F - Patient Reported Outcomes

EORTC QLQ-C30, EORTC QLQ-LC13, PRO-CTCAE, EQ-5D-5L, AND PGIC

EORTC QLQ-C30 (VERSION 3)

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

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

During the past week:

6.	Were you limited in doing either your work or other daily activities?	1	2	3	4
7.	Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8.	Were you short of breath?	1	2	3	4
9.	Have you had pain?	1	2	3	4
10.	Did you need to rest?	1	2	3	4
11.	Have you had trouble sleeping?	1	2	3	4
12.	Have you felt weak?	1	2	3	4
13.	Have you lacked appetite?	1	2	3	4

14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4
16. Have you been constipated?	1	2	3	4

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

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

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

1	2	3	4	5	6	7
Very poor						Excellent

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

1	2	3	4	5	6	7
Very poor						Excellent

EORTC QLQ - LC13

Patients sometimes report that they have the following symptoms or problems.

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

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

NCI- PRO-CTCAE ITEMS

As individuals go through treatment for their cancer they sometimes experience different symptoms and side effects. For each question, please check or mark an ⓧ in the one box that best describes your experiences over the past 7 days…

RASH				
Did you have any RASH?				
o Yes	o No			
HAIR LOSS				
Did you have any HAIR LOSS?				
o Not at all	o A little bit	o Somewhat	o Quite a bit	o Very much
HAND-FOOT SYNDROME (A RASH OF THE HANDS OR FEET THAT CAN CAUSE CRACKING, PEELING, REDNESS OR PAIN)				
What was the SEVERITY of your HAND-FOOT SYNDROME (A RASH OF THE HANDS OR FEET THAT CAN CAUSE CRACKING, PEELING, REDNESS OR PAIN) at its WORST?				
o None	o Mild	o Moderate	o Severe	o Very severe
ITCHY SKIN				
What was the SEVERITY of your ITCHY SKIN at its WORST?				
o None	o Mild	o Moderate	o Severe	o Very severe
ARM OR LEG SWELLING				
How often did you have ARM OR LEG SWELLING?				
o Never	o Rarely	o Occasionally	o Frequently	o Almost Constantly
What was the SEVERITY of your ARM OR LEG SWELLING at their WORST?				
o None	o Mild	o Moderate	o Severe	o Very severe
How much did ARM OR LEG SWELLING INTERFERE with your usual or daily activities?				
o Not at all	o A little bit	o Somewhat	o Quite a bit	o Very much
NAUSEA				
How often did you have NAUSEA?				
o Never	o Rarely	o Occasionally	o Frequently	o Almost Constantly
What was the SEVERITY of your NAUSEA at its WORST?				
o None	o Mild	o Moderate	o Severe	o Very severe
VOMITING				
How often did you have VOMITING?				
o Never	o Rarely	o Occasionally	o Frequently	o Almost Constantly
What was the SEVERITY of your VOMITING at its WORST?				

○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
LOOSE OR WATERY STOOLS (DIARRHEA)				
How often did you have LOOSE OR WATERY STOOLS (DIARRHEA)?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly

Please think back over the past 7 days…

NUMBNESS OR TINGLING IN YOUR HANDS OR FEET				
What was the SEVERITY of your NUMBNESS OR TINGLING IN YOUR HANDS OR FEET at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did NUMBNESS OR TINGLING IN YOUR HANDS OR FEET INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much
MOUTH AND THROAT SORES				
What was the SEVERITY of your MOUTH AND THROAT SORES at their WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did MOUTH AND THROAT SORES INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much
SHIVERING OR SHAKING CHILLS				
How often did you have SHIVERING OR SHAKING CHILLS?				
○ Never	○ Rarely	○ Occasionally	○ Frequently	○ Almost Constantly
What was the SEVERITY of your SHIVERING OR SHAKING CHILLS at their WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
PAIN, SWELLING, REDNESS AT A SITE OF DRUG INJECTION OR IV				
Did you have any PAIN, SWELLING, REDNESS AT A SITE OF DRUG INJECTION OR IV?				
○ Yes	○ No			
DIZZINESS				
What was the SEVERITY of your DIZZINESS at its WORST?				
○ None	○ Mild	○ Moderate	○ Severe	○ Very severe
How much did DIZZINESS INTERFERE with your usual or daily activities?				
○ Not at all	○ A little bit	○ Somewhat	○ Quite a bit	○ Very much

Developed by the National Cancer Institute and used with permission

Health Questionnaire
English version for the UK

Under each heading, please tick the ONE box that best describes your health TODAY

MOBILITY

I have no problems in walking about □
I have slight problems in walking about □
I have moderate problems in walking about □
I have severe problems in walking about □
I am unable to walk about □

SELF-CARE

I have no problems washing or dressing myself □
I have slight problems washing or dressing myself □
I have moderate problems washing or dressing myself □
I have severe problems washing or dressing myself □
I am unable to wash or dress myself □

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities*

I have no problems doing my usual activities □
I have slight problems doing my usual activities □
I have moderate problems doing my usual activities □
I have severe problems doing my usual activities □
I am unable to do my usual activities □

PAIN / DISCOMFORT

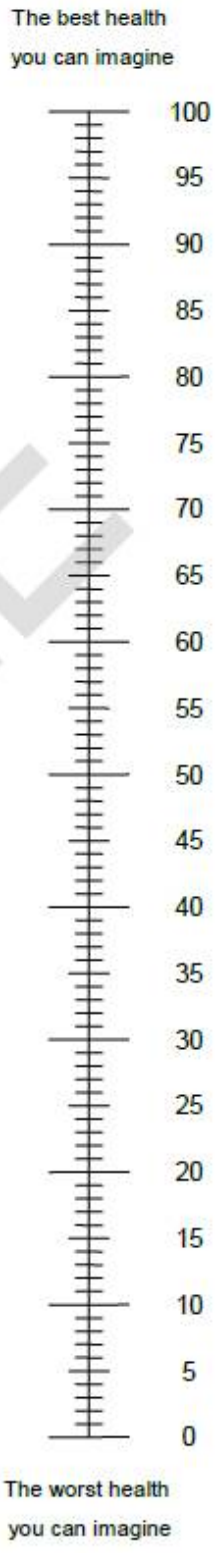
I have no pain or discomfort □
I have slight pain or discomfort □
I have moderate pain or discomfort □
I have severe pain or discomfort □
I have extreme pain or discomfort □

ANXIETY / DEPRESSION

I am not anxious or depressed □
I am slightly anxious or depressed □
I am moderately anxious or depressed □
I am severely anxious or depressed □
I am extremely anxious or depressed □

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



PATIENT GLOBAL IMPRESSION OF CHANGE (PGIC)

Since the start of the treatment I have received in this study, my overall health status is:

Please tick (✓) one box only:

- ☐ Very Much Improved
- ☐ Much Improved
- ☐ Minimally Improved
- ☐ No Change
- ☐ Minimally Worse
- ☐ Much Worse
- ☐ Very Much Worse

Appendix G - Dosing modification and toxicity management guidelines

Dosing Modification and Toxicity Management Guidelines for Immune-Mediated, Infusion-Related, and Non-Immune–Mediated Reactions (MEDI4736 Monotherapy or Combination Therapy With Tremelimumab or Tremelimumab Monotherapy) 1 November 2017 Version

General Considerations

Dose Modifications	Toxicity Management
Drug administration modifications of study drug/study regimen will be made to manage potential immune-related AEs based on severity of treatment-emergent toxicities graded per NCI CTCAE v4.03. In addition to the criteria for permanent discontinuation of study drug/study regimen based on CTC grade/severity (table below), permanently discontinue study drug/study regimen for the following conditions: • Inability to reduce corticosteroid to a dose of ≤10 mg of prednisone per day (or equivalent) **within 12 weeks** after last dose of study drug/study regimen • Recurrence of a previously experienced Grade 3 treatment-related AE following resumption of dosing **Grade 1** No dose modification **Grade 2** Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1. If toxicity worsens, then treat as Grade 3 or Grade 4. Study drug/study regimen can be resumed once event stabilizes to Grade ≤1 after completion of steroid taper. Patients with endocrinopathies who may require prolonged or continued steroid replacement can be retreated with study drug/study regimen on the following conditions: 1. The event stabilizes and is controlled. 2. The patient is clinically stable as per Investigator or treating physician's clinical judgement. 3. Doses of prednisone are at ≤10 mg/day or equivalent. **Grade 3** Depending on the individual toxicity, study drug/study regimen may be permanently discontinued. Please refer to guidelines below. **Grade 4** Permanently discontinue study drug/study regimen. Note: For Grade ≥3 asymptomatic amylase or lipase levels, hold study drug/study regimen, and if complete work up shows no evidence of pancreatitis, study drug/study regimen may be continued or resumed.	It is recommended that management of immune-mediated adverse events (imAEs) follows the guidelines presented in this table: – It is possible that events with an inflammatory or immune mediated mechanism could occur in nearly all organs, some of them not noted specifically in these guidelines. – Whether specific immune-mediated events (and/or laboratory indicators of such events) are noted in these guidelines or not, patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression, concomitant medications, and infections) to a possible immune-mediated event. In the absence of a clear alternative etiology, all such events should be managed as if they were immune related. General recommendations follow. – Symptomatic and topical therapy should be considered for low-grade (Grade 1 or 2, unless otherwise specified) events. – For persistent (>3 to 5 days) low-grade (Grade 2) or severe (Grade ≥3) events, promptly start prednisone 1 to 2 mg/kg/day PO or IV equivalent. – Some events with high likelihood for morbidity and/or mortality – e.g., myo-carditis, or other similar events even if they are not currently noted in the guidelines – should progress rapidly to high dose IV corticosteroids (methylprednisolone at 2 to 4 mg/kg/day) even if the event is Grade 2, and if clinical suspicion is high and/or there has been clinical confirmation. Consider, as necessary, discussing with the study physician, and promptly pursue specialist consultation. – If symptoms recur or worsen during corticosteroid tapering (28 days of taper), increase the corticosteroid dose (prednisone dose [e.g., up to 2 to 4 mg/kg/day PO or IV equivalent]) until stabilization or improvement of symptoms, then resume corticosteroid tapering at a slower rate (>28 days of taper). – More potent immunosuppressives such as TNF inhibitors (e.g., infliximab) (also refer to the individual sections of the imAEs for specific type of immunosuppressive) should be considered for events

Dosing Modification and Toxicity Management Guidelines for Immune-Mediated, Infusion-Related, and Non-Immune–Mediated Reactions (MEDI4736 Monotherapy or Combination Therapy With Tremelimumab or Tremelimumab Monotherapy) 1 November 2017 Version

General Considerations

Dose Modifications	Toxicity Management
Note: Study drug/study regimen should be permanently discontinued in Grade 3 events with high likelihood for morbidity and/or mortality – e.g., myocarditis, or other similar events even if they are not currently noted in the guidelines. Similarly, consider whether study drug/study regimen should be permanently discontinued in Grade 2 events with high likelihood for morbidity and/or mortality – e.g., myocarditis, or other similar events even if they are not currently noted in the guidelines – when they do not rapidly improve to Grade <1 upon treatment with systemic steroids and following full taper Note: There are some exceptions to permanent discontinuation of study drug for Grade 4 events (i.e., hyperthyroidism, hypothyroidism, Type 1 diabetes mellitus).	not responding to systemic steroids. Progression to use of more potent immunosuppressives should proceed more rapidly in events with high likelihood for morbidity and/or mortality – e.g., myocarditis, or other similar events even if they are not currently noted in the guidelines – when these events are not responding to systemic steroids. – With long-term steroid and other immunosuppressive use, consider need for *Pneumocystis jirovecii* pneumonia (PJP, formerly known as *Pneumocystis carinii* pneumonia) prophylaxis, gastrointestinal protection, and glucose monitoring. – Discontinuation of study drug/study regimen is not mandated for Grade 3/Grade 4 inflammatory reactions attributed to local tumor response (e.g., inflammatory reaction at sites of metastatic disease and lymph nodes). Continuation of study drug/study regimen in this situation should be based upon a benefit-risk analysis for that patient.

AE Adverse event; CTC Common Toxicity Criteria; CTCAE Common Terminology Criteria for Adverse Events; imAE immune-mediated adverse event; IV intravenous; NCI National Cancer Institute; PO By mouth.

Pediatric Considerations

Dose Modifications	Toxicity Management
The criteria for permanent discontinuation of study drug/study regimen based on CTC grade/severity is the same for pediatric patients as it is for adult patients, as well as to permanently discontinue study drug/study regimen if unable to reduce corticosteroid ≤ a dose equivalent to that required for corticosteroid replacement therapy **within 12 weeks** after last dose of study drug/study regimen	– All recommendations for specialist consultation should occur with a pediatric specialist in the specialty recommended. – The recommendations for dosing of steroids (i.e., mg/kg/day) and for IV IG and plasmapheresis that are provided for adult patients should also be used for pediatric patients. – The infliximab 5 mg/kg IV dose recommended for adults is the same as recommended for pediatric patients ≥ 6 years old. For dosing in children younger than 6 years old, consult with a pediatric specialist.

Pediatric Considerations

Dose Modifications	Toxicity Management
	– For pediatric dosing of mycophenolate mofetil, consult with a pediatric specialist. – With long-term steroid and other immunosuppressive use, consider need for PJP prophylaxis, gastrointestinal protection, and glucose monitoring.

Specific Immune-Mediated Reactions

Adverse Events	Severity Grade of the Event (NCI CTCAE version 4.03)	Dose Modifications	Toxicity Management
Pneumonitis/Interstitial Lung Disease (ILD)	**Any Grade**	**General Guidance**	**For Any Grade:** – Monitor patients for signs and symptoms of pneumonitis or ILD (new onset or worsening shortness of breath or cough). Patients should be evaluated with imaging and pulmonary function tests, including other diagnostic procedures as described below. – Initial work-up may include clinical evaluation, monitoring of oxygenation via pulse oximetry (resting and exertion), laboratory work-up, and high- resolution CT scan.
	Grade 1 (asymptomatic, clinical or diagnostic observations only; intervention not indicated)	No dose modifications required. However, consider holding study drug/study regimen dose as clinically appropriate and during diagnostic work-up for other etiologies.	**For Grade 1 (radiographic changes only):** – Monitor and closely follow up in 2 to 4 days for clinical symptoms, pulse oximetry (resting and exertion), and laboratory work-up and then as clinically indicated. – Consider Pulmonary and Infectious disease consult.
	Grade 2 (symptomatic; medical intervention indicated; limiting instrumental ADL)	Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1. • If toxicity worsens, then treat as Grade 3 or Grade 4. • If toxicity improves to Grade ≤1, then the decision to reinitiate study drug/study regimen will be based upon treating physician's clinical judgment and after completion of steroid taper.	**For Grade 2 (mild to moderate new symptoms):** – Monitor symptoms daily and consider hospitalization. – Promptly start systemic steroids (e.g., prednisone 1 to 2 mg/kg/day PO or IV equivalent). – Reimage as clinically indicated. – If no improvement within 3 to 5 days, additional workup should be considered and prompt treatment with IV methylprednisolone 2 to 4 mg/kg/day started – If still no improvement within 3 to 5 days despite IV methylprednisolone at 2 to 4 mg/kg/day, promptly start immunosuppressive therapy such as TNF inhibitors (e.g., infliximab at 5 mg/kg every 2 weeks). Caution: It is important to rule out sepsis

			and refer to infliximab label for general guidance before using infliximab. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, or anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation])[a] – Consider pulmonary and infectious disease consult. – Consider, as necessary, discussing with study physician.
	Grade 3 or 4 (Grade 3: severe symptoms; limiting self-care ADL; oxygen indicated) (Grade 4: life-threatening respiratory compromise; urgent intervention indicated [e.g., tracheostomy or intubation])	Permanently discontinue study drug/study regimen.	**For Grade 3 or 4 (severe or new symptoms, new/worsening hypoxia, life-threatening):** – Promptly initiate empiric IV methylprednisolone 1 to 4 mg/kg/day or equivalent. – Obtain Pulmonary and Infectious disease consult; consider, as necessary, discussing with study physician. – Hospitalize the patient. – Supportive care (e.g., oxygen). – If no improvement within 3 to 5 days, additional workup should be considered and prompt treatment with additional immunosuppressive therapy such as TNF inhibitors (e.g., infliximab at 5 mg/kg every 2 weeks' dose) started. Caution: rule out sepsis and refer to infliximab label for general guidance before using infliximab. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and, in particular, anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Diarrhea/Colitis	**Any Grade**	**General Guidance**	**For Any Grade:** – Monitor for symptoms that may be related to diarrhea/enterocolitis (abdominal pain, cramping, or changes in bowel habits such as increased frequency over baseline or blood in stool) or

		related to bowel perforation (such as sepsis, peritoneal signs, and ileus). – Patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression, other medications, or infections), including testing for clostridium difficile toxin, etc. – Steroids should be considered in the absence of clear alternative etiology, even for low-grade events, in order to prevent potential progression to higher grade event. – Use analgesics carefully; they can mask symptoms of perforation and peritonitis.
Grade 1 (Diarrhea: stool frequency of <4 over baseline per day) (Colitis: asymptomatic; clinical or diagnostic observations only)	No dose modifications.	**For Grade 1:** – Monitor closely for worsening symptoms. – Consider symptomatic treatment, including hydration, electrolyte replacement, dietary changes (e.g., American Dietetic Association colitis diet), and loperamide. Use probiotics as per treating physician's clinical judgment.
Grade 2 (Diarrhea: stool frequency of 4 to 6 over baseline per day) (Colitis: abdominal pain; mucus or blood in stool)	Hold study drug/study regimen until resolution to Grade ≤1 • If toxicity worsens, then treat as Grade 3 or Grade 4. • If toxicity improves to Grade ≤1, then study drug/study regimen can be resumed after completion of steroid taper.	**For Grade 2:** – Consider symptomatic treatment, including hydration, electrolyte replacement, dietary changes (e.g., American Dietetic Association colitis diet), and loperamide and/or budesonide. – Promptly start prednisone 1 to 2 mg/kg/day PO or IV equivalent. – If event is not responsive within 3 to 5 days or worsens despite prednisone at 1 to 2 mg/kg/day PO or IV equivalent, GI consult should be obtained for consideration of further workup, such as imaging and/or colonoscopy, to confirm colitis and rule out perforation, and prompt treatment with IV methylprednisolone 2 to 4 mg/kg/day started. – If still no improvement within 3 to 5 days despite 2 to 4 mg/kg IV methylprednisolone, promptly start immunosuppressives such as infliximab at 5 mg/kg once every 2 weeks[a]. **Caution:** it is important to

[a]ASCO Educational Book 2015 "Managing Immune Checkpoint Blocking Antibody Side Effects" by Michael Postow MD.

			rule out bowel perforation and refer to infliximab label for general guidance before using infliximab. – Consider, as necessary, discussing with study physician if no resolution to Grade ≤1 in 3 to 4 days. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
	Grade 3 or 4 (Grade 3 diarrhea: stool frequency of ≥7 over baseline per day; Grade 4 diarrhea: life threatening consequences) (Grade 3 colitis: severe abdominal pain, change in bowel habits, medi-cal intervention indi-cated, peritoneal signs; Grade 4 colitis: life-threatening consequences, urgent intervention indicated)	**Grade 3** Permanently discontinue study drug/study regimen for Grade 3 if toxicity does not improve to Grade ≤1 within 14 days; study drug/study regimen can be resumed after completion of steroid taper. **Grade 4** Permanently discontinue study drug/study regimen.	**For Grade 3 or 4:** – Promptly initiate empiric IV methylprednisolone 2 to 4 mg/kg/day or equivalent. – Monitor stool frequency and volume and maintain hydration. – Urgent GI consult and imaging and/or colonoscopy as appropriate. – If still no improvement within 3 to 5 days of IV methylprednisolone 2 to 4 mg/kg/day or equivalent, promptly start further immunosuppressives (e.g., infliximab at 5 mg/kg once every 2 weeks). **Caution**: Ensure GI consult to rule out bowel perforation and refer to infliximab label for general guidance before using infliximab. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Hepatitis (elevated LFTs) Infliximab should not be used for management of immune-related hepatitis.	**Any Grade**	**General Guidance**	**For Any Grade:** – Monitor and evaluate liver function test: AST, ALT, ALP, and TB. – Evaluate for alternative etiologies (e.g., viral hepatitis, disease progression, concomitant medications).

[b]FDA Liver Guidance Document 2009 Guidance for Industry: Drug Induced Liver Injury – Premarketing Clinical Evaluation.

Grade 1 (AST or ALT >ULN and ≤3.0×ULN and/or TB > ULN and ≤1.5×ULN)	• No dose modifications. • If it worsens, then treat as Grade 2 event.	**For Grade 1:** – Continue LFT monitoring per protocol.
Grade 2 (AST or ALT >3.0×ULN and ≤5.0×ULN and/or TB >1.5×ULN and ≤3.0×ULN)	• Hold study drug/study regimen dose until Grade 2 resolution to Grade ≤1. • If toxicity worsens, then treat as Grade 3 or Grade 4. • If toxicity improves to Grade ≤1 or baseline, resume study drug/study regimen after completion of steroid taper.	**For Grade 2:** – Regular and frequent checking of LFTs (e.g., every 1 to 2 days) until elevations of these are improving or resolved. – If no resolution to Grade ≤1 in 1 to 2 days, consider, as necessary, discussing with study physician. – If event is persistent (>3 to 5 days) or worsens, promptly start prednisone 1 to 2 mg/kg/day PO or IV equivalent. – If still no improvement within 3 to 5 days despite 1 to 2 mg/kg/day of prednisone PO or IV equivalent, consider additional work up and start prompt treatment with IV methylprednisolone 2 to 4 mg/kg/day. – If still no improvement within 3 to 5 days despite 2 to 4 mg/kg/day of IV methylprednisolone, promptly start immunosuppressives (i.e., mycophenolate mofetil).[a] Discuss with study physician if mycophenolate mofetil is not available. **Infliximab should NOT be used.** – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Grade 3 or 4 (Grade 3: AST or ALT >5.0×ULN and ≤20.0×ULN and/or TB >3.0×ULN and ≤10.0×ULN)	**For Grade 3:** For elevations in transaminases ≤8 × ULN, or elevations in bilirubin ≤5 × ULN:	**For Grade 3 or 4:** – Promptly initiate empiric IV methylprednisolone at 1 to 4 mg/kg/day or equivalent. – If still no improvement within 3 to 5 days despite 1 to 4 mg/kg/day methylprednisolone IV or equivalent, promptly start treatment with immunosuppressive therapy (i.e., mycophenolate

	(Grade 4: AST or ALT >20×ULN and/or TB >10×ULN)	• Hold study drug/study regimen dose until resolution to Grade ≤1 or baseline • Resume study drug/study regimen if elevations downgrade to Grade ≤1 or baseline within 14 days and after completion of steroid taper. • Permanently discontinue study drug/study regimen if the elevations do not downgrade to Grade ≤1 or baseline within 14 days For elevations in transaminases >8 × ULN or elevations in bilirubin >5 × ULN, discontinue study drug/study regimen. Permanently discontinue study drug/study regimen for any case meeting Hy's law criteria (AST and/or ALT >3 × ULN + bilirubin >2 × ULN without initial findings of cholestasis (i.e., elevated alkaline P04) and in the absence of any alternative cause.[b] **For Grade 4:** Permanently discontinue study drug/study regimen.	mofetil). Discuss with study physician if mycophenolate is not available. **Infliximab should NOT be used.** – Perform hepatology consult, abdominal workup, and imaging as appropriate. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Nephritis or renal dysfunction (elevated serum creatinine)	**Any Grade**	**General Guidance**	**For Any Grade:** – Consult with nephrologist. – Monitor for signs and symptoms that may be related to changes in renal function (e.g., routine urinalysis, elevated serum BUN and creatinine, decreased creatinine clearance, electrolyte

		imbalance, decrease in urine output, or proteinuria). – Patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression or infections). – Steroids should be considered in the absence of clear alternative etiology even for low-grade events (Grade 2), in order to prevent potential progression to higher grade event.
Grade 1 (Serum creatinine > 1 to 1.5 × baseline; > ULN to 1.5 × ULN)	No dose modifications.	**For Grade 1:** – Monitor serum creatinine weekly and any accompanying symptoms. • If creatinine returns to baseline, resume its regular monitoring per study protocol. • If creatinine worsens, depending on the severity, treat as Grade 2, 3, or 4. – Consider symptomatic treatment, including hydration, electrolyte replacement, and diuretics.
Grade 2 (serum creatinine >1.5 to 3.0 × baseline; >1.5 to 3.0 × ULN)	Hold study drug/study regimen until resolution to Grade ≤1 or baseline. • If toxicity worsens, then treat as Grade 3 or 4. • If toxicity improves to Grade ≤1 or baseline, then resume study drug/study regimen after completion of steroid taper.	**For Grade 2:** – Consider symptomatic treatment, including hydration, electrolyte replacement, and diuretics. – Carefully monitor serum creatinine every 2 to 3 days and as clinically warranted. – Consult nephrologist and consider renal biopsy if clinically indicated. – If event is persistent (>3 to 5 days) or worsens, promptly start prednisone 1 to 2 mg/kg/day PO or IV equivalent. – If event is not responsive within 3 to 5 days or worsens despite prednisone at 1 to 2 mg/kg/day PO or IV equivalent, additional workup should be considered and prompt treatment with IV methylprednisolone at 2 to 4 mg/kg/day started. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic

			antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a] – When event returns to baseline, resume study drug/study regimen and routine serum creatinine monitoring per study protocol.
	Grade 3 or 4 (Grade 3: serum creatinine >3.0 × baseline; >3.0 to 6.0 × ULN; Grade 4: serum creatinine >6.0 × ULN)	Permanently discontinue study drug/study regimen.	**For Grade 3 or 4:** – Carefully monitor serum creatinine on daily basis. – Consult nephrologist and consider renal biopsy if clinically indicated. – Promptly start prednisone 1 to 2 mg/kg/day PO or IV equivalent. – If event is not responsive within 3 to 5 days or worsens despite prednisone at 1 to 2 mg/kg/day PO or IV equivalent, additional workup should be considered and prompt treatment with IV methylprednisolone 2 to 4 mg/kg/day started. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Rash (excluding bullous skin formations)	**Any Grade** (refer to NCI CTCAE v 4.03 for definition of severity/grade depending on type of skin rash)	**General Guidance**	**For Any Grade:** – Monitor for signs and symptoms of dermatitis (rash and pruritus). – IF THERE IS ANY BULLOUS FORMATION, THE STUDY PHYSICIAN SHOULD BE CONTACTED **AND STUDY DRUG DISCONTINUED.**
	Grade 1	No dose modifications.	**For Grade 1:** – Consider symptomatic treatment, including oral antipruritics (e.g., diphenhydramine or hydroxyzine) and topical therapy (e.g., urea cream).

	Grade 2	For persistent (>1 to 2 weeks) Grade 2 events, hold scheduled study drug/study regimen until resolution to Grade ≤1 or baseline. • If toxicity worsens, then treat as Grade 3. • If toxicity improves to Grade ≤1 or baseline, then resume drug/study regimen after completion of steroid taper.	**For Grade 2:** – Obtain dermatology consult. – Consider symptomatic treatment, including oral antipruritics (e.g., diphenhydramine or hydroxyzine) and topical therapy (e.g., urea cream). – Consider moderate-strength topical steroid. – If no improvement of rash/skin lesions occurs within 3 to 5 days or is worsening despite symptomatic treatment and/or use of moderate strength topical steroid, consider, as necessary, discussing with study physician and promptly start systemic steroids such as prednisone 1 to 2 mg/kg/day PO or IV equivalent. – Consider skin biopsy if the event is persistent for >1 to 2 weeks or recurs.
	Grade 3 or 4	**For Grade 3:** Hold study drug/study regimen until resolution to Grade ≤1 or baseline. If temporarily holding the study drug/study regimen does not provide improvement of the Grade 3 skin rash to Grade ≤1 or baseline within 30 days, then permanently discontinue study drug/study regimen. **For Grade 4:** Permanently discontinue study drug/study regimen.	**For Grade 3 or 4:** – Consult dermatology. – Promptly initiate empiric IV methylprednisolone 1 to 4 mg/kg/day or equivalent. – Consider hospitalization. – Monitor extent of rash [Rule of Nines]. – Consider skin biopsy (preferably more than 1) as clinically feasible. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a] – Consider, as necessary, discussing with study physician.
Endocrinopathy (e.g., hyperthyroidism, hypothyroidism, Type 1	**Any Grade** (depending on the type of endocrinopathy, refer to NCI	**General Guidance**	**For Any Grade:**

diabetes mellitus, hypophysitis, hypopituitarism, and adrenal insufficiency; exocrine event of amylase/lipase increased also included in this section)	CTCAE v4.03 for defining the CTC grade/severity)		– Consider consulting an endocrinologist for endocrine events. – Consider, as necessary, discussing with study physician. – Monitor patients for signs and symptoms of endocrinopathies. Non-specific symptoms include headache, fatigue, behavior changes, changed mental status, vertigo, abdominal pain, unusual bowel habits, polydipsia, polyuria, hypotension, and weakness. – Patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression including brain metastases, or infections). – Depending on the suspected endocrinopathy, monitor and evaluate thyroid function tests: TSH, free T3 and free T4 and other relevant endocrine and related labs (e.g., blood glucose and ketone levels, HgA1c). – For modest asymptomatic elevations in serum amylase and lipase, corticosteroid treatment is not indicated as long as there are no other signs or symptoms of pancreatic inflammation. – If a patient experiences an AE that is thought to be possibly of autoimmune nature (e.g., thyroiditis, pancreatitis, hypophysitis, or diabetes insipidus), the investigator should send a blood sample for appropriate autoimmune antibody testing.
	Grade 1	No dose modifications.	**For Grade 1 (including those with asymptomatic TSH elevation):** – Monitor patient with appropriate endocrine function tests. – For suspected hypophysitis/hypopituitarism, consider consultation of an endocrinologist to guide assessment of early-morning ACTH, cortisol, TSH and free T4; also consider gonadotropins, sex hormones, and prolactin levels, as well as cosyntropin stimulation test (though it may not be

		useful in diagnosing early secondary adrenal insufficiency). – If TSH < 0.5 × LLN, or TSH >2 × ULN, or consistently out of range in 2 subsequent measurements, include free T4 at subsequent cycles as clinically indicated and consider consultation of an endocrinologist.
Grade 2	For Grade 2 endocrinopathy other than hypothyroidism and Type 1 diabetes mellitus, hold study drug/study regimen dose until patient is clinically stable. • If toxicity worsens, then treat as Grade 3 or Grade 4. Study drug/study regimen can be resumed once event stabilizes and after completion of steroid taper. Patients with endocrinopathies who may require prolonged or continued steroid replacement (e.g., adrenal insufficiency) can be retreated with study drug/study regimen on the following conditions: 1. The event stabilizes and is controlled. 2. The patient is clinically stable as per investigator or treating physician's clinical judgement. 3. Doses of prednisone are ≤10 mg/day or equivalent.	**For Grade 2 (including those with symptomatic endocrinopathy):** – Consult endocrinologist to guide evaluation of endocrine function and, as indicated by suspected endocrinopathy and as clinically indicated, consider pituitary scan. – For all patients with abnormal endocrine work up, except those with isolated hypothyroidism or Type 1 DM, and as guided by an endocrinologist, consider short-term corticosteroids (e.g., 1 to 2 mg/kg/day methylprednisolone or IV equivalent) and prompt initiation of treatment with relevant hormone replacement (e.g., hydrocortisone, sex hormones). – Isolated hypothyroidism may be treated with replacement therapy, without study drug/study regimen interruption, and without corticosteroids. – Isolated Type 1 diabetes mellitus (DM) may be treated with appropriate diabetic therapy, without study drug/study regimen interruption, and without corticosteroids. – Once patients on steroids are improving, gradually taper immunosuppressive steroids (as appropriate and with guidance of endocrinologist) over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a] – For patients with normal endocrine workup (laboratory assessment or MRI scans), repeat laboratory assessments/MRI as clinically indicated.

	Grade 3 or 4	For Grade 3 or 4 endocrinopathy other than hypothyroidism and Type 1 diabetes mellitus, hold study drug/study regimen dose until endocrinopathy symptom(s) are controlled. Study drug/study regimen can be resumed once event stabilizes and after completion of steroid taper. Patients with endocrinopathies who may require prolonged or continued steroid replacement (e.g., adrenal insufficiency) can be retreated with study drug/study regimen on the following conditions: 1. The event stabilizes and is controlled. 2. The patient is clinically stable as per investigator or treating physician's clinical judgement. 3. Doses of prednisone are ≤10 mg/day or equivalent.	**For Grade 3 or 4:** – Consult endocrinologist to guide evaluation of endocrine function and, as indicated by suspected endocrinopathy and as clinically indicated, consider pituitary scan. Hospitalization recommended. – For all patients with abnormal endocrine work up, except those with isolated hypothyroidism or Type 1 DM, and as guided by an endocrinologist, promptly initiate empiric IV methylprednisolone 1 to 2 mg/kg/day or equivalent, as well as relevant hormone replacement (e.g., hydrocortisone, sex hormones). – For adrenal crisis, severe dehydration, hypotension, or shock, immediately initiate IV corticosteroids with mineralocorticoid activity. – Isolated hypothyroidism may be treated with replacement therapy, without study drug/study regimen interruption, and without corticosteroids. – Isolated Type 1 diabetes mellitus may be treated with appropriate diabetic therapy, without study drug/study regimen interruption, and without corticosteroids. – Once patients on steroids are improving, gradually taper immunosuppressive steroids (as appropriate and with guidance of endocrinologist) over ≥28 days and consider prophylactic antibiotics, antifungals, and anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Neurotoxicity (to include but not be limited to limbic encephalitis and autonomic neuropathy,	**Any Grade** (depending on the type of neurotoxicity, refer to NCI CTCAE v4.03 for defining the CTC grade/severity)	**General Guidance**	**For Any Grade:** – Patients should be evaluated to rule out any alternative etiology (e.g., disease progression, infections, metabolic syndromes, or medications). – Monitor patient for general symptoms (headache, nausea, vertigo, behavior change, or weakness).

excluding Myasthenia Gravis and Guillain-Barre)			– Consider appropriate diagnostic testing (e.g., electromyogram and nerve conduction investigations). – Perform symptomatic treatment with neurological consult as appropriate. –
	Grade 1	No dose modifications.	**For Grade 1:** – See "Any Grade" recommendations above.
	Grade 2	For acute motor neuropathies or neurotoxicity, hold study drug/study regimen dose until resolution to Grade ≤1. For sensory neuropathy/neuropathic pain, consider holding study drug/study regimen dose until resolution to Grade ≤1. If toxicity worsens, then treat as Grade 3 or 4. Study drug/study regimen can be resumed once event improves to Grade ≤1 and after completion of steroid taper.	**For Grade 2:** – Consider, as necessary, discussing with the study physician. – Obtain neurology consult. – Sensory neuropathy/neuropathic pain may be managed by appropriate medications (e.g., gabapentin or duloxetine). – Promptly start systemic steroids prednisone 1 to 2 mg/kg/day PO or IV equivalent. – If no improvement within 3 to 5 days despite 1 to 2 mg/kg/day prednisone PO or IV equivalent, consider additional workup and promptly treat with additional immunosuppressive therapy (e.g., IV IG).
	Grade 3 or 4	**For Grade 3:** Hold study drug/study regimen dose until resolution to Grade ≤1. Permanently discontinue study drug/study regimen if Grade 3 imAE does not resolve to Grade ≤1 within 30 days. **For Grade 4:** Permanently discontinue study drug/study regimen.	**For Grade 3 or 4:** – Consider, as necessary, discussing with study physician. – Obtain neurology consult. – Consider hospitalization. – Promptly initiate empiric IV methylprednisolone 1 to 2 mg/kg/day or equivalent. – If no improvement within 3 to 5 days despite IV corticosteroids, consider additional workup and promptly treat with additional immunosuppressants (e.g., IV IG). – Once stable, gradually taper steroids over ≥28 days.

Peripheral neuromotor syndromes (such as Guillain-Barre and myasthenia gravis)	**Any Grade**	**General Guidance**	**For Any Grade:** – The prompt diagnosis of immune-mediated peripheral neuromotor syndromes is important, since certain patients may unpredictably experience acute decompensations that can result in substantial morbidity or in the worst case, death. Special care should be taken for certain sentinel symptoms that may predict a more severe outcome, such as prominent dysphagia, rapidly progressive weakness, and signs of respiratory insufficiency or autonomic instability. – Patients should be evaluated to rule out any alternative etiology (e.g., disease progression, infections, metabolic syndromes or medications). It should be noted that the diagnosis of immune-mediated peripheral neuromotor syndromes can be particularly challenging in patients with underlying cancer, due to the multiple potential confounding effects of cancer (and its treatments) throughout the neuraxis. Given the importance of prompt and accurate diagnosis, it is essential to have a low threshold to obtain a neurological consult. – Neurophysiologic diagnostic testing (e.g., electromyogram and nerve conduction investigations, and "repetitive stimulation" if myasthenia is suspected) are routinely indicated upon suspicion of such conditions and may be best facilitated by means of a neurology consultation. – It is important to consider that the use of steroids as the primary treatment of Guillain-Barre is not typically considered effective. Patients requiring treatment should be started with IV IG and followed by plasmapheresis if not responsive to IV IG.
	Grade 1	No dose modifications.	**For Grade 1:** – Consider, as necessary, discussing with the study physician.

		– Care should be taken to monitor patients for sentinel symptoms of a potential decompensation as described above. – Obtain a neurology consult.
Grade 2	Hold study drug/study regimen dose until resolution to Grade ≤1. Permanently discontinue study drug/study regimen if it does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency or autonomic instability.	**For Grade 2:** – Consider, as necessary, discussing with the study physician. – Care should be taken to monitor patients for sentinel symptoms of a potential decompensation as described above. – Obtain a neurology consult – Sensory neuropathy/neuropathic pain may be managed by appropriate medications (e.g., gabapentin or duloxetine). *MYASTHENIA GRAVIS:* o Steroids may be successfully used to treat myasthenia gravis. It is important to consider that steroid therapy (especially with high doses) may result in transient worsening of myasthenia and should typically be administered in a monitored setting under supervision of a consulting neurologist. o Patients unable to tolerate steroids may be candidates for treatment with plasmapheresis or IV IG. Such decisions are best made in consultation with a neurologist, taking into account the unique needs of each patient. o If myasthenia gravis-like neurotoxicity is present, consider starting AChE inhibitor therapy in addition to steroids. Such therapy, if successful, can also serve to reinforce the diagnosis. *GUILLAIN-BARRE:* o It is important to consider here that the use of steroids as the primary treatment

			of Guillain-Barre is not typically considered effective. o Patients requiring treatment should be started with IV IG and followed by plasmapheresis if not responsive to IV IG.
	Grade 3 or 4	**For Grade 3:** Hold study drug/study regimen dose until resolution to Grade ≤1. Permanently discontinue study drug/study regimen if Grade 3 imAE does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency or autonomic instability. **For Grade 4:** Permanently discontinue study drug/study regimen.	**For Grade 3 or 4 (severe or life-threatening events):** – Consider, as necessary, discussing with study physician. – Recommend hospitalization. – Monitor symptoms and obtain neurological consult. *MYASTHENIA GRAVIS:* o Steroids may be successfully used to treat myasthenia gravis. They should typically be administered in a monitored setting under supervision of a consulting neurologist. o Patients unable to tolerate steroids may be candidates for treatment with plasmapheresis or IV IG. o If myasthenia gravis-like neurotoxicity present, consider starting AChE inhibitor therapy in addition to steroids. Such therapy, if successful, can also serve to reinforce the diagnosis. *GUILLAIN-BARRE:* o It is important to consider here that the use of steroids as the primary treatment of Guillain-Barre is not typically considered effective. o Patients requiring treatment should be started with IV IG and followed by plasmapheresis if not responsive to IV IG.
Myocarditis	**Any Grade**	**General Guidance**	**For Any Grade:**

	Discontinue drug permanently if biopsy-proven immune-mediated myocarditis.	– The prompt diagnosis of immune-mediated myocarditis is important, particularly in patients with baseline cardiopulmonary disease and reduced cardiac function. – Consider, as necessary, discussing with the study physician. – Monitor patients for signs and symptoms of myocarditis (new onset or worsening chest pain, arrhythmia, shortness of breath, peripheral edema). As some symptoms can overlap with lung toxicities, simultaneously evaluate for and rule out pulmonary toxicity as well as other causes (e.g., pulmonary embolism, congestive heart failure, malignant pericardial effusion). A Cardiology consultation should be obtained early, with prompt assessment of whether and when to complete a cardiac biopsy, including any other diagnostic procedures. – Initial work-up should include clinical evaluation, BNP, cardiac enzymes, ECG, echocardiogram (ECHO), monitoring of oxygenation via pulse oximetry (resting and exertion), and additional laboratory work-up as indicated. Spiral CT or cardiac MRI can complement ECHO to assess wall motion abnormalities when needed. – Patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression, other medications, or infections)
Grade 1 (asymptomatic with laboratory (e.g., BNP) or cardiac imaging abnormalities)	No dose modifications required unless clinical suspicion is high, in which case hold study drug/study regimen dose during diagnostic work-up for other etiologies. If study drug/study regimen is held, resume after complete resolution to Grade 0.	**For Grade 1 (no definitive findings):** - Monitor and closely follow up in 2 to 4 days for clinical symptoms, BNP, cardiac enzymes, ECG, ECHO, pulse oximetry (resting and exertion), and laboratory work-up as clinically indicated. - Consider using steroids if clinical suspicion is high.
Grade 2, 3 or 4	- If Grade 2 -- Hold study drug/study regimen dose until	**For Grade 2-4:** – Monitor symptoms daily, hospitalize.

	(Grade 2: Symptoms with mild to moderate activity or exertion) (Grade 3: Severe with symptoms at rest or with minimal activity or exertion; intervention indicated) (Grade 4: Life-threatening consequences; urgent intervention indicated (e.g., continuous IV therapy or mechanical hemodynamic support))	resolution to Grade 0. If toxicity rapidly improves to Grade 0, then the decision to reinitiate study drug/study regimen will be based upon treating physician's clinical judgment and after completion of steroid taper. If toxicity does not rapidly improve, permanently discontinue study drug/study regimen. If Grade 3-4, permanently discontinue study drug/study regimen.	– Promptly start IV methylprednisolone 2 to 4 mg/kg/day or equivalent after Cardiology consultation has determined whether and when to complete diagnostic procedures including a cardiac biopsy. – Supportive care (e.g., oxygen). – If no improvement within 3 to 5 days despite IV methylprednisolone at 2 to 4 mg/kg/day, promptly start immunosuppressive therapy such as TNF inhibitors (e.g., infliximab at 5 mg/kg every 2 weeks). Caution: It is important to rule out sepsis and refer to infliximab label for general guidance before using infliximab. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, or anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Myositis/Polymyositis ("Poly/myositis")	**Any Grade**	**General Guidance**	**For Any Grade:** – Monitor patients for signs and symptoms of poly/myositis. Typically, muscle weakness/pain occurs in proximal muscles including upper arms, thighs, shoulders, hips, neck and back, but rarely affects the extremities including hands and fingers; also difficulty breathing and/or trouble swallowing can occur and progress rapidly. Increased general feelings of tiredness and fatigue may occur, and there can be new-onset falling, difficulty getting up from a fall, and trouble climbing stairs, standing up from a seated position, and/or reaching up. – If poly/myositis is suspected, a Neurology consultation should be obtained early, with prompt guidance on diagnostic procedures. Myocarditis may co-occur with poly/myositis; refer to guidance under Myocarditis. Given breathing complications, refer to guidance under Pneumonitis/ILD. Given possibility of an existent (but previously unknown) autoimmune disorder, consider Rheumatology consultation.

		– Consider, as necessary, discussing with the study physician. – Initial work-up should include clinical evaluation, creatine kinase, aldolase, LDH, BUN/creatinine, erythrocyte sedimentation rate or C-reactive protein level, urine myoglobin, and additional laboratory work-up as indicated, including a number of possible rheumatological/antibody tests (i.e., consider whether a rheumatologist consultation is indicated and could guide need for rheumatoid factor, antinuclear antibody, anti-smooth muscle, antisynthetase [such as anti-Jo-1], and/or signal-recognition particle antibodies). Confirmatory testing may include electromyography, nerve conduction studies, MRI of the muscles, and/or a muscle biopsy. Consider Barium swallow for evaluation of dysphagia or dysphonia. Patients should be thoroughly evaluated to rule out any alternative etiology (e.g., disease progression, other medications, or infections).
Grade 1 (mild pain)	- No dose modifications.	**For Grade 1:** – Monitor and closely follow up in 2 to 4 days for clinical symptoms and initiate evaluation as clinically indicated. – Consider Neurology consult. – Consider, as necessary, discussing with the study physician.
Grade 2 (moderate pain associated with weakness; pain limiting instrumental activities of daily living [ADLs])	Hold study drug/study regimen dose until resolution to Grade ≤1. - Permanently discontinue study drug/study regimen if it does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency.	**For Grade 2:** – Monitor symptoms daily and consider hospitalization. – Obtain Neurology consult, and initiate evaluation. – Consider, as necessary, discussing with the study physician. – If clinical course is rapidly progressive (particularly if difficulty breathing and/or trouble swallowing), promptly start IV methylprednisolone

		2 to 4 mg/kg/day systemic steroids along with receiving input from Neurology consultant – If clinical course is *not* rapidly progressive, start systemic steroids (e.g., prednisone 1 to 2 mg/kg/day PO or IV equivalent); if no improvement within 3 to 5 days, continue additional work up and start treatment with IV methylprednisolone 2 to 4 mg/kg/day – If after start of IV methylprednisolone at 2 to 4 mg/kg/day there is no improvement within 3 to 5 days, consider start of immunosuppressive therapy such as TNF inhibitors (e.g., infliximab at 5 mg/kg every 2 weeks). Caution: It is important to rule out sepsis and refer to infliximab label for general guidance before using infliximab. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, or anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]
Grade 3 or 4 (pain associated with severe weakness; limiting self-care ADLs)	**For Grade 3:** Hold study drug/study regimen dose until resolution to Grade ≤1. Permanently discontinue study drug/study regimen if Grade 3 imAE does not resolve to Grade ≤1 within 30 days or if there are signs of respiratory insufficiency. **For Grade 4:** - Permanently discontinue study drug/study regimen.	**For Grade 3 or 4 (severe or life-threatening events):** – Monitor symptoms closely; recommend hospitalization. – Obtain Neurology consult, and complete full evaluation. – Consider, as necessary, discussing with the study physician. – Promptly start IV methylprednisolone 2 to 4 mg/kg/day systemic steroids along with receiving input from Neurology consultant. – If after start of IV methylprednisolone at 2 to 4 mg/kg/day there is no improvement within 3 to 5 days, consider start of immunosuppressive therapy such as TNF inhibitors (e.g., infliximab at 5 mg/kg every 2 weeks). Caution: It is important to rule out sepsis and refer to infliximab label for general guidance before using infliximab.

	– Consider whether patient may require IV IG, plasmapheresis. – Once the patient is improving, gradually taper steroids over ≥28 days and consider prophylactic antibiotics, antifungals, or anti-PJP treatment (refer to current NCCN guidelines for treatment of cancer-related infections [Category 2B recommendation]).[a]

AChE Acetylcholine esterase; ADL Activities of daily living; AE Adverse event; ALP Alkaline phosphatase test; ALT Alanine aminotransferase; AST Aspartate aminotransferase; BUN Blood urea nitrogen; CT Computed tomography; CTCAE Common Terminology Criteria for Adverse Events; ILD Interstitial lung disease; imAE immune-mediated adverse event; IG Immunoglobulin; IV Intravenous; GI Gastrointestinal; LFT Liver function tests; LLN Lower limit of normal; MRI Magnetic resonance imaging; NCI National Cancer Institute; NCCN National Comprehensive Cancer Network; PJP *Pneumocystis jirovecii* pneumonia (formerly known as *Pneumocystis carinii* pneumonia); PO By mouth; T3 Triiodothyronine; T4 Thyroxine; TB Total bilirubin; TNF Tumor necrosis factor; TSH Thyroid-stimulating hormone; ULN Upper limit of normal.

Infusion-Related Reactions

Severity Grade of the Event (NCI CTCAE version 4.03)	Dose Modifications	Toxicity Management
Any Grade	General Guidance	**For Any Grade:** – Manage per institutional standard at the discretion of investigator. – Monitor patients for signs and symptoms of infusion-related reactions (e.g., fever and/or shaking chills, flushing and/or itching, alterations in heart rate and blood pressure, dyspnea or chest discomfort, or skin rashes) and anaphylaxis (e.g., generalized urticaria, angioedema, wheezing, hypotension, or tachycardia).
Grade 1 or 2	**For Grade 1:** The infusion rate of study drug/study regimen may be decreased by 50% or temporarily interrupted until resolution of the event. **For Grade 2:** The infusion rate of study drug/study regimen may be decreased 50% or temporarily interrupted until resolution of the event. Subsequent infusions may be given at 50% of the initial infusion rate.	**For Grade 1 or 2:** – Acetaminophen and/or antihistamines may be administered per institutional standard at the discretion of the investigator. – Consider premedication per institutional standard prior to subsequent doses. – Steroids should not be used for routine premedication of Grade ≤2 infusion reactions.
Grade 3 or 4	**For Grade 3 or 4:** Permanently discontinue study drug/study regimen.	**For Grade 3 or 4:** – Manage severe infusion-related reactions per institutional standards (e.g., IM epinephrine, followed by IV diphenhydramine and ranitidine, and IV glucocorticoid).

CTCAE Common Terminology Criteria for Adverse Events; IM intramuscular; IV intravenous; NCI National Cancer Institute.

Non–Immune-Mediated Reactions

Severity Grade of the Event (NCI CTCAE version 4.03)	Dose Modifications	Toxicity Management
Any Grade	Note: Dose modifications are not required for AEs not deemed to be related to study treatment (i.e., events due to underlying disease) or for laboratory abnormalities not deemed to be clinically significant.	Treat accordingly, as per institutional standard.
Grade 1	No dose modifications.	Treat accordingly, as per institutional standard.
Grade 2	Hold study drug/study regimen until resolution to ≤Grade 1 or baseline.	Treat accordingly, as per institutional standard.
Grade 3	Hold study drug/study regimen until resolution to ≤Grade 1 or baseline. For AEs that downgrade to ≤Grade 2 within 7 days or resolve to ≤Grade 1 or baseline within 14 days, resume study drug/study regimen administration. Otherwise, discontinue study drug/study regimen.	Treat accordingly, as per institutional standard.
Grade 4	Discontinue study drug/study regimen (Note: For Grade 4 labs, decision to discontinue should be based on accompanying clinical signs/symptoms, the Investigator's clinical judgment, and consultation with the Sponsor.).	Treat accordingly, as per institutional standard.

Note: As applicable, for early phase studies, the following sentence may be added: "Any event greater than or equal to Grade 2, please discuss with Study Physician."
AE Adverse event; CTCAE Common Terminology Criteria for Adverse Events; NCI National Cancer Institute.

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: d419ac00001-csp-8		
Document Title:	D419AC00001 Clinical Study Protocol Version 8	
Document ID:	Doc ID-003403622	
Version Label:	5.0 CURRENT LATEST APPROVED	
Server Date **(dd-MMM-yyyy HH:mm 'UTC'Z)**	**Signed by**	**Meaning of Signature**
		Content Approval
		Content Approval
		Qualified Person Approval
		Content Approval

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

Protocol amendments and other significant changes to study conduct

Number	Key details of amendment	Reason for amendment	Person(s)/ group(s) responsible for amendment[a]
Amendments made before the start of patient recruitment			
Amendment 1 Protocol version 2.0	To provide additional overall information in the toxicity management table regarding when to permanently discontinue the study treatments in the event of immune-related AEs	Based on FDA feedback	AstraZeneca
	To provide additional detail about the criteria required to continue treatment through progression	Based on FDA feedback	AstraZeneca
	To require that initial PRO assessments be completed at screening	Missed timepoint	AstraZeneca
	To provide additional detail about the assessments required during re-treatment with durvalumab monotherapy or durvalumab + tremelimumab	Re-treatment patients do require repeated collection of PK, ADA, sPD-L1, SNP genotyping, and MDSC assessments	AstraZeneca
Amendments made after the start of patient recruitment			
Amendment 2 Protocol version 3.0	To change irRECIST from a secondary to an exploratory objective	The protocol includes a partial (not full) BICR, so the analysis of irRECIST is more appropriate as an exploratory objective	AstraZeneca
	To clarify that the in-line hierarchical testing strategy is fully described in the Methods for Statistical Analysis section, while only the 3 key tests are described in the Synopsis	Based on EMA feedback	AstraZeneca
Amendment 3 Protocol version 4.0	To change the frequency of pregnancy tests to every visit for patients receiving study treatment	Based on EMA feedback	AstraZeneca

Number	Key details of amendment	Reason for amendment	Person(s)/ group(s) responsible for amendment[a]
Amendment 4 Protocol version 5.0	To update the primary and secondary objectives to reflect changes to endpoint measures; ie, BICR rather than Investigator assessments, as well as to assess the treatment benefit and efficacy of durvalumab as suggested by emerging immuno-oncology data. Investigator assessments will be used for sensitivity analysis. Text of BICR-based assessments was changed to reflect BICR-RECIST 1.1 analysis on all radiological scans of all patients rather than a random sample	Consistent with program decisions and emerging data	AstraZeneca
	Numbers of patients enrolled, randomized, and per treatment group were updated to 1850, 1092, and 364, respectively	Consistent with program decisions	AstraZeneca
	The assessments of PFS and OS were nominated as multiple primary objectives. Primary and secondary objectives, endpoints, rationale, statistical considerations, and analyses were modified accordingly	Consistent with program decisions and emerging data	AstraZeneca
	Duration of treatment was modified so that patients in all groups can continue therapy until disease progression rather than stopping at 12 months	Emerging data from ongoing durvalumab studies suggest that some patients are losing clinical benefit after they complete the 12 months of therapy	AstraZeneca
	The rationale for the re-treatment option was amended to enable patients in the durvalumab + tremelimumab group who complete 4 dosing cycles (providing clinical benefit per Investigator judgement) and subsequently have PD during treatment with durvalumab monotherapy to restart combination treatment if they also meet eligibility criteria	Consistent with program decisions	AstraZeneca
Amendment 5 Protocol version 6.0	To clarify that, in the SoC chemotherapy group, treatment can be continued, at Investigator's discretion, until disease progression is confirmed. Patients in the SoC chemotherapy group are not allowed to continue treatment once disease progression is confirmed	Clarification	AstraZeneca

Number	Key details of amendment	Reason for amendment	Person(s)/ group(s) responsible for amendment[a]
Amendment 6 Protocol version 7.0	The definitions for the PD-L1 tumor membrane expression analysis sets are provided For durvalumab + tremelimumab, the assessment of PFS and OS were nominated as multiple primary objectives in NSCLC patients with ≥25% PD-L1 membrane expression in tumoral tissue For durvalumab monotherapy, the assessment of OS is nominated as a primary objective in NSCLC patients with ≥25% PD-L1 membrane expression in tumoral tissue	Emerging science suggests the PD-L1 high population may experience the largest benefit. This change ensures this population, for which patients were stratified at enrollment, is tested first in the statistical hierarchy (Hellmann et al 2016, Reck et al 2014, Socinski et al 2016). A recent Phase III study in the first line Stage IV NSCLC population has demonstrated a clinically relevant and statistically significant improvement in OS for monotherapy; this change ensures a comparable analysis is performed in MYSTIC.	AstraZeneca
Amendment 7 Protocol version 8.0	Survival status for withdrawn consent and lost to follow-up patients were updated	To clarify which analysis sets require survival status data and how and when to obtain these data.	AstraZeneca
	Potential risks and AESIs were updated	To align with the most current IB.	AstraZeneca
	Appendix G was updated.	To include the most current version of the dosing modification and toxicity management guidelines.	AstraZeneca

[a] All protocol amendments were approved by AstraZeneca before being submitted to a regulatory authority and/or an IRB/IEC.

ADA anti-drug antibody; AE adverse event; AESI adverse event of special interest; BICR Blinded Independent Central Review; EMA European Medicines Agency; FDA Food and Drug Administration; IB investigator's brochure; IEC Independent Ethics Committee; IRB Independent Review Board; irRECIST immune-related Response Evaluation Criteria in Solid

Tumors, version 1.1; MDSC myeloid-derived suppressor cell; NSCLC non-small cell lung cancer; OS overall survival; PD progressive disease; PD-L1 programmed cell death ligand 1; PFS progression-free survival; PK pharmacokinetic; PRO patient-reported outcomes; RECIST 1.1 Response Evaluation Criteria in Solid Tumors, version 1.1; SNP single nucleotide polymorphism; SoC standard of care; sPD-L1 soluble programmed cell death ligand 1.

AstraZeneca

Statistical Analysis Plan

Study Code D419AC00001

Edition Number 1.0

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
ALK	Anaplastic lymphoma kinase
APF12	Proportion of patients alive and progression free at 12 months from first dose
Baseline	Refers to the most recent assessment of any variable prior to dosing with study treatment/randomisation (as appropriate)
BICR	Blinded independent central review
BoR	Best objective response
CI	Confidence Interval
CR	Complete Response
CRF/eCRF	Case Report Form (electronic)
CSR	Clinical Study Report
CTC/CTCAE	Common Terminology Criteria for Adverse Event (National Institutes of Health, National Cancer Institute)
DoR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDoR	Expected Duration of Response
EGFR	Epidermal growth factor receptor
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire
EORTC QLQ-LC13	European Organisation for Research and Treatment of Cancer 13-item lung cancer-specific quality of life questionnaire. Module used as a supplement to EORTC QLQ-C30
EQ-5D	EuroQoL 5-dimension utility index
EQ-5D-3L	EuroQoL 5-dimension, 3-level health state utility index
EQ-5D-5L	EuroQoL 5-dimension, 5-level health state utility index
FAS	Full analysis set
HR	Hazard ratio
HRQoL	Health-related Quality of Life
IDMC	Independent data monitoring committee

Abbreviation or special term	Explanation
IP	Investigational Product
ITT	Intention to Treat
IVRS	Interactive Voice Response System
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MRI	Magnetic Resonance Imaging
MTP	Multiple testing procedure
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
NSCLC	Non-small cell lung cancer
NTL	Non-target lesions
OAE	Other Significant Adverse Event
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive disease
PD-L1	Programmed death ligand 1
PDx	Pharmacodynamic(s)
PFS	Progression-free survival
PFS2	Time from randomisation to second progression
PR	Partial Response
PRO	Patient Reported Outcome
QoL	Quality of Life
QLQ-LC13	Quality of Life Lung Cancer Module; 13 item self administered questionnaire from the EORTC for lung cancer
QTcF	QT interval (corrected for heart rate using Fredericia's correction)
RECIST 1.1	Response Evaluation Criteria In Solid Tumours, Version 1.1
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
TL	Target lesions
WHO	World Health Organization

AMENDMENT HISTORY

Date	Brief description of change
N/A	N/A

1. STUDY DETAILS

1.1 Study objectives

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1 (programmed death ligand 1)-positive and/or PD-L1–negative NSCLC.

1.1.1 Primary objective

Primary Objective:	**Outcome Measure:**
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with NSCLC	PFS using Investigator assessments according to RECIST 1.1a

[a] Sensitivity analyses of PFS will be performed based on BICR assessment according to RECIST 1.1 (subset of patients with BICR assessment only) and Investigator assessment according to RECIST 1.1 modified for confirmation of progression. If bias cannot be excluded based upon the sample BICR assessment, then an independent evaluation of all radiographic images will be required for the assessment of the primary PFS endpoint.

1.1.2 Secondary objectives

Secondary Objectives:	**Outcome Measures:**
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, ORR, DoR, APF12, PFS2, and OS	PFS in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1 ORR, DoR, and APF12 using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice[a] OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, ORR, PFS2, and OS	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS in patients with PD-L1–positive NSCLC using Investigator assessments according to RECIST 1.1 PFS2 using local standard clinical practice[a] OS
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS and ORR	PFS and ORR using Investigator assessments according to RECIST 1.1 PFS and ORR in patients with PD-L1–negative NSCLC using Investigator assessments according to RECIST 1.1

To assess disease-related symptoms and HRQoL in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC QLQ-C30 v3 and the LC13 module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.
To assess the PK of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of ADAs for MEDI4736 and tremelimumab

[a] PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death.

1.1.3 Safety objective

Safety Objective:	**Outcome Measures:**
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of advanced or metastatic NSCLC patients	Adverse events (AE), physical examinations, laboratory findings, and vital signs

1.1.4 Exploratory objectives

Exploratory Objectives:	**Outcome Measures:**
To explore irRECIST as an assessment methodology for clinical benefit of MEDI4736 + tremelimumab compared to SoC with assessment by BICR	PFS and ORR using BICR assessment according to irRECIST
To assess AEs by patient self-reporting of specific CTCAE symptoms	Collection of approximately 20 patient-reported outcomes version of CTCAE (PRO-CTCAE) symptoms via an electronic device solution
To assess patients' overall impression of the change in their health status since the start of study treatment	Patients' Global Impression of Change (PGIC) item will be collected directly from patients via an electronic device solution
To investigate the relationship between PK exposure and clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze PK exposure and the relationship with clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, as deemed appropriate

To describe and evaluate resource use associated with assigned treatments and underlying disease during assigned treatment	Health resource utilization measures including hospitalization, outpatient visits, or emergency department visits
To explore the impact of treatment and disease state on health state utility using the EQ-5D-5L during assigned treatment	EQ-5D-5L health state utility index will be used to derive health state utility based on patient-reported data
To investigate associations between pre-treatment peripheral myeloid-derived suppressor cells (MDSCs) measures and clinical activity	A graphical and/or a data modeling approach will be used to analyze the relationship between MDSC counts with clinical outcomes and/or with tumor lesion measurements
To investigate the relationship between biomarkers and clinical outcomes, efficacy, AEs, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze biomarkers (eg, IFNγ and/or PD-L1 status defined under alternative methods) and the relationship with clinical outcomes, efficacy, AEs, and/or safety parameters, as deemed appropriate

Note: Exploratory objective analyses may be reported separately from the main clinical study report.

1.2 Study design

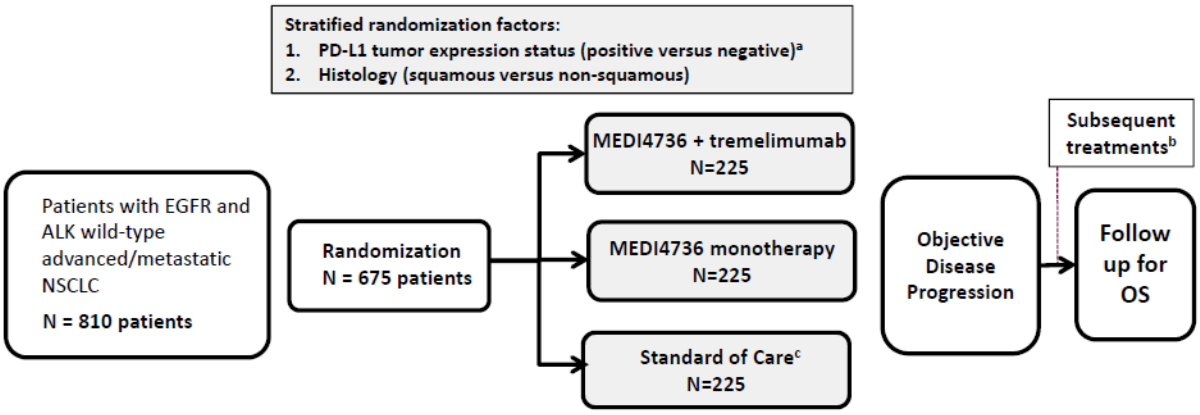
This is a randomized, open-label, multi-center, global Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with EGFR and ALK wild-type advanced or metastatic NSCLC. A schematic diagram of the overall study design is shown in Figure 1, and a detailed study flow chart is shown in Figure 2.

This study will enroll approximately 810 patients at sites in North America, Asia, Australia, and Europe to randomize approximately 675 patients (including at a minimum 438 patients with PD-L1–negative NSCLC) to treatment.

Patients will provide a tumor tissue sample at screening (newly acquired or archived sample <3 months old) to determine PD-L1 expression status (defined by an immunohistochemistry [IHC] assay developed by in which ≥25% PD-L1–membrane expression in tumoral tissue is considered positive and <25% is considered negative for PD-L1 expression; referred to hereafter as patients with PD-L1-positive or PD-L1-negative tumors, respectively).

Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (as described above) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy. Doses and treatment regimens are described in Section 7.2 of the CSP. Assessments will be conducted as indicated in Table 2, Table 3, and Table 4 in the CSP.

Figure 1 Overall study design



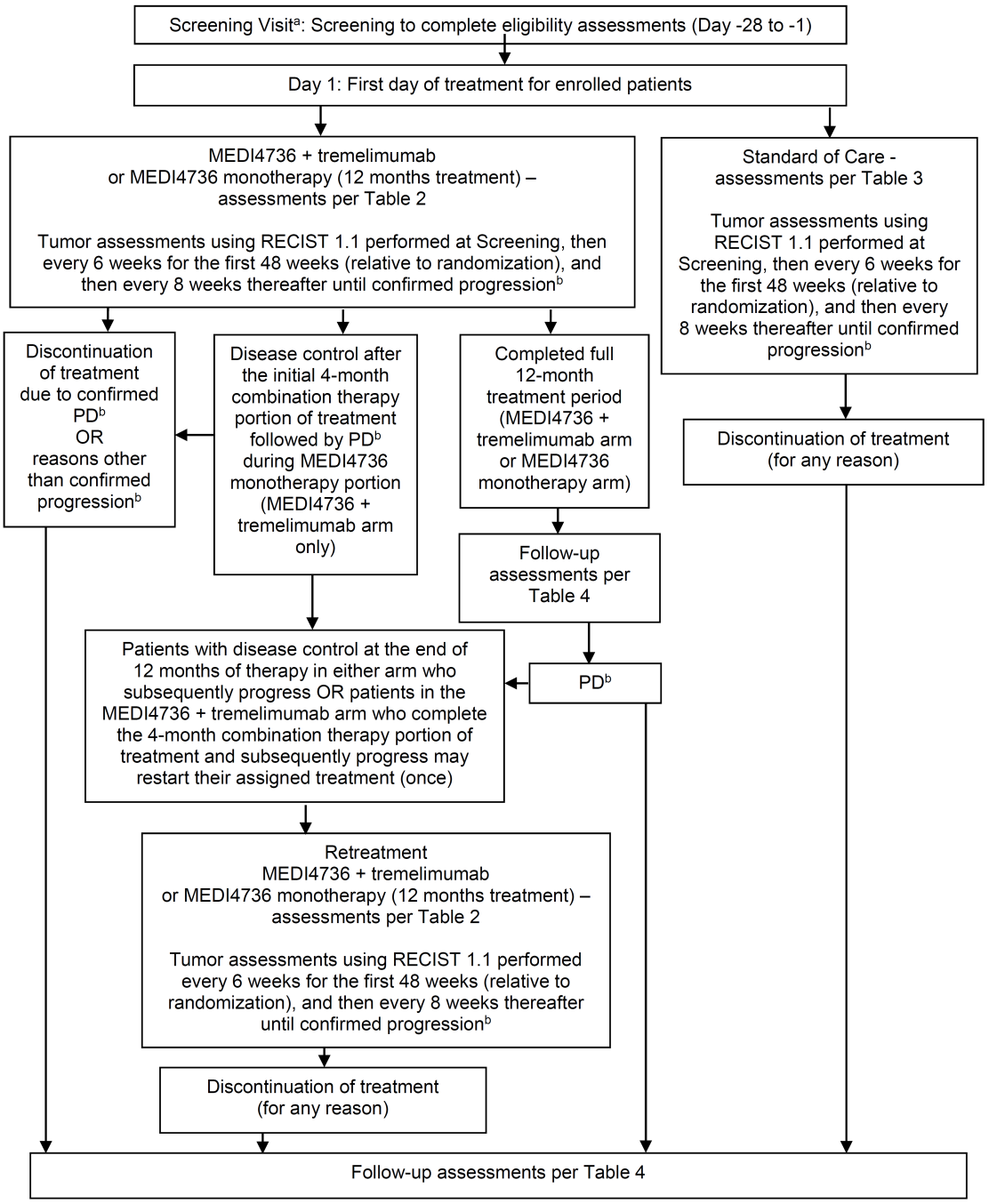
[a] Sites will be supplied with PD-L1 status upon request at disease progression.

[b] Offer of standard chemotherapy per Investigator discretion.

[c] Standard of Care is an Investigator choice from the following: paclitaxel + carboplatin, gemcitabine + cisplatin (or carboplatin) (squamous only), pemetrexed + cisplatin (or carboplatin) (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only following pemetrexed/platinum induction).

Figure 2 Study flow chart

Tables referred to in this figure are in reference to the CSP.



[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization.

[b] A confirmatory scan is always required following the initial demonstration of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). (See Section 5.1 of CSP for more information.)

Independent Data Monitoring Committee (IDMC)

An IDMC comprised of independent experts will be convened and will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review safety assessments and make recommendations to continue, amend, or stop the study based on safety findings. The committee will meet at least every 6 months thereafter.

Details on the IDMC are provided in Section 5.1 and full details of the IDMC procedures and processes can be found in the IDMC Charter.

1.3 Number of patients

The study will plan to enroll approximately 810 patients in order to randomize 675 eligible patients 1:1:1 to MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC. The 675 patients will comprise at a minimum 438 patients who have PD-L1-negative tumors.

The study is sized to characterize the PFS benefit of MEDI4736 in combination with tremelimumab versus SoC in patients with EGFR and ALK wild-type advanced or metastatic NSCLC (ie, regardless of PD-L1 tumor expression status) and patients with PD-L1–negative tumors. The sizing assumes a 3-month delay in separation of the PFS curves between each arm, hence the use of average hazard ratios (HRs).

The primary analysis of PFS will be performed when

- (approximately) 338 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups (75% maturity) AND

- (approximately) 223 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1 negative tumors (76% maturity)

MEDI4736 + tremelimumab versus SoC (primary endpoint)

If PFS at 12 months was 33.5% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average PFS HR is 0.59, the trial will have >90% power to demonstrate statistical significance at the 5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.81. With a 15-month recruitment period and a minimum follow up period of 6.5 months assumed, it is anticipated that this analysis will be performed 21.5 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PD-L1–negative population)

If the boundary is crossed for the primary PFS analysis, then PFS in the PD-L1–negative population of MEDI4736 + tremelimumab versus SoC will be tested at the 5% alpha level (in line with the hierarchical testing strategyexplained in Section 4.2.1). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and SoC treatment groups and a true average PFS HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 90% power to demonstrate statistical significance at the 5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.77.

MEDI4736 monotherapy versus SoC (ITT and PD-L1–positive populations)

If the boundary is crossed for the primary PFS analysis and for PFS of MEDI4736 + tremelimumab versus SoC in the PD-L1–negative population, then a statistical analysis of MEDI4736 monotherapy versus SoC will be performed in all randomized patients as well as in patients with PD-L1-positive tumors (in line with the hierarchical testing strategy):

- **ITT population:** With approximately 450 patients randomized across the MEDI4736 monotherapy and SoC treatment groups and a true average PFS HR of 0.59, an estimated 338 progression/death events (75% maturity) are expected to have occurred at 21.5 months from "first patient in," providing >90% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.78.

- **PD-L1-positive population:** With approximately 158 patients with PD-L1-positive tumors randomized across the MEDI4736 monotherapy and SoC treatment groups and an true average PFS HR of 0.51, an estimated 114 progression/death events (72% maturity) are expected to have occurred at 21.5 months from "first patient in," providing >90% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that could be statistically significant being an average HR of 0.65.

MEDI4736 + tremelimumab versus MEDI4736 monotherapy (PD-L1–negative population)

If the boundary is crossed for the primary PFS analysis and for PFS of MEDI4736 + tremelimumab versus SoC in the PD-L1–negative population, then a statistical analysis of MEDI4736 + tremelimumab versus MEDI4736 monotherapy will be performed in the PD-L1-negative population (in line with the hierarchical testing strategy). With approximately 292 patients with PD-L1-negative tumors randomized across the MEDI4736 + tremelimumab and MEDI4736 monotherapy treatment groups, if PFS at 12 months was 30% with MEDI4736 + tremelimumab and 13% with MEDI4736 monotherapy (~5.8-month median PFS) and assuming an true average PFS HR of 0.64, an estimated 223 progression/death events (76% maturity) are expected to have occurred at 21.5 months from "first patient in," providing approximately 86% power to demonstrate statistical significance at the 2.5% level (using a 2-sided test), with the smallest treatment difference that would be statistically significant being an average HR of 0.74.

MEDI4736 + tremelimumab versus SoC (overall survival)

The final analysis of OS will be performed when approximately 333 deaths have been observed. An interim analysis of OS will be performed at the time of the primary PFS analysis. The alpha will be split between the 2 OS analyses using a bespoke spending function where a fixed significance level will be assigned at the first analysis and the remaining significance level assigned to the final analysis, taking account of correlation between analyses (Stone 2010). If OS at 18 months was 49% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9-month median OS and assuming the true average OS HR is 0.70, an estimated 333 death events (74% maturity) are expected to have occurred at 36.5 months from "first patient in." With 333 deaths, the study will have approximately 84% power to demonstrate statistical significance using an overall alpha level of 2.5% (2-sided test). The smallest treatment difference that could be statistically significant will be an average HR of 0.78.

Table 1 provides a summary of the statistical assumptions.

Table 1 Summary of statistical assumptions

Endpoint	Analysis population	Events (number of patients)	Alpha (%)	Power (%)	Average hazard ratio to detect
PFS MEDI+tremelimumab versus SoC	FAS (ITT population)	338 (450)	5	>90	0.59
PFS MEDI+tremelimumab versus SoC	PD-L1-negative population	223 (292)	5	90	0.64

Endpoint	Analysis population	Events (number of patients)	Alpha (%)	Power (%)	Average hazard ratio to detect
PFS MEDI monotherapy versus SoC	FAS (ITT population)	338 (450)	2.5	>90	0.59
PFS MEDI monotherapy versus SoC	PD-L1-positive population	114 (158)	2.5	>90	0.51
PFS MEDI+tremelimumab versus MEDI monotherapy	PD-L1-negative population	223 (292)	2.5	86	0.64
OS MEDI+tremelimumab versus SoC	FAS (ITT population)	333 (450)	2.5	84	0.70

PFS Progression free survival; OS Overall Survival; PD-L1 Programmed death ligand 1.

2. ANALYSIS SETS

2.1 Definition of analysis sets

Full analysis set (Intention to treat (ITT))

The full analysis set (FAS) will include all randomized patients. The full analysis set will be used for all efficacy analyses (including PROs). Treatment groups will be compared on the basis of randomized study treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the analysis in the treatment group to which they were randomized.

PD-L1-negative analysis set

The PD-L1-negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 negative as defined by an IHC assay developed by (ie, <25% PD-L1– membrane expression in tumoral tissue).

PD-L1-positive analysis set

The PD-L1-positive analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive as defined by an IHC assay developed by (ie, ≥25% PD-L1–membrane expression in tumoral tissue).

Safety analysis set

The safety analysis set (SAS) will consist of all patients who received at least 1 dose of study treatment. Safety data will not be formally analyzed but summarized using the safety analysis set, according to the treatment received, that is, erroneously treated patients (e.g., those randomized to treatment A but actually given treatment B) will be summarized according to the treatment they actually received.

Pharmacokinetic analysis set

All patients who received at least 1 dose of investigational product (IP) per the protocol for whom any post-dose data are available and who do not violate or deviate from the protocol in ways that would significantly affect the pharmacokinetic (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. Definitions of the analysis sets for each outcome variable are provided in Table 2.

Table 2 Summary of outcome variables and analysis populations

Outcome variable	Population
Efficacy data	
PFS	Full analysis set (ITT population)
ORR, DoR, APF12, PFS2, OS, PROs, and symptom endpoints	Full analysis set (ITT population)
PFS, ORR	PD-L1-negative analysis set
PFS	PD-L1-positive analysis set
Demography	Full analysis set (ITT population)
PK data	PK analysis Set
Safety Data	
Exposure	Safety analysis Set
AEs	Safety analysis Set
Laboratory measurements	Safety analysis Set
Vital signs	Safety analysis Set
ECGs	Safety analysis Set

2.2 Violations and deviations

The important protocol deviations will be listed and summarised by randomised treatment group. Deviation 1, below, will lead to exclusion from the Safety analysis set. Deviation 4, below, will lead to exclusion from the FAS analysis set. None of the other deviations will lead to patients being excluded from the analysis sets described in Section 2.1 (with the exception of the PK analysis set, if the deviation is considered to impact upon PK). A per-protocol analysis excluding patients with significant protocol deviations is not planned; however, a 'deviation bias' sensitivity analysis will be performed excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients:

- did not have the intended disease or indication or
- did not receive any randomised therapy.

The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock and will be documented prior to the primary analysis being conducted.

Eligibility criteria deviations are deviations from the protocol inclusion and exclusion criteria. Post-entry deviations are deviations from the protocol that occurred after the patient was assigned to the study.

The following general categories will be considered important deviations and be listed and discussed in the CSR as appropriate for the study. If a 'deviation bias' sensitivity analysis is conducted then patients with these deviations will be excluded from the sensitivity analysis:

- Deviation 1: Patients randomised but who did not receive study treatment.
- Deviation 2: Patients who deviate from key entry criteria as per the CSP. These are inclusion criteria 3,5 and exclusion criteria 3,4, 8,18.
- Deviation 3: Baseline RECIST scan > 42 days before date of randomisation.
- Deviation 4: No baseline RECIST 1.1 assessment on or before date of first dose.
- Deviation 5: Received prohibited concomitant medications (including other anti-cancer agents). Please refer to the Clinical Study Protocol (CSP) section 7.7 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.
- Deviation 6: Patients randomised who received treatment other than that to which they were randomised to.

The categorisation of these as important deviations is not automatic and will depend on duration and the perceived effect on efficacy.

In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification of deviations will be made at the blinded data review meeting (BDRM) prior to database lock or data freeze. Decisions made at the BDRM will be documented and approved by AstraZeneca prior to analysis.

Misrandomisations in terms of errors in treatment dispensing, in addition to incorrect stratifications, will also be summarised and listed separately to the important protocol deviations. A misrandomisation is when a patient is not randomised or treated according to the randomisation schedule. It is envisaged that there will be 2 sub categories of this:

- Patients who receive no treatment whatsoever for a period of time due to errors in dispensing of medication. Note, this is not due to tolerability issues where patients may stop taking drug.

- The patient receives a treatment pack with a different code to their randomisation code. However, the actual treatment may still match the randomised treatment. For example, a patient is given randomisation code 0001, which according to the randomisation schedule is MEDI4736. However, at the randomisation visit they are given treatment pack 0003, but this still contains MEDI4736.

The summary will include all patients with a dispensing error but will also include information on how many of those patients received at least one dose of the wrong treatment at any time. Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1. During the study, decisions on how to handle misrandomisations will be made on an individual basis with written instruction from the study team leader and/or statistician.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST Visit Responses

For all patients, the RECIST version 1.1 (see further details in Appendix F in the CSP) tumour response data will be used to determine each patient's visit response. It will also be used to determine if and when a patient has progressed and also their best objective response.

The baseline assessment should be performed no more than 28 days before the start of IP treatment and ideally as close as possible to the start of the assigned IP. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; Table 2 in the CSP for MEDI4736 + tremelimumab or MEDI4736 monotherapy, Table 3 in the CSP for SoC, and Table 4 in the CSP for patients

who have completed/discontinued randomized treatment) then every 8 weeks thereafter, until confirmed objective disease progression per RECIST 1.1 (irrespective of the reason for stopping treatment or subsequent therapy). 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.

For patients who discontinue IP (including SoC) due to toxicity in the absence of confirmed objective progression, objective tumor assessments per the scheduled assessments should be continued every 6 weeks for 48 weeks (relative to randomization) and then every 8 weeks until confirmed objective disease progression.

A confirmatory scan is required for all patients following the initial demonstration of PD. The confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD and preferably at the next scheduled visit in the absence of clinically significant deterioration. Treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC may continue between the initial assessment of progression and confirmation of progression. Progression would be considered confirmed per RECIST 1.1 criteria available in Appendix F in the CSP using Investigator assessments.

If a patient discontinues treatment (and/or receives a subsequent anticancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

Categorization of objective tumor response assessment will be based on the RECIST 1.1 criteria of response: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). Target lesion progression will be calculated in comparison to when the tumor burden was at a minimum (ie, smallest sum of diameters previously recorded on study). In the absence of progression, tumor response (CR or PR) and SD will be calculated in comparison to the baseline tumor measurements obtained before starting treatment.

Following confirmed progression, patients should continue to be followed up for survival every 1 to 2 months as outlined in the study plan (Table 4 in the CSP). Exceptions are patients with confirmed PD who continue to receive IP at the discretion of the Investigator (after consultation with AstraZeneca); these patients can receive treatment for a maximum of 12 months and will have scans for RECIST 1.1 assessments every 6 weeks (relative to the date of randomization per Table 2 and Table 3 in the CSP) for the first 48 weeks of treatment and then every 8 weeks until disease progression. Subsequent anticancer therapy information will be collected at the timepoints indicated in Table 4 in the CSP.

Patients in the MEDI4736 + tremelimumab or MEDI4736 monotherapy groups who will receive retreatment must have a baseline tumor assessment within 28 days of restarting treatment and additional scans every 6 weeks for the first 48 weeks relative to the date of randomization, and then every 8 weeks thereafter until disease progression. All assessments in Table 2 will be followed for patients who receive retreatment.

3.1.1 Investigator RECIST 1.1-based assessments: Target lesions (TLs)

At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD, or PD depending on the status of their disease compared with baseline and previous assessments. Baseline will be assessed within the 28 days prior to randomization. If a patient has had a tumor assessment that cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE; unless there is evidence of progression in which case the response will be assigned as PD).

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (except lymph nodes which must have short axis ≥ 15 mm) with CT or MRI and which is suitable for accurate repeated measurements.

A patient can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (representative of all lesions involved suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to the date of randomisation 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 non-target lesions (NTL) at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits.

Measurable disease (ie at least one TL) is one of the entry criteria for the study. However, if a patient with non-measurable disease is enrolled in the study, the evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new lesions. If a patient does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 3 TL visit responses

Visit Reponses	Description
Complete Response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.

Visit Reponses	Description
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not Evaluable (NE)	Only relevant in certain situations (i.e. if any of the target lesions were not assessed or not evaluable or had a lesion intervention at this visit; and scaling up could not be performed for lesions with interventions). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a target lesion response
Not applicable (NA)	No TLs are recorded at baseline

Rounding of TL data

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

Missing TL data

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

- A new lesion is recorded.
- A NTL visit response of PD is recorded.
- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared.

Note: the nadir can only be taken from assessments where all the TLs had a lesion diameter recorded.

Lymph nodes

For lymph nodes, if the size reduces to < 10mm 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 < 10mm and all other TLs are 0mm then although the sum may be >0mm 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 a value of 5mm will be entered into the database and used in TL calculations, unless the radiologist has indicated and 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 blinded to treatment assignment.

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolisation), should be handled in the following way 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 tumours:

- 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 subject would be assigned a TL response of PD.

- Step 3: If after both steps PD has not been assigned, then, if appropriate, a scaled sum of diameters will be calculated (as long as ≤ 1/3 of the TLs have missing measurements), treating the lesion with intervention as missing, 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 also have a value of 0 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 where appropriate (as per step 2 above).

Scaling (applicable only for irradiated lesions/lesion intervention)

If > 1/3 of TL measurements are treated as missing (because of intervention) then TL response will be NE, unless the sum of diameters of non-missing TL would result in PD (ie 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 treated as 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.

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

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

$$\frac{26}{26.8} \times 29.3 = 28.4cm$$

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

Lesions that split in two

If a TL splits in 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.

3.1.2 Investigator RECIST 1.1-based assessments: Non-target lesions (NTLs) and new lesions

At each visit an overall assessment of the NTL response should be recorded by the investigator. 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 4 NTL Visit Responses

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

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

Details of any new lesions 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 tumour.

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 tumour assessments where possible until objective disease progression is observed.

3.1.3 Investigator RECIST 1.1-based assessments: Overall visit response

Table 5 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 5 Overall visit responses

Target lesions	Non-target lesions	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

Target lesions	Non-target lesions	New lesions	Overall visit response
NA	NE	No (or NE)	NE

3.1.4 Blinded Independent Central Review (BICR) of RECIST 1.1-based assessments

A BICR of radiological scans will be performed to confirm the robustness of the PFS endpoint. This may initially either be performed on all patients or just on a random sample of patients (study team to confirm approach) . If a sample BICR is chosen, a separate document will be produced which will outline the justification for the percentage of patients needed in the sample BICR, the method used to identify the subset of patients, the method for comparing the PFS results obtained by local review with the PFS results of the sample BICR, and the criteria for determining whether BICR needs tobe performed on all patients. If bias cannot be excluded based upon an initial sample BICR assessment, then an independent evaluation of all radiographic images will be required for assessment of the primary PFS endpoint.

All images will be collected centrally. The imaging scans will be reviewed by 2 independent radiologists using RECIST 1.1 and will be adjudicated, if required. For each patient, the BICR will define the overall visit response data (CR, PR, SD, PD, or NE) and the relevant scan dates for each time point (ie, for visits where response or progression is/is not identified). 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). Endpoints (of PFS) will be derived from the overall visit response date and the scan dates. If there are significant discrepancies between the site and BICR evaluations, a BICR of all patients will be performed.

The definitions of irCR, irPR, irSD, irPD, and irNED (ie responses according to irRC), as outlined by Wolchok et al 2009, will be outlined in the BICR charter, but a brief description of the methodology is given here. In this project irRC using a RECIST base will be implemented where the target lesions will be measured unidimensionally.

In irRC the presence of new lesions will not automatically trigger a declaration of Progressive Disease, but instead the new lesions will be measured and these measurements will be added

to the sum of diameters of the target lesions. Based on the sum of these measurements and % calculations thereof, the target lesion response assessment will be derived. The overall response assessment (irCR, irPR, irSD, irPD, irNE or irNED) will be obtained at the BICR and confirmation of irPD is required.

PFS and ORR by irRECIST criteria using BICR assessments will also be performed for exploratory purposes. This will be on either the patient subset or the full set of assessments, as deemed neccessary. If a full BICR is chosen or deemed necessary then these exploratory analyses will have a greater importance when interpreting the results.

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

3.2 Outcome Variables

The analysis of the primary endpoint, PFS, and the analyses of the secondary endpoints, ORR, DoR, and APF12, will be based on Investigator tumor assessments according to RECIST 1.1. In addition, time to secondary progression (PFS2) will be defined by local clinical practice and OS will also be evaluated.

All RECIST assessments, 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.

3.2.1 Primary endpoint - progression-free survival

Progression free survival (PFS) (per RECIST 1.1 using Investigator assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression. 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. If the patient has no evaluable visits or does not have baseline data, they will be censored at 0 days unless they die within 2 visits (2x6 weeks for tumor assessments + 2x7 days for visit window) of baseline.

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 investigational assessments, the date of progression will be determined based on the earliest of the RECIST assessment/scan dates of the component that indicates 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 target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the target lesions, and similarly for non-target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the non-target lesions.

PFS will also be determined using BICR assessments according to RECIST 1.1.

Exploratory analyses

PFS based on RECIST 1.1 modified for confirmation of progression will be performed for exploratory purposes using the algorithm described above for the RECIST 1.1 Investigator assessments, but following a modification whereby any objective disease progression must be

confirmed by the next scheduled scan. The confirmatory scan must be no sooner than 4 weeks after the initial suspected progression and preferably at the next scheduled visit in the absence of clinically significant deterioration. If disease progression is confirmed (or disease progression occurs and no further scans are recorded) then the date of progression will be when it was originally observed. Patients with a single disease progression and no further tumor assessment scans will be treated as PD in the analysis. In the absence of significant clinical deterioration, the investigational site is advised to continue the patient on their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy until progression has been confirmed. If progression is not confirmed, the patient should continue their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy treatment and on-treatment assessments. Treatment through PD in the SoC group is at the Investigator's discretion; however, a confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

PFS by irRECIST criteria using BICR assessments will also be reported.

3.2.2 Secondary endpoints

3.2.2.1 Objective response rate

ORR (per RECIST 1.1 using Investigator assessments) is defined as the number (%) of patients with at least 1 visit response of CR or PR. If any patients do not have measurable disease at baseline then the analysis of ORR will exclude these patients, so that the denominator is a subset of the Intent-to-Treat (ITT) population who have measurable disease at baseline. 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.

An exploratory analysis of ORR by irRECIST criteria using BICR assessments will also be reported.

3.2.2.2 Duration of response

DoR (per RECIST 1.1 using Investigator assessments) will be defined as the time from the date of first documented response 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 denominator for DoR will be defined as described for ORR (see Section 3.2.2.1).

The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of CR or PR.

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.

3.2.2.3 Time from randomization to second progression

Time from randomisation to second progression (PFS2) is defined as the time from the date of randomisation to the earliest of the progression events (subsequent to that used for the primary variable PFS) or death (ie date of PFS2 event or censoring – date of randomisation + 1). The date of the first progression will be programmatically determined from investigator assessed data (See Section 3.2.1 for details.) The date of second progression will be recorded by the Investigator in the eCRF and defined according to local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression or death. RECIST assessments will not be collected for assessment of PFS2. The date of the PFS2 assessment and investigator opinion of progression status (progressed or nonprogressed) at each assessment will be recorded in the electronic case report form (eCRF).
Second progression status will be reviewed (every 6 weeks for the first 48 weeks relative to the date of randomization and then every 8 weeks thereafter) following the progression event used for the primary variable PFS (the first progression) and status recorded. Patients alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression, that is, censored at the latest of the PFS or PFS2 assessment date if the patient has not had a second progression or death.

3.2.2.4 Proportion of patients alive and progression free at 12 months

The proportion of patients alive and progression free at 12 months (APF12) will be defined as the Kaplan-Meier estimate of PFS (per RECIST 1.1 as assessed using Investigator assessments) at 12 months.

3.2.2.5 Overall survival

OS is defined as the time from the date of randomization until death due to any cause (ie, date of death or censoring – date of randomisation + 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 (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made in the week following the date of Data Cut Off (DCO) for the analysis, and if patients are confirmed to be alive or if the death date is post the DCO date these patients 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 CRF fields to determine the

last recorded date on which the patient was known to be alive for those patients still on treatment (since the SURVIVE module is only completed for patients off treatment).

3.2.2.6 Best objective response

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in Appendix F in the CSP. It is the best response a patient has had following randomization during their time in the study up until RECIST progression or the last evaluable assessment in the absence of RECIST progression.

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

BoR will be determined programmatically based on RECIST 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 assessments prior to death.

For patients who die with no evaluable RECIST assessments, if the death occurs ≤90 days (ie, 2*(6 weeks ±3 days)) after randomization, then BoR will be assigned to the progression (PD) category. For patients who die with no evaluable RECIST assessments, if the death occurs >90 days (ie, 2*(6 weeks ±3 days)) after the date of randomization then BoR will be assigned to the NE category.

Progression events that have been censored due to them being >90 days after the last evaluable assessment will not contribute to the BoR derivation.

3.2.2.7 Change in tumour size

For supportive purposes percentage change from baseline in tumour size will be derived at each scheduled tumour assessment visit (ie, week 6, week 12 etc hereafter referred to as week X for convenience). Best percentage change from baseline in tumour size will also be derived as the biggest decrease or the smallest increase in tumour size from baseline.

This is based on RECIST target lesion measurements taken at baseline and at the timepoint of interest. Tumour size is defined as the sum of the longest diameters of the target lesions for the BICR data based upon RECIST assessments. Target lesions are measurable tumour lesions. Baseline for RECIST is defined to be the last evaluable assessment prior to starting treatment. The change in target lesion tumour size at week X will be obtained for each patient by taking the difference between the sum of the target lesions at week X and the sum of the target lesions at baseline. To obtain the percentage change in target lesion tumour size at week X the change in target lesion tumour size is divided by the sum of the target lesions at baseline and multiplied by 100 (i.e. (week X - baseline) / baseline * 100). More details on target lesions and measurements can be found in Section 3.1.

Apply a window around the week X visit: Whenever tumour size data for the week X visit (Note: or visit at which progression was documented if before week X) is available then this should be used in the analysis. A windowing rule will be applied and will follow the protocol

allowed visit window; therefore any RECIST scan performed within ± 7 days of the protocol scheduled visit will be used for that visit.

The above derivations will be programmed for the investigator assessments based upon RECIST analysis.

3.3 Patient-reported outcome (PRO) variables

Patient reported outcome (PRO) questionnaires will be assessed using the EORTC QLQ-C30 with the QLQ-LC13 module (HRQoL and lung cancer specific symptoms), PRO-CTCAE, PGIC, and EQ-5D-5L. All items/questionnaires will be scored according to published scoring guidelines or the developer's guidelines, if published guidelines are not available. All PRO analyses will be based on the Full Analysis Set (FAS; ITT population), unless stated.

3.3.1 EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea/vomiting), 5 individual items (dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a global measure of health status. The EORTC QLQ-C30 will be scored according to the EORTC QLQ-C30 scoring manual (Fayers et al 2001). An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, each of the functional scales, and the global health status scale in the EORTC QLQ-C30 according to the EORTC QLQ-C30 Scoring Manual. Higher scores on the global health status and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity.

Baseline will be defined as the last non-missing assessment prior to first dose for symptoms and summaries.

The change from baseline in HRQoL will be assessed using the EORTC-QLQ-C30 global QoL scale which includes 2 items from the QLQ-C30: "How would you rate your overall health during the past week? (Item 29) and "How would you rate your overall QoL during the past week? (Item 30).

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥10 for scales/items from the EORTC QLQ-C30 (Osoba et al 1998). For example, a clinically meaningful improvement in physical function (as assessed by EORTC QLQ-C30) is defined as an increase in the score from baseline of ≥10, whereas a clinically meaningful deterioration is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, the change in symptoms/functioning from baseline will be categorized as improvement, no change or deterioration as shown in Table 6.

Table 6 Mean change and visit response in health-related quality of life

Score	Change from baseline	Visit response
EORTC QLQ-C30 Global quality of life score	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change
EORTC QLQ-C30 symptom scales/items	≥+10	Deterioration
	≤-10	Improvement
	Otherwise	No change
EORTC QLQ-C30 functional scales	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change

For each subscale, if <50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded. If there is evidence that the missing data are systematic, missing values will be handled to ensure that any possible bias is minimized.

For the visit level summaries of Improvement/Deterioration/No change then all patients with a baseline and post-baseline score will be included thus the denominator may differ from the time to deterioration and improvement rate endpoints derived below.

3.3.1.1 Time to HRQoL/function deterioration

For the following HRQoL items of the EORTC QLQ-C30, time to deterioration will be analyzed:

The Global Health Status/ QoL scale consisting of items 29 and 30 of the EORTC QLQ C30.

- Item 29: How would you rate your overall health during the past week?
- Item 30: How would you rate your overall quality of life during the past week?

Patients are asked to rate their overall health and overall quality of life on a scale from 1 (very poor) to 7 (excellent).

Time to deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful deterioration (a decrease in the function scales or the global health status/HRQoL from baseline of ≥10) that is confirmed at a subsequent visit or death (by

any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to HRQoL/function deterioration. (ie date of HRQoL/function deterioration event or censoring-date of randomisation + 1). Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the HRQoL/function change could be evaluated.

Patients whose HRQoL (as measured by EORTC QLQ-C30) has not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the HRQoL/function could be evaluated. Also, if HRQoL deteriorates after 2 or more missed PRO assessment visits (using the same definitions for two missed visits as used in the 'Time to Symptom deterioration' derivation in section 3.3.2.1) or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated. If the patient has no evaluable visits or does not have baseline data they will be censored at 1 day unless they die within 2 visits of baseline (16 weeks plus 1 week allowing for a late assessment within the visit window).

The population for analysis of time to HRQoL/function deterioration will include a subset of the ITT population who have baseline scores ≥ 10. In the primary analysis, RECIST 1.1 progression will not be considered as HRQoL/function deterioration and data will not be affected by RECIST progression. However a sensitivity analyses will be performed for time to HRQoL deterioration for global health status only where RECIST 1.1 progression is considered as an event of HRQoL deterioration. If a patient has both RECIST 1.1 progression and HRQoL deterioration, the earliest date of the two will be used.

3.3.1.2 Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥10 for EORTC QLQ-C30 symptom scales/items) in that symptom from baseline. The denominator will consist of a subset of the ITT population who have a baseline symptom score >10.

3.3.1.3 HRQoL/function improvement rate

The HRQoL/function improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (an increase from baseline score ≥10 for EORTC QLQ-C30 functional scales and global health status/HRQoL) in that scale from baseline. The denominator will consist of a subset of the ITT population who have a baseline HRQoL/function score ≤90.

3.3.2 Lung cancer module (EORTC QLQ-LC13)

The QLQ-LC13 is a lung cancer specific module from the EORTC for lung cancer comprising 13 questions to assess lung cancer symptoms (cough, hemoptysis, dyspnea, and site-specific pain), treatment-related side-effects (sore mouth, dysphagia, peripheral neuropathy, and alopecia), and pain medication.

The LC-13 incorporates symptom scales including:

- Dyspnea (multi-item scale based on 3 questions: were you short of breath when you rested; walked; climbed stairs)
- Cough: 1 item (how much did you cough?)
- Haemoptysis: 1 item (did you cough up blood?)
- Pain: 3 individual items (Have you had pain in your chest; your arm or shoulder; other parts of your body?)

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC-13 is identical in principle to that for the symptom scales/single items of the EORTC-QLQ-C30.

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥10 for scales/items from the QLQ-LC13 (Osoba et al 1998). For example, a clinically meaningful deterioration or worsening in chest pain (as assessed by the QLQ-LC13) is defined as an increase in the score from baseline of ≥10, whereas a clinically meaningful improvement is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, the change in symptoms from baseline will be categorized as an improvement, no change or deterioration as shown in Table 7.

Table 7 Visit response for health-related quality of life (HRQoL) and disease-related symptoms

Score	Change from baseline	Visit response
QLQ-LC13 symptom scales/items	≥10	Deterioration
	≤10	Improvement
	Otherwise	No change

QLQ-LC13 Lung Cancer Module.

For the visit level summaries of Improvement/Deterioration/No change then all patients with a baseline and post-baseline score will be included thus the denominator may differ from the time to deterioration and improvement rate endpoints derived below.

3.3.2.1 Time to symptom deterioration

For each of the following key symptom scales/items in the QLQ-LC13, time to deterioration will be analyzed:

- Dyspnea (multi-item scale based on 3 questions: were you short of breath when you rested; walked; climbed stairs)
- Cough: 1 item (how much did you cough?)
- Hemoptysis: 1 item (did you cough up blood?)
- Pain (3 individual items): a) Have you had pain in your chest; b) your arm or shoulder; c) other parts of your body?)

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales/single items of the EORTC QLQ-C30.

Time to symptom deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful symptom deterioration (an increase in the score from baseline of ≥10) that is confirmed at a subsequent visit or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to symptom deterioration (ie date of symptom deterioration event or censoring – date of randomisation + 1). Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 visits of the last PRO assessment where the symptom change could be evaluated.

Patients whose symptoms (as measured by QLQ-LC13) have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the symptom could be evaluated. Also, if symptoms deteriorate after 2 or more missed PRO assessment visits or the patient dies after 2 or more missed PRO assessment visits, the patient will be censored at the time of the last PRO assessment where the symptom could be evaluated. If the patient has no evaluable visits or does not have baseline data they will be censored at 1 day unless they die within 2 visits of baseline (126 days (i.e. 16 weeks x7 days) plus 2x7 days allowing for a late assessment within the visit window).

The population for analysis of time to symptom deterioration will include a subset of the ITT population who have baseline scores ≤ 90.

In the primary analysis, RECIST 1.1 progression will not be considered as symptom deterioration and data will not be affected by RECIST progression. However a sensitivity analyses will be performed for QLQ-LC13 time to symptom deterioration for each of chest pain, arm/shoulder pain, and other pain where RECIST 1.1 progression is considered as an event of symptom deterioration.

If a patient has both RECIST 1.1 progression and symptom deterioration, the earliest date of the two will be used.

3.3.2.2 Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥10 for QLQ-LC13 symptom scales/items) in that symptom from baseline. The denominator will consist of a subset of the ITT population who have a baseline symptom score >10.

3.3.3 Healthy state utility (EQ-5D-5L)

The EQ-5D is a standardised measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal. Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care.

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and unable to/extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied (Oemar and Janssen 2013). In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

The evaluable population will comprise a subset of the ITT analysis set who have a baseline EQ-5D-5L assessment.

3.3.4 PRO Compliance Rates

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30, LC13 and EQ-5D-5L respectively. These will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least one individual item completed.

- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time e.g. a questionnaire from a patient who has not withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. For patients that have progressed, the latest of progression and safety follow-up will be used to assess whether the patient is still under HRQoL follow-up at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.

- Evaluable questionnaire = a questionnaire with a completion date and at least one subscale that is non-missing.

- Overall PRO compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.

- Overall patient compliance rate is defined for each randomised treatment group as: Total number of patients with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above), divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

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

3.4 Safety

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

Data from the initial treatment period (ie, the initial 12 months of treatment) on the immunotherapy agents (MEDI4736 or MEDI4736+tremelimumab) will be compared against SoC in the main presentations of safety data and safety data from the retreatment period may also be summarised separately (see Section 4.1). ‘On treatment’ will be defined as assessments between date of start dose and 90 days following last dose of the immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) on each period of treatment and between date of start dose and 90 days following last dose of the Standard of Care agents. Note that for one version of the safety outputs the period of time after the

administration of subsequent therapy will not be considered 'on treatment' (see further Section 4.2.12).

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

3.4.1 Adverse events (AEs)

AEs and SAEs will be collected throughout the study, from date of first dose and 90 days after the last dose of immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) and between date of first dose and 90 days following last dose of the Standard of Care agents. The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03). A treatment emergent adverse event (TEAE) is an AE with an onset date or a pre-existing AE worsening following the first dose of study treatment through to 90 days after the last dose of immunotherapy agents (ie, the last dose of MEDI4736, tremelimumab or MEDI4736+tremelimumab) or 90 days after the last dose of the Standard of Care agents. For the MEDI4736+tremelimumab arm, in the unlikely event of the components being administered separately then date of first dose/last dose will be considered as the earliest/latest dosing date of either component.

Other significant adverse events

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

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

AEs of special interest

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

AESIs for MEDI4736 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 AESIs are being closely monitored in clinical studies with MEDI4736 monotherapy

and combination therapy. An immune-related adverse event (irAE) is defined as an adverse event that is associated with drug exposure and is consistent with an immune-mediated mechanism of action and where there is no clear alternate aetiology. Serologic, immunologic, and histologic (biopsy) data, as appropriate, should be used to support an irAE diagnosis. Appropriate efforts should be made to rule out neoplastic, infectious, metabolic, toxin, or other etiologic causes of the irAE.

Some clinical concepts (including some selected individual preferred terms and higher level terms) have been considered "AEs of special interest" (AESI) to the MEDI4736 program. These AESIs have been identified as Pneumonitis, Colitis, Hepatitis, Hypothyroidism, Hyperthyroidism, Hypnophysitis, Adrenal Insufficiency, Dermatitis, Nephritis /Acute Renal Failure, Pancreatitis, Neuropathy, Infusion-related Reactions and Hypersitivity / Anaphylactic Reactions. Other categories may be added or existing terms may be merged as necessary. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which preferred terms contribute to each AESI. A further review will take place prior to Database lock (DBL) to ensure any further terms not already included are captured within the categories.

3.4.2 Treatment exposure

Exposure will be defined separately for the initial treatment period and for the re-treatment period for the immunotherapy agents as follows:

Total (or intended) exposure of study treatment

- Total (or intended) exposure = last dose date where dose > 0 mg – first dose date + 1

Actual exposure of study treatment

- Actual exposure = intended exposure – total duration of dose delays, where intended exposure will be calculated as above.

Dose reductions are not permitted per Section 6.7 of the CSP for the immunotherapy agents (MEDI4736 or MEDI4736+tremelimumab). The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

Exposure will also be measured by the number of cycles received. For all five choices of SoC regimen, a cycle corresponds to a period of 21 days, but for each immunotherapy agent a cycle corresponds to one dose of treatment. If a cycle is prolonged due to toxicity, this should still be counted as one cycle. A cycle will be counted if treatment is started even if the full dose is not delivered.

Patients who permanently discontinue during a dose delay

If a decision is made to permanently discontinue study treatment in-between cycles or during a dose delay then the date of last administration of study medication recorded will be used in the programming.

3.4.3 Dose intensity

Dose intensity will be derived separately for the initial treatment period and the re-treatment period for the immunotherapy agents. It will also be derived for the SOC agents. Relative dose intensity (RDI) is the percentage of the actual dose intensity delivered relative to the intended dose intensity through to treatment discontinuation. Percentage intended dose (PID) is the percentage of the actual dose delivered relative to the intended dose through to progression or treatment completion. Both will be derived using the date of objective disease progression as defined by RECIST using the investigator site assessments. If the investigator considered that it was in the patient's best interest to continue study treatment past this time, this was not included in the derivation of RDI and PID.

Relative dose intensity (RDI) and percent intended dose (PID) will be defined as follows for MEDI4736, tremelimumab and all Standard of Care therapy:

- RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the earlier of progression (or a censoring event) or the actual last day of dosing and D is the intended cumulative dose up to the earlier of progression (or a censoring event) or the actual last day of dosing.

- PID = 100% * d/D, where d is the actual cumulative dose delivered up to progression (or a censoring event) and D is the intended cumulative dose up to progression (or a censoring event). D is the total dose that would be delivered, if there were no modification to dose or schedule.

The standard dose intensity calculations above are focused upon standard agents that treat to progression and as the primary analysis is based upon RECIST 1.1 it is still appropriate to calculate these. However, due to the fact that there is the intention to treat through progression per the protocol there will be additional measures of dose intensity to investigate the entire treatment period including the time period after progression:

- Relative dose intensity (treatment through progression) (RDI2) = 100% * d/D, where d is the actual cumulative dose delivered up to the later of progression (or a censoring event) or the actual last day of dosing and D is the intended cumulative dose up to the later of progression (or a censoring event) or the actual last day of dosing.

- Percentage intended dose (treatment through progression) (PID2) = 100% * d/D, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the actual last day of dosing.

As the treatment period is for a finite period of time for the immunotherapy agents, additional considerations need to be taken into account for the above calculations. For the each of the components of the combination arm and also the MEDI4736 monotherapy arm, where the last dose is on the week 48 visit and if there are scans post week 48, the censoring of progression should occur at week 48 for the purposes of RDI, PID, RDI2 and PID2.

Example of dose intensity for MEDI4736 monotherapy

			Study Day									
RDI	PID	Patient	1	29	57	85	113	141	169	197	225	
100%		1	X	X	X	X	X	X	X	X	X	PD
100%	89%	2	X	X	X	X	X	X	X	X[D]		PD
56%		3	X		X		X	O	X	X		PD
67%		4	X	X	O	X	X	X	O	X	O	PD

X: Dose of 10mg/kg taken; O: Dose missed; [D]: Dose discontinued; PD: Progressive Disease

Patients 1-4 progressed on Day 227, so the intended dose through to progression was 9 * 20mg/kg of MEDI4736 = 180mg/kg.

Patient 1 received a total of 180 mg/kg of MEDI4736, whereas other patients received less treatment due to:

- Early stopping prior to PD (Patient 2)
- Dosing delays (Patient 3)
- Missed doses (Patient 4)

The examples of Patients 2 and 4 illustrate that for RDI, the end of actual dosing period is calculated based on the smallest recovery period after the last non-zero dose.

Patient 1: RDI = PID = (9 * 20 mg/kg)/ 180 mg/kg = 100%

Patient 2: RDI = (8 * 20 mg/kg)/ 160 mg/kg = 100%

PID = (8 * 20 mg/kg)/ 180 mg/kg = 89%

Patient 3: RDI = PID = (5 * 20 mg/kg)/ 180 mg/kg = 56%

Patient 4: RDI = PID = (6 * 20 mg/kg)/ 180 mg/m^2 = 67%

3.4.4 Laboratory data

Laboratory data will be collected throughout the study, from screening to the follow-up visits as described in the CSP. Blood and urine samples for determination of haematology, clinical chemistry, and urinalysis will be collected as described in Section 5.2.1 of the CSP. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 3.4.7 below will be used.

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

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

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

Absolute values will be compared to the project reference range 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-dose value at any time.

Local reference ranges will be used for the primary interpretation of laboratory data at the local laboratory. Project reference ranges will be used throughout for reporting purposes. The denominator used in laboratory summaries of CTC 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 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 CTCAE criterion does not consider changes from baseline, to be evaluable the patient need only have 1 post dose-value recorded.

3.4.5 Time to first subsequent therapy from discontinuation of study treatment

Time to subsequent therapy from date of last dose is defined as the time from the date of discontinuation of study treatment to the start date of the first subsequent therapy after

discontinuation of treatment. Any patient not known to have had a first subsequent therapy will not have this calculation performed.

3.4.6 ECGs

ECG data obtained up until the 30 days from date of last dose of study treatment will be used for reporting. For derivation of post baseline visit values considering visit window and to handle multiple records present in any visit window, derivation rules as described in Section 3.4.7 below will be used.

At each time point the Investigator's assessment of the ECG will be collected locally. Heart rate, duration of QRS complex, RR, PR and QT intervals will be collected centrally via a digital read. This digital copy of all ECGs will be held centrally by a central ECG provider, and the data from this review will be stored for analysis if necessary at the end of the study. If it is necessary to analyse this data then QTcF (Fridericia) will be calculated programmatically using the reported ECG values (RR and QT).

QTcF = QT/RR^ (1/3) where RR is in seconds

For triplicate ECGs, the mean of the three ECG assessments will be used to determine the value at that time point.

3.4.7 Vital signs

Vital signs data obtained up until the 30 days from date of last dose of study treatment will be used for reporting. Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 3.4.7 below will be used.

The denominator in vital signs data should include only those patients with recorded data.

3.4.8 General considerations for safety assessments

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

- The time windows should be exhaustive so that data recorded at any time point has the potential to be summarised. 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 half way 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 for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy (with 4 weeks between scheduled assessments) are:

Day 29, visit window 2 – 42

Day 57, visit window 43 – 70

Day 85, visit window 71 – 98

Day 113, visit window 99 – 126

Day 141, visit window 127 – 154

Day 169, visit window 155 – 182

Day 197, visit window 183 – 210

Day 225, visit window 211 – 238

Day 253, visit window 239 – 266

Day 281, visit window 267 – 294

Day 309, visit window 295 – 322

Day 337, visit window 323 – 350

Note: Due to the differing assessment schedules the visit windows will be different for the different study treatments and endpoints.

- 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 summarised, 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 post baseline

values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.

– To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group visit data should only be summarised if the number of observations is greater than the minimum of 20 and > 1/3 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.

- Baseline will be defined as the last non-missing measurement prior to dosing with study treatment. For the re-treatment period for immunotherapy agents then baseline is similarly defined as the last non-missing measurement prior to the first dose on the re-treatment period. For laboratory data, any assessments made on day 1 will be considered pre-dose. Alternatively, 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 with no washout or other intervention in the screening period), the average can be taken as a baseline value. For non-numeric laboratory tests (ie 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 day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. Where safety data are summarised over time, study day will be calculated in relation to date of first treatment.

Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e., below the lower limit of quantification) or > x (i.e., above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

3.5 Biomarker Variables

PD-L1 expression status (positive, negative) is defined according to following criteria:

- Positive:- ≥ 25% tumour cell membrane positivity for PD-L1 at any intensity above background staining as noted on the corresponding negative control.

- Negative:- < 25% tumour cell membrane positivity for PD-L1 at any intensity above background staining as noted on the corresponding negative control.

3.6 Pharmacokinetic and Immunogenicity variables

Analyses to evaluate the pharmacokinetics and immunogenicity of MEDI4736 and tremeluminab will be performed by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

3.6.1 Population pharmacokinetics and exposure-response/safety analysis

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. The results of such an analysis will be reported separately from the main CSR. The PK, pharmacodynamic (PDx), 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-PDx methods.

3.6.2 Pharmacokinetic non-compartmental analysis

The PK analyses will be performed at AstraZeneca. The actual sampling times will be used in the PK calculations. PK concentration data and summary statistics will be tabulated. Individual and mean blood concentration-time profiles will be generated. PK parameters will be determined using standard non-compartmental methods. The following PK parameters will be determined after the first and steady-state doses: peak and trough concentration (as data allow). Samples below the lower limit of quantification will be treated as missing in the analyses.

3.6.3 Immunogenicity analysis

Immunogenicity results will be analyzed descriptively by summarizing the number and percentage of patients who develop detectable ADAs against MEDI4736 and tremelimumab. The immunogenicity titer and presence of neutralizing ADAs will be reported for samples confirmed positive for the presence of ADAs. Summaries will be based upon all patients from the safety population. The effect of immunogenicity on PK, pharmacodynamics, efficacy, and safety will be evaluated, if the data allow, but will be reported in a separate report

3.7 Health Resource Use

Health resource use outcome variables include the following:

- Length of hospital stay
- Reasons for hospitalisation
- Length of any time spent in an intensive care unit (ICU)

The length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation (length of hospital stay = end date of

hospitalisation – start date of hospitalisation + 1). Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation.

Sum of total duration of hospital stay will be considered for analysis if >20% of patients who were hospitalised were admitted to hospital more than one time during study period.

The length of ICU stay will be calculated in the using the same method as detailed above for the length of hospital stay.

4. ANALYSIS METHODS

4.1 General principles

The formal statistical analysis will be performed to test the main hypotheses:

- H0: No difference between MEDI4736 + tremelimumab and SoC
- H1: Difference between MEDI4736 + tremelimumab and SoC

The primary endpoint is PFS in all patients regardless of PD-L1 status using Investigator assessments per RECIST 1.1. The study has been sized to characterize the PFS benefit of MEDI4736 + tremelimumab versus SoC. The analysis will be performed when:

- (approximately) 338 PFS events have occurred across the MEDI4736+tremelimumab and SoC treatment groups (75% maturity) AND
- (approximately) 223 PFS events have occurred across the MEDI4736+tremelimumab and SoC treatment groups in patients with PD-L1 negative tumors (76% maturity).

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

- Descriptive statistics will be used for all variables, as appropriate, and will be presented by treatment group. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum, and maximum. For log transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarized by frequency counts and percentages for each category.
- Unless otherwise stated, percentages will be calculated out of the population total for the corresponding treatment arm.
- For continuous data the mean and median will be rounded to 1 additional decimal place compared to the original data. The standard deviation will be rounded to 2

additional decimal places compared to the original data. Minimum and maximum will be displayed with the same accuracy as the original data.

- For categorical data, percentages will be rounded to 1 decimal place.
- SAS® version 9.2 will be used for all analyses.

Baseline will be the last assessment of the variable under consideration prior to the intake of the first dose of IP, except for efficacy variables. For efficacy variables, baseline is defined as the last visit prior to randomization.

All data collected will be listed. Efficacy and PRO data will be summarized and analyzed based on the FAS. PK data will be summarized and analyzed based on the PK Analysis Set. Safety and treatment exposure data will be summarized on the Safety Analysis Set. Study population and demography data will be summarized based upon the FAS.

All outputs will be summarized by treatment arm for all randomized patients (ITT) and where required, for all randomized patients within the PD-L1-negative and PD-L1-positive subgroups.

Safety data will be summarised from the initial treatment period (ie the initial 12 months of treatment) only for the immunotherapy agents alongside the SOC agents. Safety data from the re-treatment period may also be summarised via a small set of headline summaries should there be sufficient number of patients re-treated to warrant this. Any safety summaries representing the re-treatment period will be based upon a subset of the safety analysis set representing patients who have had at least one dose of study treatment in the re-treatment period.

4.2 Analysis methods

Results of all statistical analysis will be presented using a 95% confidence interval (CI) and 2-sided p-value, unless otherwise stated.

The following table (Table 8) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint. Note, all endpoints compare MEDI4736 + tremelimumab versus SoC in all randomized patients (ITT population), unless otherwise indicated.

Table 8 Pre-planned statistical and sensitivity analyses to be conducted

Endpoints analyzed	Notes
Progression-free survival	Stratified log-rank tests for: • Primary analysis using Investigator RECIST 1.1 assessments • Secondary analysis using Investigator assessments (RECIST 1.1): – MEDI4736 + tremelimumab versus SoC for PD-L1-negative population (stratified only for histology) – MEDI4736 monotherapy versus SoC (ITT population) – MEDI4736 monotherapy versus SoC for PD-L1-positive population (stratified only for histology) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-negative population (stratified only for histology) – Sensitivity analyses using BICR assessments (RECIST 1.1) for a subset of patients sampled from the ITT population for BICR – Sensitivity analysis based on RECIST 1.1 modified for confirmation of progression using Investigator assessments – Exploratory analysis using BICR data for irRECIST for a subset of patients sampled from the ITT population for BICR

Endpoints analyzed	Notes
Objective response rate	Logistic regression for: • Secondary analysis for the ITT population and PD-L1-negative population using Investigator RECIST 1.1 assessments • Exploratory analysis using BICR data for irRECIST for a subset of patients sampled from the ITT population for BICR
Duration of response	Analysis methods as described by Ellis et al 2008 for: • Secondary analysis using Investigator assessments (RECIST 1.1)
Proportion of patients alive and progression free at 12 months	Hazard ratio using the Kaplan Meier estimates of progression free survival at 12 months (following method described by Klein et al 2007)
Time from randomization to second progression	Stratified log-rank test
Overall survival	Stratified log-rank test
Time to deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Stratified log-rank test

4.2.1 Multiple testing strategy

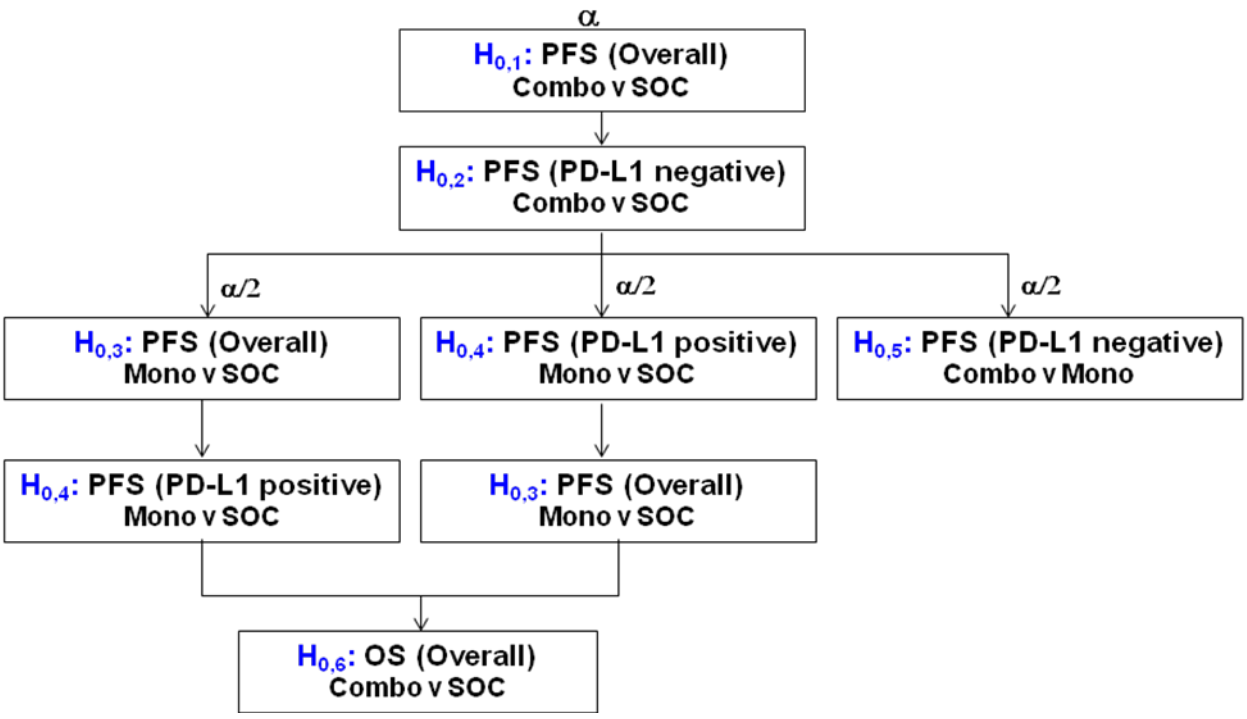
The multiple testing procedure (MTP) (as shown in Figure 3) will define which significance levels should be applied to the interpretation of the raw p-values for the primary endpoint of PFS and the key secondary endpoints intended for label claims.

Hypotheses will be tested using a MTP with an alpha-exhaustive recycling strategy (Burman et al 2009). With this approach, hypotheses will be tested in a pre-defined order, where PFS for MEDI4736 + tremelimumab versus SoC (ITT population) is tested first, followed by test of PFS for MEDI4736 + tremelimumab versus SoC (PD-L1 negative population). The other hypotheses will then be tested in the MTP using alpha (test mass) splitting and alpha recycling, where the test mass that becomes available after each rejected hypothesis is recycled to secondary hypotheses not yet rejected. This testing procedure stops when the entire test mass is allocated to non-rejected hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses. Figure 3 shows the MTP framework.

If the tests for MEDI4736 + tremelimumab versus SoC is statistically significant at the 0.05 level in both the PD-L1 all comers and the PD-L1-negative populations, then the

following 3 comparisons of (i) MEDI4736 monotherapy versus SoC in the all comer, (ii) MEDI4736 monotherapy versus SoC in the PD-L1 positive, and (iii) the MEDI4736 + tremelimumab versus MEDI4736 monotherapy in the PD-L1 negatives will be conducted, each at the 2.5% (2-sided) level.[1]

Figure 3 Multiple testing procedures for controlling the type 1 error rate



Combo MEDI4736 + tremelimumab combination therapy; Mono MEDI4736 monotherapy.

1. If either the null hypothesis for MEDI4736 + tremelimumab combination therapy versus SoC in the all comer (H1) or the PD-L1 negative population (H2) or both are true, then strong control of familywise error rate is maintained, since these tests will be conducted sequentially at the 0.05 level and will serve as gate keepers.

If H1 and H2 are both not true, ie, PFS in MEDI4736 + tremelimumab combination therapy is different from SoC in the all comer and the PD-L1-negative population, then the only opportunity where the overall type I error will not be preserved at the 0.05 level, by comparing the 3 hypotheses in the third level of the hierarchy at the 0.025 level, is if the null hypothesis is true in all the 3 cases; that is, there is no difference between the 2 arms in each comparison. However, given that H1 and H2 are both not true, it will not be possible for the remaining 3 hypotheses to be simultaneously true. Therefore, the overall false positive error from making these 3 comparisons at the 0.025 level should not be greater than 0.05. Hence the family-wise error rate will be maintained at the 0.05 level.

The details on how the alpha will be spent / controlled in all the possible scenarios are outlined below.

1. Test $H_{0,1}$ at level α

(a) If not significant, accept $H_{0,1}$ and stop procedure.

(b) If statistically significant, reject $H_{0,1}$ and continue with point 2.

2. Test $H_{0,2}$ at level α

(a) If not significant, accept $H_{0,2}$ and stop procedure.

(b) If statistically significant, reject $H_{0,2}$ and continue with point 3.

3. Test the three hypotheses $H_{0,3}$, $H_{0,4}$ and $H_{0,5}$ at level $\alpha/2$ each.

(a) If $H_{0,5}$ is not significant at level $\alpha/2$, accept $H_{0,5}$ and test $H_{0,3}$ and $H_{0,4}$.

(i) If neither $H_{0,3}$ nor $H_{0,4}$ is statistically significant at level $\alpha/2$, accept both $H_{0,3}$ and $H_{0,4}$ and stop the procedure.

(ii) If one of the two hypotheses $H_{0,3}$ and $H_{0,4}$ is statistically significant at level $\alpha/2$, then reject this hypothesis (but accept the other) and stop the procedure.

(iii) If both $H_{0,3}$ and $H_{0,4}$ are statistically significant at level $\alpha/2$, reject both $H_{0,3}$ and $H_{0,4}$ and continue testing $H_{0,6}$ at level $\alpha/2$.

- If $H_{0,6}$ is statistically significant at level $\alpha/2$, then reject $H_{0,6}$. Otherwise, accept $H_{0,6}$.

(b) If $H_{0,5}$ is statistically significant at level $\alpha/2$, reject $H_{0,5}$ and test $H_{0,3}$ and $H_{0,4}$ using Holm's procedure.

(i) If neither $H_{0,3}$ nor $H_{0,4}$ is statistically significant at level $\alpha/2$, accept both $H_{0,3}$ and $H_{0,4}$ and stop the procedure.

(ii) If one of the two hypotheses $H_{0,3}$ and $H_{0,4}$ is statistically significant at level $\alpha/2$, then reject this hypothesis and test the other hypothesis of the two at level α.

- If the other hypothesis is not significant at level α, accept this hypothesis and stop the procedure.

(iii) If both the two hypotheses $H_{0,3}$ and $H_{0,4}$ are statistically significant (either at level $\alpha/2$ each, or one of them having a p-value below $\alpha/2$ and the other a p-value below α), then , reject both $H_{0,3}$ and $H_{0,4}$ and continue testing $H_{0,6}$ at level α.

4. If $H_{0,6}$ is statistically significant at level α, then reject $H_{0,6}$. Otherwise, accept $H_{0,6}$.

Note also that 2 analyses of OS are planned:

1. At the time of the primary PFS analysis

2. A final OS analysis at 74% maturity

The alpha will be split between the 2 OS analyses using a bespoke spending function where a fixed significance level will be assigned at the first analysis and the remaining significance level assigned to the final analysis, taking account of correlation (Stone 2010):

- 1/10α available allocated to the first analysis,

- The significance level for the final analysis will be calculated using the software package EAST by selecting: boundary family, p-value, and Haybittle-Peto and specifying the fixed significance level used at the first analysis as well as the information fraction for each analysis. The information fraction is calculated as the number of events at the analysis time-point divided by the total number of events at the final analysis time-point.

Any non-statistically significant analyses at the time of the PFS analysis will not preclude further testing of OS. However, if the boundary is not crossed for the primary and key secondary endpoints, AstraZeneca may decide not to perform the final OS analysis.

4.2.2 Primary endpoint - progression-free survival

The primary PFS analysis will be based on the programmatically derived RECIST 1.1 using the Investigator tumor assessments. The analysis will be performed in the ITT population using a stratified log-rank test adjusting for PD-L1 tumor expression (positive versus negative) and histology (squamous versus non-squamous). The effect of MEDI4736 + tremelimumab versus SoC treatment will be estimated by the HR together with its corresponding 95% CI and p-value.

The covariates in the statistical modeling will be based on the values entered into IVRS at randomization, even if it is subsequently discovered that these values were incorrect.

The HR and its CI can be estimated from the stratified log-rank as follows (Berry et al 1991, Collett 2003, Selke and Siegmund 1983):

$$HR = \exp\left(\frac{U}{V}\right)$$

$$95\%\,\text{CI for HR} = \left(\exp\left\{\frac{U}{V} - \frac{1.96}{\sqrt{V}}\right\}, \exp\left\{\frac{U}{V} + \frac{1.96}{\sqrt{V}}\right\}\right)$$

Where $U = \sum_k U_k = \sum_k \sum_i (d_{1ki}, -e_{1ki})$ is the stratified log-rank test statistic obtained from the SAS LIFETEST procedure , $\sqrt{V} = \sqrt{\sum_k V_k}$, is its standard deviation, k denotes the stratum and d_{1ki} and e_{1ki} are the observed and expected events in Group 1, stratum k.

A secondary analysis of PFS will be performed to compare MEDI4736 monotherapy versus SoC as well as to compare MEDI4736 + tremelimumab versus MEDI4736 monotherapy. These analyses will be performed using the same methodology as for primary endpoint described above.

Kaplan-Meier plots of PFS will be presented by treatment arm, and by treatment arm and PD-L1 tumor status subgroup, where appropriate. Summaries of the number and percentage of patients experiencing a PFS event and the type of event (RECIST 1.1 or death) will be provided along with median PFS for each treatment.

The assumption of proportionality will be assessed, initially only with regards to the primary treatment comparison. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, by fitting a time-dependent covariate to assess the extent to which this represents random variation. If a lack of proportionality is evident, the variation in treatment effect will be described by presenting piecewise HR calculated over distinct time-periods. In such circumstances, the HR can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. If lack of proportionality is found, this may be a result of treatment-by-covariate interactions, which will be investigated. In addition, the KM curve along with landmark analyses (e.g., one year PFS rate) will also help in understanding the treatment benefit.

The treatment status at progression of patients at the time of analysis will be summarised. 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. This will also provide distribution of number of days prior to progression for the patients who have discontinued treatment.

Additional supportive summaries/graphs

In addition, the number of patients prematurely censored will be summarised by treatment arm together with baseline prognostic factors of the prematurely censored patients. A patient would be defined as prematurely censored if they had not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than one scheduled tumour assessment interval plus 2 weeks (8 weeks if time period between randomisation and DCO for that patient is 48 weeks or less; 10 weeks otherwise) prior to the DCO date.

Additionally, summary statistics will be given for the number of days from censoring to data cut-off for all censored patients.

A summary of the duration of follow-up will be summarised using median time from randomisation to date of censoring (date last known to be non-progressor) in censored (not progressed) patients only, presented by treatment group.

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

Summaries of the number and percentage of patients who miss two or more consecutive RECIST assessments and the number of patients who miss one RECIST assessment will be presented for each treatment group.

All of the collected RECIST 1.1 data will be listed for all randomised patients. In addition, a summary of new lesions (i.e., sites of new lesions) will be produced.

Sensitivity Analyses

The following sensitivity analyses will only be performed for the primary treatment comparison.

The analysis will be based on the PFS programmatically derived RECIST 1.1 using the Investigator tumor assessments.

- Evaluation-Time bias

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time points. The midpoint between the time of progression and the previous evaluable RECIST assessment will be analyzed using a log-rank test. For patients whose death was treated as PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust even in highly asymmetric assessment schedules (Sun and Chen 2010).

- Attrition bias

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the PFS analysis is reversed.

- Ascertainment bias

Ascertainment bias will be assessed by analyzing the BICR assessments, but will only be performed if BICR is required for all patients. The stratified log-rank test will be repeated on these data. The HR and CI will be presented.

If there is an important discrepancy between the primary analysis using investigator assessments and this sensitivity analysis using BICR assessments, then the proportion of patients with site but no central confirmation of progression will be summarised; such patients have the potential to introduce bias in the central review due to informative censoring. An approach that imputes an event at the next visit in the central review analysis may help inform the most likely HR value (Fleischer et al 2011), but only if an important discrepancy exists.

Disagreements between investigator and central reviews of RECIST progression will be presented for each treatment group. The summary will include the early discrepancy rate which is the frequency of central review declared progressions before the investigator review as a proportion of all central review progressions and the late discrepancy rate which is the frequency of central review declared progressions after the investigator review as a proportion of all discrepancies.

Secondary Analysis

A secondary analysis of PFS based on the programmatically derived RECIST 1.1 using the Investigator tumor assessments will be performed to compare MEDI4736 + tremelimumab versus SoC in the PD-L1-negative population. This analysis will be performed using a stratified log-rank test adjusting solely for histology (squamous versus non-squamous). The effect of treatment will be estimated by the HR together with its corresponding 95% CI and p-value. The HR and CI will be estimated using the same approach as specified above for the primary analysis of PFS. In addition, a secondary analysis of PFS will be performed similarly to compare MEDI4736 + tremelimumab versus MEDI4736 monotherapy in the PD-L1-negative population as well as MEDI4736 monotherapy versus SoC in the PD-L1-positive population.

Exploratory Analyses

An exploratory analysis of PFS using Investigator assessment based on RECIST 1.1 modified for confirmation of progression as well as PFS based on BICR assessments according to irRECIST criteria will be performed. The stratified log-rank test used for the primary analysis of PFS will be repeated. The HR and CI will be presented.

A forest plot illustrating the hazard ratio and 95% confidence interval will be provided to compare the primary and sensitivity analyses of progression free survival.

Subgroup analyses and a forest plot will be generated comparing PFS between treatments.

No adjustment to the significance level for testing will be made since all these subgroup and sensitivity analyses will be considered supportive of the primary analysis of PFS.

The effect of covariates upon the HR estimate and the consistency of treatment effect between subgroups will be analysed for PFS.

Subgroup Analyses

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1 using Investigator assessments) between MEDI4736 + tremelimumab versus SoC in the following subgroups of the FAS (but not limited to):

- Sex (male versus female)
- Age at randomization (<65 versus ≥65 years of age)
  - This will be determined from the date of birth (BIRTHDAT in the DEM module) and date of randomisation (RND_DAT in the CRIT1 module) on the eCRF at screening. Patients with a partial date of birth (ie for those countries where year of birth only is given) will have an assumed date of birth of 1st Jan [given year]). Patients with a missing age value will be included using the mean age (overall FAS) and categorised accordingly.
- PD-L1 status (positive versus negative)
- Histology (squamous versus non-squamous)
- Smoking (smoker versus non-smoker [never smoker])
  - This will be determined from the response to 'Nicotine Use Occurrence' (SU module) on the eCRF at screening. Patients with a missing smoking status will be included in the 'smoker' category.
- Race (Asian versus non-Asian)

The subgroup analyses for the stratification factors will be based on the values entered into the IVRS, all other factors will be based on values recorded on the eCRF as indicated above. Other baseline variables may also be assessed if there is clinical justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors. No adjustment to the significance level for testing of the subgroup and sensitivity analyses will be made since all these analyses will be considered supportive of the analysis of PFS. These hazard ratios and associated two-sided 95% CIs will be summarised and presented on a forest plot, along with the results of the overall primary analysis.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events in a subgroup), the relationship between that subgroup and PFS will not be formally analysed. In this case, only descriptive summaries will be provided.

Effect of covariates on HR estimate

Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. Before embarking on more detailed modeling, an initial model will be constructed, containing treatment and the stratification factors alone, to ensure that any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

Consistency of treatment effect between subgroups

The presence of quantitative interactions will be assessed by means of an overall global interaction test. This will be performed in the overall population by comparing the fit of a Cox proportional hazards model including treatment, all covariates, and all covariate-by-treatment interaction terms, with one that excludes the interaction terms and will be assessed at the 2-sided 10% significance level. If a covariate does not have more than 20 events per level (of the covariate) it will be included as a covariate in the model but the covariate-by-treatment interaction term will be omitted. If the fit of the model is not significantly improved, then it will be concluded that, overall, the treatment effect is consistent across the subgroups.

If the global interaction test is found to be statistically significant, an attempt to determine the cause and type of interaction will be made. Stepwise backwards selection will be performed on the saturated model, whereby (using a 10% significance level throughout) the least significant interaction terms are removed one-by-one and any newly significant interactions re-included until a final model is reached where all included interactions are significant and all excluded interactions are non-significant. Throughout this process, all main effects will be included in the model regardless of whether the corresponding interaction term is still present. This approach will identify the factors that independently alter the treatment effect and prevent identification of multiple correlated interactions.

Any quantitative interactions identified using this procedure will then be tested to rule out any qualitative interaction using the approach of Gail and Simon 1985).

Additionally, for each subgroup, the HR (MEDI4736 + tremelimumab: SoC) and 95% CI will be calculated from a single model that contains treatment, factor (only the factor that determines the subgroup) and treatment-by-factor interaction term. These will be presented on a forest plot including the HR and 95% CI from the overall population.

4.2.3 Objective response rate

The ORR will be based on the programmatically derived RECIST using the Investigator tumor data. The ORR will be compared between MEDI4736 + tremelimumab versus SoC using logistic regression models adjusting for the same factors as the primary endpoint (PD-L1 tumor expression and histology). The results of the analysis will be presented in terms of an odds ratio (an odds ratio greater than 1 will favour MEDI4736+tremelimumab) together with its associated profile likelihood CI (e.g. using the option ‘LRCI’ in SAS procedure GENMOD) and p-value (based on twice the change in log-likelihood resulting from

the addition of a treatment factor to the model). This analysis will be performed in the ITT population and PD-L1-negative population. The analysis of the PD-L1-negative patients will be performed using a logistic regression model adjusting for only histology.

ORR by irRECIST criteria using BICR assessments will also be reported in the ITT population.

If there are not enough responses for a meaningful analysis using logistic regression then a Fisher's exact test using mid p-values will be presented.

The mid-p-value modification of the Fisher exact test amounts to subtracting half of the probability of the observed table from Fisher's p-value.

Fisher's exact test mid p-value = Two-sided p-value – (Table probability / 2)

Summaries will be produced that present the number and percentage of patients with a tumor response (CR/PR). Overall visit response data will be listed for all patients (ie, the FAS). For each treatment arm, best objective response (BoR) will be summarized by n (%) for each category (CR, PR, SD, PD, and NE). No formal statistical analyses are planned for BoR.

4.2.4 Duration of response

In order to analyze the DoR between MEDI4736 + tremelimumab and SoC, the expected duration of response (EDoR) will be derived for each treatment arm (Ellis et al 2008) using the Investigator tumor data. The EDoR is the product of the proportion of patients responding to treatment and the mean DoR in responding patients and provides an estimate based on all randomized patients. Treatments will be compared by calculating the ratio of EDoRs, using the weibull distribution (unless the data suggests otherwise) for DoR in responding patients. Additionally, descriptive data will be provided for the DoR in responding patients (i.e. median duration of response and 95% CIs) by treatment arm, including the associated Kaplan-Meier curves (without any formal comparison of treatment arms or p-value attached). This analysis will be performed on the ITT population.

4.2.5 Proportion of patients alive and progression free at 12 months

The APF12 (where 12 months equates to study day 366) will be summarized (using the Kaplan-Meier curve) and presented by treatment arm. APF12 will be compared between MEDI4736 + tremelimumab and SoC by using the Kaplan-Meier estimator of PFS at 12 months for each treatment to obtain the HR. The HR and CI will be presented using the following approach (Klein et al 2007):

- The *HR(group1:group2)* is estimated as $\frac{\ln \hat{S}_1(t)}{\ln \hat{S}_2(t)}$
- The variance for *ln(HR)* is estimated as $\frac{\hat{\sigma}_1(t)^2}{\ln^2 S_1(t)} + \frac{\hat{\sigma}_2(t)^2}{\ln^2 S_2(t)}$

where $\hat{\sigma}_i(t)^2 = \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$ is the variance for ln{S(t) derived from greenwood's formula for the variance of *S(t)* and can be estimated from standard software packages, where d_i and n_i refer to the number of events and patients at risk for each risk set.

The *ln(HR)* and its variance in each strata will be estimated and combined by weighting inversely proportionately according to each within stratum variance (Whitehead and Whitehead 1991).

This analysis will be performed in the ITT population.

4.2.6 Time from randomization to second progression

Second progression (PFS2) in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding 95% CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

For supportive purposes, the time to the start of subsequent therapy will be analyzed using the same methodology and model. The HR for the treatment effect together with its 95% CI will be presented. In addition, a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment arm and the time between progression and starting subsequent therapy will be assessed. No multiplicity adjustment will be applied as these are viewed as supportive endpoints.

A summary table of first subsequent therapies by treatment arm will be provided, as well as response to first subsequent therapy by treatment arm.

This analysis will be performed in the ITT population.

4.2.7 Overall survival

OS in the ITT population will be analyzed using a stratified log-rank tests, using the same methodology as described for the primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. 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 for each treatment.

The assumption of proportionality will be assessed in the same way as for PFS.

OS will be analysed at the time of the primary PFS analysis, and at final OS follow-up (when 333 OS events have occurred)

The alpha will be split between the 2 OS analyses using the hierarchical testing strategy as already described in Section 4.2.1.

Impact of switching (crossover outside of this study) to immunotherapies (or other potentially active investigational agents) on OS analyses

Exploratory analyses of OS adjusting for the impact of subsequent immunotherapy or other investigational treatment may be performed, if a sufficient proportion of patients switch. Methods such as Rank Preserving Structural Failure Time (Robins and Tsiatis 1991), Inverse Probability of Censoring Weighting (Robins 1993) and other methods in development will be explored. The decision to adjust and the final choice of methods will be based on a blinded review of the data and the plausibility of the underlying assumptions. Baseline and time-dependent characteristics will be explored, and summaries of baseline characteristics will be summarised by treatment arm, splitting between those that have and haven't switched at the time of the analyses. Further detail will be provided in the Payer Analysis Plan. These analyses are intended to support reimbursement appraisals.

Subsequent therapies received after discontinuation of treatment will be summarised and listed by treatment group. Patients who subsequently received an immunotherapy agent or entered an immunotherapy trial will be summarised and listed by treatment arm according to line of subsequent therapy, i.e. immediately after immunotherapy or as a later line.

4.2.8 Change in tumour size

The absolute values and percentage change in target lesion tumour size from baseline will be summarized using descriptive statistics and presented at each timepoint for each treatment arm. The best change in target lesion tumour size from baseline, (where best change in target lesion size is the maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction) will also be summarised and presented for each treatment arm.

Tumour size will also be presented graphically using waterfall plots for each treatment arm, to present each subject's best percentage change in tumour 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 tumour size levels will be added to the plots, which correspond with the definitions of progression and 'partial' response respectively. On each of the waterfall plots the histology classification (Squamous versus All other) of each patient will be indicated. Additional waterfall plots showing percentage change in tumour size at specific timepoints may be produced if it is felt that these are warranted to provide greater clarity.

The above outputs will be programmed for the investigator RECIST assessments.

4.2.9 Patient reported outcomes

The PRO endpoints that have been identified as primary are EORTC QLQ-C30 time to deterioration in Global Health Status/ QoL (items 29 and 30) and EORTC QLQ-LC13 time to deterioration in Dyspnoea, Cough, Haemoptysis, and Pain These are not part of the main multiple testing procedure and as supportive endpoints will need a Bonferroni adjustment to

the significance level to aid interpretation. Therefore, these 5 endpoints will be tested at a 1% significance level and 99% CIs will be produced.

The other time to symptom deterioration endpoints will be tested at a 5% significance level and 95% CIs will be produced.

4.2.9.1 EORTC QLQ-C30

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 3 symptom scales and the 5 individual symptom items will be produced. Similarly, a summary of QoL/function improvement rate for each of the 5 function scales (physical, role, emotional, cognitive, and social) and global health status/QoL will be produced.

Summaries of original and change from baseline values of each symptom scale/item, the global HRQoL score and each functional domain will be reported by visit for each treatment arm. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration) will also be produced for each treatment arm.

4.2.9.2 EORTC QLQ-LC13

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment arm. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 6 individual symptom items will be produced.

Summaries of original and change from baseline values of each symptom (dyspnea, cough, hemoptysis, chest pain, arm/shoulder pain, other pain) and each treatment-related side effect (sore mouth, dysphagia, peripheral neuropathy and alopecia) will be reported by visit for each treatment arm. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each visit for each ordinal

symptom item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration) will also be produced for each treatment arm.

4.2.9.3 PRO-CTCAE

Data from the PRO-CTCAE will be summarized using FAS. The number (%) of patients with each level of response for each CTCAE item at baseline and over time will be summarized. A bar chart of the incidence by visit will be presented for each CTCAE. Further summaries to explore the data (i.e. the severity of symptoms) may be produced.

4.2.9.4 Patients' Global Impression of Change

PGIC data will be presented using summaries and descriptive statistics based on the FAS.

4.2.9.5 EQ-5D-5L

Descriptive statistics, graphs, and listings will be reported for health state utility values and the visual analogue scale by visit, as well as the change in these scores from baseline. To support future economic evaluations, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post-progression, and will be outlined in the payer analysis plan.

4.2.10 Healthcare resource use

An exploratory health economic analysis of hospital episodes including type of contact (hospitalization, outpatient, or day case), reason, length of stay by ward type (including intensive care unit), procedures, and tests may be undertaken to examine the impact of disease and treatment on resource use to primarily support the economic evaluation of MEDI4736+tremelimumab in comparison to SoC, and will be outlined in the Payer Analysis Plan. This would include providing descriptive statistics as appropriate, including means, median, and ranges.

4.2.11 Safety data

Safety and tolerability data will be presented by treatment arm using the safety population. Safety data will be summarized only. No formal statistical analyses will be performed on the safety data.

Any safety summaries examining retreatment will be produced separately.

Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of MedDRA preferred terms and CTCAE grade) will be listed individually by patient. The number of patients experiencing each AE will be summarized by treatment arm and CTCAE grade. Additionally, data presentations of the rate of AEs per person-years at risk may be produced. Any safety summaries examining re-treatment with MEDI4736+tremelimumab and MEDI4736 monotherapy will be produced separately, if required.

Other safety data will be assessed in terms of physical examination, clinical chemistry, hematology, vital signs, and ECGs. Exposure to MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC will be summarized. Time on study, MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC, dose delays/interruptions and dose reductions will also be summarized. At the end of the study, appropriate summaries of all safety data will be produced.

The following sections describe the planned safety summaries. However, additional safety tables may be required to aid interpretation of the safety data.

Adverse Events

All AEs, both in terms of current MedDRA preferred term and CTCAE grade, will be summarised descriptively by count (n) and percentage (%) for each treatment group. The current MedDRA dictionary will be used for coding. The majority of the AE summaries, unless stated otherwise, will be based on TEAEs. Any AE occurring before study treatment (i.e. before the administration of the first dose on Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as 'pre-treatment'. However, any AE occurring before the administration of the first dose on Study Day 1 that increases in severity after the first dose will be regarded as treatment emergent and thus will be included in the the majority of summary tables.

AEs observed up until 90 days following discontinuation of the immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) and the Standard of Care agent or until the initiation of the first subsequent anti-cancer therapy following discontinuation of treatment (whichever occurs first) will be used for reporting of all of the AE summary tables. This will more accurately depict AEs attributable to study treatment only as a number of AEs up to 90 days following discontinuation of the immunotherapy agents and the Standard of Care agent are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, all of the AE summaries will also be produced containing AEs observed up until 90 days following discontinuation of the immunotherapy agents and the Standard of Care agent (ie without taking subsequent anti-cancer therapy into account).

A selection of AE summaries mayalso be produced containing AEs (by system organ class and preferred term) observed from the initiation of the first subsequent anti-cancer therapy following discontinuation of study treatment until 90 days following discontinuation of immunotherapy agents (ie summarising those AEs experienced by patients taking subsequent therapy during the AE collection follow-up window post discontinuation of study treatment). These outputs will only be produced if the number of AEs observed warrant the inclusion of such outputs for interpretational purposes. Any data post 90 days last dose for immunotherapy agents and Standard of Care agents will be presented in a separate summary that presents any events that occur prior to dosing or starting more than 90 days after discontinuing treatment.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved) and investigator's assessment of severity and relationship to study drug. Frequencies and percentages of patients reporting each preferred term will be presented (i.e. multiple events per patient will not be accounted for apart from any on episode level summaries which may be produced).

Summary information (the number and percent of patients by system organ class and preferred term separated by treatment group) will be tabulated for:

- All AEs
- All AEs causally related to study medication (as determined by the reporting investigator)
- AEs with CTCAE grade 3 or higher
- AEs with CTCAE grade 3 or higher, causally related to study medication (as determined by the reporting investigator)
- AEs with outcome of death
- AEs with outcome of death causally related to study medication (as determined by the reporting investigator)
- AEs by outcome
- All SAEs
- All SAEs causally related to study medication (as determined by the reporting investigator)
- SAEs leading to discontinuation of study medication
- AEs leading to discontinuation of study medication
- AEs leading to discontinuation of study medication, causally related to study medication (as determined by the reporting investigator)
- AEs leading to hospitalization
- AEs leading to dose delay of study medication
- Other significant AEs
- Other significant AEs causally related to study medication (as determined by the reporting investigator)

- Immune mediated AEs (as determined by the reporting investigator)
- Infusion reaction AEs (as determined by the reporting investigator)

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

Each AE event rate (per 100 patient years) will also be summarised by preferred term within each system organ class for the output summarising all AEs. For each preferred term, the event rate is defined as the number of patients with that AE divided by the total drug exposure of patients at risk of AE. The denominator is calculated as the total over each patient of days from first dose to the earlier of the date of onset of the event or the last day of study medication.

Summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, system organ class, preferred term and treatment group.

Fluctuations observed in CTCAE grades during study will be listed for those AEs which are CTCAE ≥ 3.

In addition, all AEs will be listed.

Deaths

Two summaries of all deaths will be provided with number and percentage of patients by treatment group, categorised as:

- Total number of death (including deaths within 90 days after last dose of study medication and deaths occur outside of 90 days post dosing window)
- Death related to disease under investigation ONLY, as determined by investigator (including deaths within 90 days after last dose of study medication and deaths occur outside of 90 days post dosing window)
- TEAE with outcome of death ONLY and onset date prior to initiation of subsequent anti-cancer therapy

- AE with outcome of death ONLY and onset date falling after 90 days follow-up period or initiation of subsequent anti-cancer therapy (whichever is earlier) (1st summary)
- Death related to disease under investigation, as determined by the investigator, and with TEAE with outcome of death and onset date prior to initiation of subsequent anti-cancer therapy
- Death related to disease under investigation, as determined by the investigator, and with AE with outcome of death and onset date falling after 90 days follow-up period or initiation of subsequent anti-cancer therapy (whichever is earlier)
- Death occurred 90 days after last dose or initiation of subsequent anti-cancer therapy (whichever is earlier), and unrelated to AE or disease under investigation
- Patients with unknown reason for death.
- Other deaths

These summaries will be produced twice; firstly accounting for subsequent therapy and, secondly, without taking subsequent therapy into consideration

Adverse events of special interest

Preferred terms used to identify adverse events of special interest (as defined in section 3.4.1) will be listed before DBL and documented in the Study Master File. Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI grouping.. Groupings will be based on preferred terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

Summaries of the above-mentioned grouped AE categories will include number (%) of patients who have:

- At least adverse event of special interest presented by outcome
- At least one adverse event of special interest causally related to study medication
- At least one adverse event of special interest leading to discontinuation of study medication

A summary of total duration (days) of AESI will be provided for events which have an end date and this will be supported by summaries of ongoing AESIs at death and, separately, at data cut-off.

Additionally, there will be several summaries of AESIs requiring concomitant treatment, and particularly the relationship of AESIs to the use of immunosuppressive agents (ie, depicting which AESI triggered immunosuppressive use) and, separately, to the use of immunosuppressive agents at high doses.

Summary of long term tolerability

To assess long term tolerability, provided that there are a sufficient number of patients with events to warrant it, prevalence plots, life table plots and cumulative incidence plots will be presented for each of the AESI grouped terms and any other events considered important after review of the safety data, provided there are ≥ 10 events.

A prevalence plot provides information on the extent to which the events may be an ongoing burden to patients. The prevalence at time t after first dose of study treatment is calculated as the number of patients experiencing the event divided by the number of patients receiving study treatment or in safety follow-up at time t; generally, t is categorised by each day after dosing. The prevalence will be plotted over time and presented for each treatment group separately. Multiple occurrences of the same event are considered for each patient but a patient is only counted in the numerator whilst they are experiencing one of the occurrences of the event. These plots will only be produced for AESIs that have ≥10 events.

For each AE, median time to first onset of the AE from the date of first dose will be presented in patients in the safety analysis set by treatment group. Patients who did not experience the AE will be censored at the end of their safety follow-up. Summary tables of time to first onset for each AE will also be produced (e.g. 1-28 days, 29-56 days, 57-84 days, 85-112 days, >112 days). Median duration of the AE will be presented in patients who experienced each AE.

A life table plot can be used to describe the time to onset of the event and specifically when patients are at most risk of first experiencing the event. The hazard, or in other words, the probability of having an AE in a specified time period (e.g. 0-1 months, 1-3 months, 3-6 months, etc.) given that the patient reaches that time period without having an event is plotted for each time period. These plots will only be produced for AESIs that have ≥10 events.

A cumulative incidence plot is a plot of the raw cumulative incidence and cumulative incidence function over time with the treatment groups presented on separate plots. The raw cumulative incidence is the actual probability that a patient will have experienced their first occurrence of the event by a given time point. The cumulative incidence function estimates the cumulative incidence if the data cut-off had not been imposed and all patients had completed safety follow-up (Pintilie M.). These plots will only be produced for AESIs that have ≥10 events.

Laboratory assessments

Data obtained up until the 90 days following discontinuation of immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) or following discontinuation of the Standard of Care agent or until the initiation of the first subsequent therapy following

discontinuation of treatment (whichever occurs first) will be used for reporting. This will more accurately depict laboratory toxicities attributable to study treatment only as a number of toxicities up to 90 days following discontinuation of immunotherapy agents or the Standard of Care agent are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, summaries of laboratory data will also be produced containing data collected up until 90 days following discontinuation of the immunotherapy agents or Standard of Care agent (ie, without taking subsequent therapy into account).

A small selection of summaries of laboratory data may also be produced containing data from initiation of the first subsequent therapy following discontinuation of study treatment until 90 days following discontinuation of immunotherapy agents or the Standard of Care agent (ie summarising the laboratory data collected on patients taking subsequent therapy during the safety collection follow-up window post discontinuation of study treatment). These outputs will only be produced if the number of laboratory toxicities observed warrant the inclusion of such outputs for interpretational purposes. Any data post 90 days last dose for immunotherapy agentsor Standard of Care agents will not be summarised.

Data summaries will be provided in International System (SI) of units.

Scatter plots (shift plots) of baseline to maximum value on treatment (i.e. on-treatment is defined as data collected between the start of treatment and the relevant follow-up period following the last dose of study treatment) may be produced for certain parameters if warranted after data review (e.g. ALT, AST, ALP, Total bilirubin, Corrected calcium, LDH, Magnesium, Sodium, Potassium, Glucose, Creatinine, Urea nitrogen, and TSH, T3 and T4).

Scatter plots (shift plots) of baseline to minimum value on treatment (defined as above) may be produced for certain parameters if warranted after data review (e.g. Haemoglobin, Lymphocyte (count, absolute); Neutrophils (count, absolute); Platelet count; Albumin, Total protein, Corrected calcium, Magnesium, Sodium, Potassium, Glucose and TSH, T3 and T4).

Box-plots of absolute values by week, and box-plots of change from baseline by week, may be presented for certain parameters if warranted after data review (e.g. Haemoglobin; Neutrophil count, absolute; Lymphocyte count, absolute; Platelet count; AST; ALT; ALP; Total bilirubin; Albumin; Total protein; Corrected calcium; LDH; Sodium; Potassium; Creatinine, Urea nitrogen and TSH, T3 and T4).

For continuous laboratory assessments including the thyroid test parameters TSH, T3 and T4, absolute value, change from baseline and percentage change from baseline will be summarised using descriptive statistics at each scheduled assessment time by actual treatment group

Shift tables for laboratory values by worst CTC 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 CTC grade shift outputs will be produced are:

- Haematology: Haemoglobin, Leukocytes, Lymphocytes, absolute count, Neutrophils, absolute count, Platelets

- Clinical chemistry: ALT, AST, ALP, Total bilirubin, Albumin, Magnesium – hypo and – hyper, Sodium – hypo and – hyper, Potassium – hypo and – hyper, Corrected calcium – hypo and – hyper, Glucose – hypo and – hyper, Creatinine

For the parameters with no CTCAE grading that are listed in the CSP, shift tables from baseline to worst value on-treatment will be provided. Additional summaries will include a shift table for urinalysis (Bilirubin, Blood, Glucose, Ketones, Protein) comparing baseline value to maximum on-treatment value.

Shift tables showing baseline to maximum and baseline to minimum will be produced for TSH, T3 and T4.

Liver Enzyme Elevations and Hy's law

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

- Elevated ALT, AST, and Total bilirubin during the study

  – ALT ≥ 3x –≤ 5x, > 5x – ≤8x, > 8x. >10x and >20x Upper Limit of Normal (ULN) during the study

  – AST ≥ 3x- –≤ 5x, > 5x – ≤8x, > 8x. >10x and >20x ULN during the study

  – Total bilirubin ≥2x-≤3x, >3x-≤5x, >5x ULN during the study

  – ALT or AST ≥3x - ≤5x, >5x - ≤8x, >8x, >10x and >20x ULN during the study

  – ALT or AST ≥3x ULN and Total bilirubin ≥2x ULN during the study (Potential Hy's law): 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 ≥ 3x ULN plus Total bilirubin ≥ 2x ULN or AST ≥ 3x ULN plus Total bilirubin ≥ 2x ULN at any visit.

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

Plots of ALT and AST vs. Total bilirubin by treatment group 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.

Assessment of Thyrotoxicity

After the discontinuation of the study medication, the thyroid function tests, TSH, T3 and T4, were evaluated at 30 days after last dose, hence, the analysis of thyroid function tests will be based on data up to 30 days after the last dose of study medication or date of initiation of subsequent therapy (whichever occurs first),Absolute value and change from baseline will be summarised using descriptive statistics at each scheduled assessment time.

Shift tables showing baseline to maximum and baseline to minimum will also be produced for TSH, T3 and T4, as deemed necessary.

ECGs

ECG data obtained up until the safety follow-up will be included in the summary tables. Overall evaluation of ECG is collected at each visit in terms of normal or abnormal, and the relevance of the abnormality is termed as “clinically significant” or “not clinically significant”. A shift table of baseline evaluation to worst evaluation will be produced.

Vital signs

Vital signs data obtained up until the 30 day safety follow-up visit will be included in the summary tables.

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

Vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, temperature, respiratory rate and weight) will be summarised over time in terms of absolute values and changes from baseline at each scheduled measurement by actual treatment group.

Time to Subsequent Therapy from discontinuation of study treatment

A descriptive summary will be produced for time to subsequent therapy from discontinuation of study treatment. This summary is supportive of the Adverse Event and Laboratory data outputs.

Physical examination

All individual physical examination data will be listed only.

Other Safety Data

Data from positive pregnancy tests will be listed only.

4.2.12 WHO performance status

All WHO performance status will be summarised over time for the ITT population.

4.2.13 PK data (MEDI4736 monotherapy and MEDI4736+tremelimumab arms only)

Pharmacokinetic concentration data will be listed for each patient and each dosing day, and a summary provided for all evaluable patients in the PK analysis population. These outputs will be produced by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

Immunogenicity analysis

Immunogenicity results will be listed by patient and a summary will be provided of the number and percentage of patients who develop detectable anti-MEDI4736 and antitremelimumab antibodies based on the safety population. The immunogenicity titre and neutralizing ADA data will be listed for samples confirmed positive for the presence of anti-MEDI4736 antibodies and/or anti-tremelimumab antibodies.

The effect of immunogenicity on PK, PDx, efficacy and safety will be evaluated if data allow. These outputs will be produced by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

4.2.14 PK/PDx relationships (MEDI4736 monotherapy and MEDI4736+tremelimumab)

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modelling approach. These outputs will be produced by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

4.2.15 Biomarker data

If applicable, the relationship of exploratory biomarkers to OS, PFS, ORR and DoR will be presented for a subset of patients in the ITT population who are evaluable for each biomarker.

This will be assessed using similar summary and graphical representations to those that are outlined for the efficacy outputs in Sections 4.2.2 to Section 4.2.7.

PD-L1 expression determined by immunohistochemistry 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.

4.2.16 Demographic and baseline characteristics data

The following will be summarised for all patients in the FAS (unless otherwise specified) by treatment group:-

- Patient disposition (including screening failures and reason for screening failure)
- Important protocol deviations
- Inclusion in analysis populations

- Demographics (age, age group[<50, >=50-< 65, ≥ 65 - <75 years and ≥ 75 years], sex, race and ethnicity)
- Patient characteristics at baseline (height, weight, weight group, body mass index (BMI) and body mass index group)
- Patient recruitment by country and centre
- Previous disease-related treatment modalities
- Number of regimens of previous chemotherapy at baseline
- Previous lung cancer therapy
- Disease characteristics at baseline (WHO performance status, primary tumour location, histology type, tumour grade and overall disease classification, best response to previous therapy)
- Extent of disease at baseline
- TNM classification at baseline
- Disease related medical history (past and current)
- Relevant surgical history
- Physical examination at baseline
- Time from most recent disease progression to start of study treatment
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use, categorised (never, current, former)
- Stratification factors as per IVRS and eCRF data

The AZ drug dictionary (AZDD) will be used for concomitant medication coding.

Patient disposition data will also be summarised at the time of OS analysis.

4.2.17 Treatment exposure

The following summaries related to study treatment will be produced for the safety analysis set by randomised treatment group:

- Total exposure of each treatment group.
- Actual exposure of each treatment group.
- Total number of cycles received.
- Reasons for dose delays and infusion interruptions of MEDI4736 and tremelimumab and reasons for dose delays/interruptions, dose reductions and dose modifications for the Standard of Care agents. Dose interruptions will be based on investigator initiated dosing decisions.
- Number of dose delaysand duration of delays of MEDI4736. In addition, delays due to AEs and due to reasons other than AEs will be summarized separately.
- Number of infusions received.
- PID and RDI (percentage intended dose and relative dose intensity) of MEDI4736, tremelimumab and Standard of Care agents.
- PID2 and RDI2 (PID and RDI for treatment through progression) of MEDI4736, tremelimumab and Standard of Care agents.
- Exposure over time will be plotted.

For patients on study treatment at the time of the PFS and OS analysis, the DCO date will be used to calculate exposure.

4.2.18 Subsequent Therapy

Subsequent therapies received after discontinuation of study treatment will have summaries produced by treatment group, together with number of regimens received.

5. INTERIM ANALYSES

Interim safety monitoring will be conducted by an IDMC. No interim analysis will be performed for efficacy.

5.1 Independent Data Monitoring Committee

This study will use an external IDMC to assess ongoing safety analyses. The committee will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review the safety data from the study. The IDMC will

meet at least every 6 months thereafter. Following each meeting, the IDMC will report to the sponsor and may recommend changes in the conduct of the study.

This committee will be composed of therapeutic area experts and biostatisticians, who are not employed by AstraZeneca/MedImmune and do not have any major conflict of interest.

Following the reviews, the IDMC will recommend whether the study should continue unchanged, be stopped, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to AstraZeneca/MedImmune. The report will include the recommendation and any potential protocol amendments, and will not contain any unblinding information.

The final decision to modify or stop the study will sit with the sponsor. The sponsor or IDMC may call additional meetings if at any time there is concern about the safety of the study.

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

The safety of all AstraZeneca/MedImmune clinical studies is closely monitored on an ongoing basis by AstraZeneca/MedImmune representatives in consultation with the Patient Safety Department. Issues identified will be addressed; this could involve, for instance, amendments to the clinical study protocol and letters to investigators.

6. CHANGES OF ANALYSIS FROM PROTOCOL

None.

7. REFERENCES

Berry et al 1991
Berry G, Kitchin RM, Mock PA. A comparison of 2 simple hazard ratio estimators based on the logrank test. Stat Med 1991;10(5):749-55.

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009;28:739-61.

Ciuleanu et al 2009
Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomized, double-blind, phase 3 study. Lancet 2009;374(9699):1432-40.

Collett 2003
Collett D. Modelling survival data in medical research: 2nd ed. Chapman and Hall/CRC; 2003.

Ellis et al 2008
Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008;29(4):456-65.

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer: 2001.

Gail and Simon 1985
Gail M, Simon R. Tests for qualitative interactions between treatment effects and patient subsets. Biometrics 1985;41(2):361–72.

Klein et al 2007
Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007;26(24):4505-19.

Oemar and Janssen 2013
Oemar M, Oppe M. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument V5.0 (October 2013). Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

Osoba et al 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16(1):139-44.

Paz-Ares et al 2013
Paz-Ares LG, de Marinis F, Dediu M, Thomas M, Pujol JL, Bidoli P, et al. PARAMOUNT: Final overall survival results of the phase III study of maintenance pemetrexed versus placebo immediately after induction treatment with pemetrexed plus cisplatin for advanced nonsquamous non-small-cell lung cancer. J Clin Oncol 2013;31(23):2895-902.

Robins 1993
Robins JM. Information Recovery and Bias Adjustment in Proportional Hazards Regression Analysis of Randomized Trials Using Surrogate Markers. Proceedings of the Biopharmaceutical Section, American Statistical Association 1993; 24-33.

Robins and Tsiatis 1991
Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics-Theory and Methods 1991; 20(8):2609-31.

Scagliotti et al 2008
Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-

naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008;26(21)3543-51.

Selke and Siegmund 1983
Selke T, Siegmund D. Sequential analysis of the proportional hazards model. Biometrika 1983;70:315-26.

Stone 2010
Stone A. The application of bespoke spending functions in group-sequential designs and the effect of delayed treatment switching in survival trials. Pharm Stat 2010;9(2):151-61.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method with the conventional approach for interval-censored data analysis. Stat Biopharm Res 2010;2(1):97-108.

Whitehead and Whitehead 1991
Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.

8. APPENDIX

None.

AstraZeneca

Statistical Analysis Plan

Study Code D419AC00001

Edition Number 5

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

A Phase III Randomized, Open-Label, Multi-Center, Global Study of MEDI4736 in Combination with Tremelimumab Therapy or MEDI4736 Monotherapy Versus Standard of Care Platinum-Based Chemotherapy in First-Line Treatment of Patients with Advanced or Metastatic Non-Small-Cell Lung Cancer (NSCLC) (MYSTIC)

TABLE OF CONTENTS PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation or special term	Explanation
AE	Adverse event
ALK	Anaplastic lymphoma kinase
APF12	Proportion of patients alive and progression free at 12 months from first dose
Baseline	Refers to the most recent assessment of any variable prior to dosing with study treatment/randomisation (as appropriate)
BICR	Blinded independent central review
BoR	Best objective response
CI	Confidence Interval
CR	Complete Response
CRF/eCRF	Case Report Form (electronic)
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CTC/CTCAE	Common Terminology Criteria for Adverse Event (National Institutes of Health, National Cancer Institute)
DoR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EDoR	Expected Duration of Response
EGFR	Epidermal growth factor receptor
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire
EORTC QLQ-LC13	European Organisation for Research and Treatment of Cancer 13-item lung cancer-specific quality of life questionnaire. Module used as a supplement to EORTC QLQ-C30
EQ-5D	EuroQoL 5-dimension utility index
EQ-5D-3L	EuroQoL 5-dimension, 3-level health state utility index
EQ-5D-5L	EuroQoL 5-dimension, 5-level health state utility index
FAS	Full analysis set
HR	Hazard ratio
HRQoL	Health-related Quality of Life

Abbreviation or special term	Explanation
IDMC	Independent data monitoring committee
IP	Investigational Product
ITT	Intention to Treat
IVRS	Interactive Voice Response System
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MRI	Magnetic Resonance Imaging
MTP	Multiple testing procedure
NCI	National Cancer Institute
NE	Not evaluable
NED	No evidence of disease
NSCLC	Non-small cell lung cancer
NTL	Non-target lesions
OAE	Other Significant Adverse Event
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive disease
PD-L1	Programmed death ligand 1
PDx	Pharmacodynamic(s)
PFS	Progression-free survival
PFS2	Time from randomisation to second progression
PR	Partial Response
PRO	Patient Reported Outcome
QoL	Quality of Life
QLQ-LC13	Quality of Life Lung Cancer Module; 13 item self administered questionnaire from the EORTC for lung cancer
QTcF	QT interval (corrected for heart rate using Fredericia's correction)
RECIST 1.1	Response Evaluation Criteria In Solid Tumours, Version 1.1
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
TL	Target lesions

Abbreviation or special term	Explanation
WHO	World Health Organization

AMENDMENT HISTORY

Date	Brief description of change
	• Clarification on the interim efficacy boundary calculation for the OS comparison of MEDI4736 + tremelimumab versus SoC in PD-L1–positive$_{1\%}$ population
	• Updates to the definition of corresponding safety analysis set in PD-L1-positive$_{25\%}$ population and PD-L1-positive$_{1\%}$ population • Updates to Table 2 on the analysis of the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations • Updates to study population and demographic data will be summarized in FSA, and FAS for patients within the positive$_{25\%}$ and positive$_{1\%}$ populations, correspondingly. • Updates to safety analyses will be performed in Safety Analysis Set, and safety analysis set for patients within the positive$_{25\%}$ and positive$_{1\%}$ populations, correspondingly. • Updates to PFS/OS/ORR additional analyses • Clarification that PFS subgroup analysis will also be performed for the comparison between MEDI4736 monotherapy versus SoC • Updates to the PFS multivariate Cox model to evaluate the treatment effect adjusting for any potential imbalances in baseline prognostic factors • Clarification that OS subgroup analysis will also be performed for the comparison between MEDI4736 monotherapy versus SoC • Updates to additional analysis using Cox proportional hazards models to determine the effect of covariates on the OS HR estimate • Updates to additional analysis using Cox proportional hazards models to determine the consistency of OS treatment effect • Updates to ORR subgroup analysis

Date	Brief description of change
	In line with the Clinical Study Protocol (CSP) Amendments v6-7: • Clarifications in Objectives section to note inclusion of additional co-primary endpoint in OS and additional populations for the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations in the primary and secondary objectives. • Study design section updated to reflect the introduction of the additional populations for the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations in the revised primary and secondary objectives. • Updates to the power calculations appropriate to these updated objectives. Inclusion of modified rule for the timing of the primary PFS analysis allowing for approximately 44 weeks from time of last patient randomized • Update to the definitions of the analysis sets due to introduction of the additional populations for the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations in the revised primary and secondary objectives • Updates to Table 2 to reflect the changes to the objectives and the introduction of the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations • Clarification in Secondary Endpoints section that ORR for investigator assessments is based on subjects with measurable disease at baseline • Clarification in Secondary Endpoints section of the details of the algorithm for determining time to second progression • Clarification in Secondary Endpoints section of the details of the algorithm for determining Best Objective Response • Clarification that PRO endpoints will be analysed using PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations as well as FAS • Clarifications in PRO endpoints section regarding specific endpoints for symptom deterioration. Clarification that PRO data is collected at assessment periods (via ePRO tool) and not collected at visits and specific details provided on how to determine 2 missed assessments. • Confirmation in Biomarker Variables section on the definition of the various PD-L1 expression subgroups. • Update of Analysis Methods section to reflect use of the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations for the revised co-primary and secondary objectives/endpoints and requisite updates to the multiple testing strategy • Update to analysis of co-primary endpoint PFS to reflect the primary and secondary endpoints based on PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations and the changes to relevant confidence intervals due to the revised multiple testing strategy

Date	Brief description of change
	• Subgroup analyses updated to include PD-L1 tumour cell expression for the additional cut-off points of 1%, 10% and 50%, plus inclusion of 25% cut-off for immune cell expression • Clarification of the analysis method for the subgroup analyses based on separate Cox models for each subgroup using only treatment and sub-group factors separately • Update to analysis of co-primary endpoint OS to reflect the primary and secondary endpoints based on PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations and the changes to relevant confidence intervals due to the revised multiple testing strategy • Update to analysis of all secondary endpoints to reflect the secondary objectives based on PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations • Inclusion of mixed models repeated measures analysis for the PRO symptoms to provide alternative analyses of these endpoints • Analyses of safety data updated to include appropriate analyses based on the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations. • Interim analysis details updated to reflect the change in primary objectives now based on the PD-L1-positive$_{25\%}$ population as primary In line with project wide developments • Update to Adverse Event analyses to confirm sub-analyses of CTC of grade 3 or higher will only summaries grades 3 and 4 (and exclude grade 5)

Date	Brief description of change
	In line with the Clinical Study Protocol (CSP) Amendments v3-v5: • Clarifications in Objectives section to note inclusion of co-primary endpoints, use of BICR for PFS and updates to set of secondary objectives. • Study design section updated to reflect increase in sample size and allowing treatment to progression for both combination and monotherapy treatment groups • Changes in Primary and Secondary variable section to reflect use of co-primary endpoint and use of BICR. • Sample size section updated to reflect increase in sample size based on changes to the treatment comparisons of interest and inclusion of 2 formal interim analysis for OS endpoint. • Change in requirements for re-treatment only in the combination arm • Update of Analysis Methods section to reflect use of co-primary endpoints, the updates to secondary endpoints and requisite updates to the multiple testing strategy • Changes to Analysis Methods to note use of Cox proportional hazards for calculation of HR and CI and addition of additional subgroup for PD-L1 status (using 10% cut-off for positive/negative split) • Change to definition of key and supportive PRO endpoints. • Interim analysis for Overall Survival added In line with project wide developments • Changes in duration of Adverse Event and laboratory reporting period to include up to 90 days for immunotherapy agents and 30 days for SoC • Removal of PID, PID2 and RDI2 and increased clarification of the derivations of RDI • Clarification of summaries required for deaths Other Changes • Edits to category definitions for elevated liver enzymes • Clarification of definition of total and actual exposure for immunotherapy and SoC treatments and calculation of dose delays

1. STUDY DETAILS

1.1 Study objectives

All objectives will be evaluated for all patients, unless otherwise indicated, and where appropriate for patients with PD-L1 (programmed death ligand 1)-positive$_{25\%}$, PD-L1 positive$_{1\%}$ and/or PD-L1–low/negative NSCLC. The assessment of progression-free survival (PFS) and overall survival (OS)) in patients with PD-L1positive$_{25\%}$ will be considered co-primary objectives. Section 3.5 provides the definitions of these PD-L1 subsets:

- Positive$_{25\%}$:- ≥ 25% PD-L1 membrane-expression in tumoral tissue
- Positive$_{1\%}$:- ≥ 1% PD-L1 membrane-expression in tumoral tissue
- Low/Negative:- < 25% PD-L1 membrane-expression in tumoral tissue
- Negative:- < 1% PD-L1 membrane-expression in tumoral tissue

1.1.1 Primary objective

Primary Objective:	**Outcome Measure:**
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS in patients with PD-L1-positive$_{25\%}$ NSCLC	PFS using Blinded Independent Central Review (BICR) assessments according to RECIST 1.1
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of OS in patients with PD-L1-positive$_{25\%}$ NSCLC	OS
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of OS in patients with PD-L1 positive$_{25\%}$ NSCLC	OS

1.1.2 Secondary objectives

Secondary Objectives:	Outcome Measures:
To further assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to SoC in terms of PFS, OS, objective response rate (ORR), duration of response (DoR), proportion of patients alive and progression free at 12 months from randomization (APF12) and time to second progression (PFS2)	OS in patients with PD-L1- positive$_{1\%}$ NSCLC and all patients PFS in patients with PD-L1-positive$_{1\%}$ NSCLC and all patients, using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1-positive$_{25\%}$NSCLC, patients with PD-L1-positive$_{1\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1-positive$_{25\%}$ NSCLC, patients with PD-L1-positive$_{1\%}$ NSCLC and all patients using local standard clinical practice[a]
To assess the efficacy of MEDI4736 monotherapy compared to SoC in terms of PFS, OS, ORR, DoR, APF12 and PFS2	PFS in patients with PD-L1-positive$_{25\%}$ NSCLC using BICR assessments according to RECIST 1.1 ORR, DoR, and APF12 in patients with PD-L1-positive$_{25\%}$NSCLC using BICR assessments according to RECIST 1.1 PFS2 in patients with PD-L1-positive$_{25\%}$ NSCLC using local standard clinical practice
To assess the efficacy of MEDI4736 + tremelimumab combination therapy compared to MEDI4736 monotherapy in terms of PFS, OS and ORR	PFS and ORR in patients with PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ NSCLC and all patients using BICR assessments according to RECIST 1.1 OS in patients with PD-L1-positive$_{25\%,}$and positive$_{1\%,}$ NSCLC and all patients
To assess disease-related symptoms and health related quality of life (HRQoL) in patients treated with MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC using the EORTC Q 30-item core quality of life questionnaire, Version 3 (LQ-C30 v3) and the 13-item lung cancer quality of life questionnaire (QLQ-LC13) module	EORTC QLQ-C30 EORTC QLQ-LC13 Changes in World Health Organization (WHO)/Eastern Cooperative Oncology Group (ECOG) performance status will also be assessed.
To assess the pharmacokinetics (PK) of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy	Concentration of MEDI4736 and tremelimumab in blood and non-compartmental PK parameters, such as peak concentration and trough (as data allow; sparse sampling)
To investigate the immunogenicity of MEDI4736 and tremelimumab	Presence of anti-drug antibodies (ADAs) for MEDI4736 and tremelimumab

[a] PFS2 will be defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death based on Investigator tumor assessments

1.1.3 Safety objective

Safety Objective:	Outcome Measures:
To assess the safety and tolerability profile of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy compared to SoC in the first-line setting for treatment of advanced or metastatic patients with NSCLC	Adverse events (AE), physical examinations, laboratory findings, and vital signs

1.1.4 Exploratory objectives

Exploratory Objectives:	Outcome Measures:
To explore irRECIST 1.1 as an assessment methodology for clinical benefit of MEDI4736 + tremelimumab compared to SoC with assessment by BICR	PFS and ORR using BICR assessment according to irRECIST 1.1
To assess AEs by patient self-reporting of specific CTCAE symptoms	Collection of approximately 20 patient-reported outcomes version of CTCAE (PRO-CTCAE) symptoms via an electronic device solution
To assess patients' overall impression of the change in their health status since the start of study treatment	Patients' Global Impression of Change (PGIC) item will be collected directly from patients via an electronic device solution
To investigate the relationship between PK exposure and clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze PK exposure and the relationship with clinical outcomes, efficacy, AEs, exploratory biomarkers, and/or safety parameters, as deemed appropriate
To describe and evaluate resource use associated with assigned treatments and underlying disease during assigned treatment	Health resource utilization measures including hospitalization, outpatient visits, or emergency department visits
To explore the impact of treatment and disease state on health state utility using the EQ-5D-5L during assigned treatment	EQ-5D-5L health state utility index will be used to derive health state utility based on patient-reported data
To investigate associations between pre-treatment peripheral myeloid-derived suppressor cells (MDSCs) measures and clinical activity	A graphical and/or a data modeling approach will be used to analyze the relationship between MDSC counts with clinical outcomes and/or with tumor lesion measurements
To investigate the relationship between biomarkers and clinical outcomes, efficacy, AEs, and/or safety parameters, if deemed appropriate	A graphical and/or a data modeling approach will be used to analyze biomarkers (eg,tumour mutational burden, IFNγ and/or PD-L1 status defined under alternative methods) and the relationship with clinical outcomes, efficacy, AEs, and/or safety parameters, as deemed appropriate

Note: Exploratory objective analyses may be reported separately from the main clinical study report.

1.2 Study design

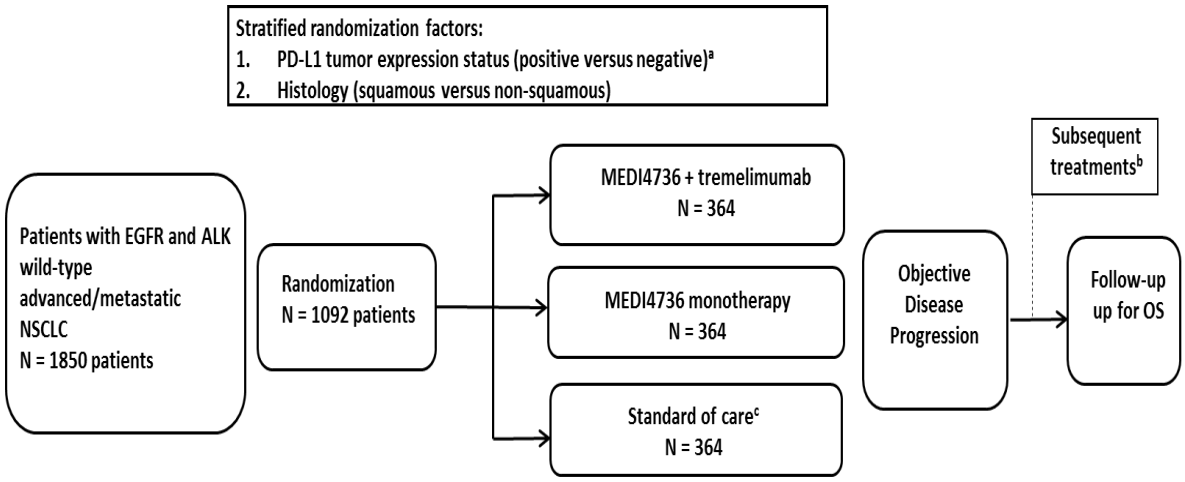
This is a randomized, open-label, multi-center, global Phase III study to determine the efficacy and safety of MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy versus platinum-based SoC chemotherapy in the first-line treatment of patients with EGFR and ALK wild-type advanced or metastatic NSCLC. A schematic diagram of the overall study design is shown in Figure 1, and a detailed study flow chart is shown in Figure 2.

This study will enroll approximately 1850 patients at sites in North America, Asia, Australia, and Europe to randomize approximately 1092 patients to treatment (including approximately 160 patients with PD-L1-positive$_{25\%}$ NSCLC in each treatment group).

Patients will provide a tumor tissue sample at enrollment (newly acquired or archived sample <3 months old) to determine PD-L1 expression status (defined by the SP263 PD-L1 immunohistochemistry assay in which ≥25% and ≥1% PD-L1–membrane expression in tumoral tissue are considered relevant positive subgroups, <25% is considered low/negative for PD-L1 expression and <1% is considered as negative for PD-L1 expression; these are referred to hereafter as patients with PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$, PD-L1-low/negative and PD-L1-negative tumors, respectively).

Patients will be randomized in a 1:1:1 ratio in a stratified manner according to PD-L1 tumor expression status (PD-L1-positive$_{25\%}$, versus PD-L1-low/negativetod) and histology (squamous versus non-squamous) to receive treatment with MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC therapy. Doses and treatment regimens are described in Section 7.2 of the CSP. Assessments will be conducted as indicated in Table 2, Table 3, and Table 4 in the CSP.

Figure 1 Overall study design



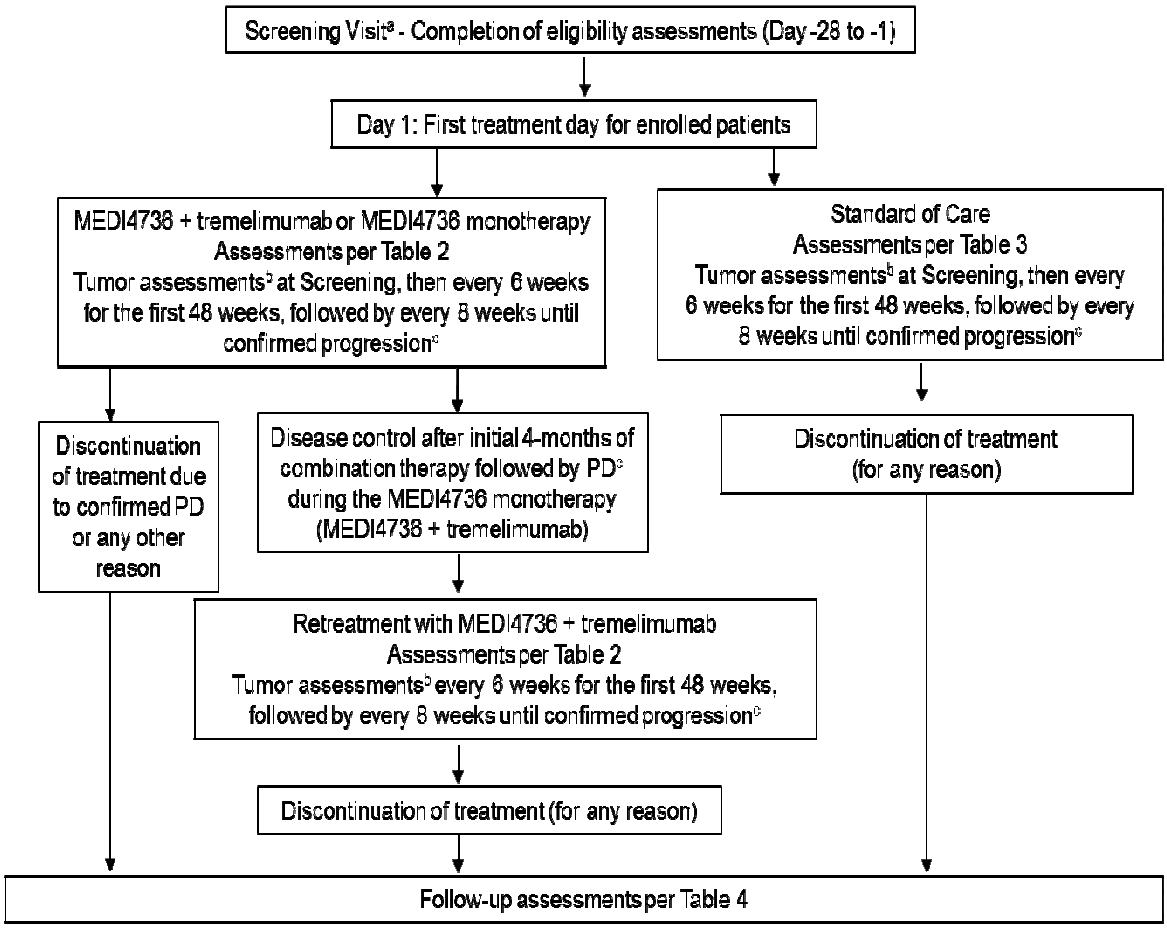
[a] Stratification by PD-L1 membrane-expression in tumoural tissue (≥25% , <25%). Sites will be supplied with PD-L1 status upon request at disease progression.

[b] Offer of standard chemotherapy per Investigator discretion.

[c] Standard of Care is an Investigator choice from the following: paclitaxel + carboplatin, gemcitabine + cisplatin (or carboplatin) (squamous only), pemetrexed + cisplatin (or carboplatin) (non-squamous only), and for eligible patients, pemetrexed maintenance (non-squamous only following pemetrexed/platinum induction).

Figure 2 Study flow chart

Tables referred to in this figure are in reference to the CSP.



[a] Informed consent of study procedures and tumor biopsy sample may be obtained prior to the 28-day screening window, if necessary, in order to permit tumor biopsy sample acquisition and analysis prior to randomization.

[b] Tumour assessments performed using RECIST 1.1.

[c] A confirmatory scan is always required following the initial demonstration of PD and preferably at the next scheduled visit (in the absence of clinically significant deterioration). (See Section 5.1 of CSP for more information.)

Independent Data Monitoring Committee (IDMC)

An IDMC comprised of independent experts will be convened and will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review safety assessments and make recommendations to continue, amend, or stop the study based on safety findings. The committee will meet at least every 6 months thereafter.

Details on the IDMC are provided in Section 5.1 and full details of the IDMC procedures and processes can be found in the IDMC Charter.

1.3 Number of patients

The study will plan to enroll approximately 1850 patients in order to randomize 1092 eligible patients 1:1:1 to MEDI4736 + tremelimumab combination therapy, MEDI4736 monotherapy, or SoC. The 1092 patients will comprise approximately 160 patients who have PD-L1-positive$_{25\%}$ tumors in each treatment group).

The study is sized to characterize the PFS and OS benefit of MEDI4736 in combination with tremelimumab versus SoC in patients with EGFR and ALK wild-type advanced or metastatic NSCLC in patients with PD-L1-positive$_{25\%}$ tumours and OS benefit of MEDI4736 monotherapy versus SoC in patients with PD-L1-positive$_{25\%}$ tumors. The sizing for PFS and OS assumes a 3-month delay in separation of the PFS and OS curves between each group, hence the use of average hazard ratios (HRs).

Two interim analyses of OS will be performed; the first at the time of the primary PFS analysis and the second when 80% of the target OS events have occurred. The alpha will be split between the 3 OS analyses using the Lan and DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, with the boundaries for the treatment comparison derived based upon the exact number of OS events at the time of analysis.

The primary PFS analysis for superiority will be performed when both of the following conditions have been met:

- Approximately 231 PFS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (72% maturity) AND

- Approximately 44 weeks follow up from last patient randomized to the

The final (primary) analysis of OS will be performed when the following conditions have been met:

- Approximately 225 OS events have occurred across the MEDI4736 + tremelimumab and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (70% maturity) AND

- Approximately 225 OS events have occurred across the MEDI4736 monotherapy and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (70% maturity)

MEDI4736 + tremelimumab versus SoC (PFS in PD-L1-positive$_{25\%}$ population)

If PFS at 12 months was 34.1% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS and assuming the true average PFS HR is 0.59, with 231 PFS events (72% maturity), the trial will have 88% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an

average HR of 0.69. With a 11-month recruitment period and a minimum follow up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited

MEDI4736 + tremelimumab versus SoC (OS in PD-L1-positive$_{25\%}$ population)

If OS at 18 months was 53.2% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9-month median OS [Ciuleanu et al 2009, Paz-Ares et al 2013, Scagliotti et al 2008]) and assuming the true average OS HR is 0.62, an estimated 225 death events (70% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum 225 deaths, the study will have 86% power to demonstrate statistical significance at the 2-sided alpha level of 1.32% (with overall alpha for OS of 1.5%), allowing for 2 interim analyses conducted at approximately 68% (at time of final PFS analysis) and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.72.

MEDI4736 monotherapy versus SoC (OS in PD-L1-positive$_{25\%}$ population)

If OS at 18 months was 53.2% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9-month median OS and assuming the true average OS HR is 0.62, an estimated 225 death events (70% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum 225 deaths, the study will have 90% power to demonstrate statistical significance at the 2-sided alpha level of 2.58% (with overall alpha for OS of 3%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.74.

MEDI4736 + tremelimumab versus SoC (OS in PD-L1–positive$_{1\%}$ population)

The overall alpha level for this comparison will be 1.5% (if only the OS comparison of combination versus SoC in PD-L1-positive$_{25\%}$ population is significant) or 3% (if only the OS comparison of monotherapy versus SoC in PD-L1- positive$_{25\%}$ population is significant) or 4.5% (if both these hypotheses are significant) (See Multiple Testing Procedure in Section 4.2.1).

With approximately 236 patients with PD-L1- positive$_{1\%}$ tumors per treatment arm (*i.e.,* 65% prevalence rate), if OS at 18 months was 51.1% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS), and assuming the true average OS HR is 0.66, an estimated 337 death events (71% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum of 337 deaths, the study will have at least 90% power to demonstrate statistical significance at the 2-sided alpha level of 1.32%, (with overall alpha for OS 1.5%,), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an at least average HR of 0.76.

MEDI4736 + tremelimumab versus SoC (OS, all patients)

Similary as the comparison of OS between MEDI4736 + tremelimumab versus SoC arm in PD-L1- positive$_{1\%}$ population, the overall alpha level for this comparison will be 1.5% (if only the OS comparisons of comination versus SoC in PD-L1- positive$_{25\%}$ population and in PD-L1- positive$_{1\%}$ population are significant) or 3% (if only the OS comparison of monotherapay versus SoC in PD-L1- positive$_{25\%}$ population and the OS comparison of combination versus SoC in PD-L1-positive$_{1\%}$ population are significant) or 4.5% (if all these hypotheses are significant) (See Multiple Testing Procedure in Section 8.5). If OS at 18 months was 49% with MEDI4736 + tremelimumab and 36% with SoC (with a 12.9 month median OS) and assuming the true average OS HR is 0.70, an estimated 528 death events (73% maturity) are expected to have occurred at 33 months from "first patient in." With a minimum of 528 deaths, the study will have at least 90% power to demonstrate statistical significance at the 2-sided alpha level of 1.32% (with overall alpha for OS 1.5%), allowing for 2 interim analyses conducted at approximately 68% and 80% of the target events, respectively. The smallest treatment difference that could be statistically significant will be an average HR of 0.81.

MEDI4736 monotherapy versus SoC (PFS in PD-L1-positive$_{25\%}$ population)

If PFS at 12 months was 34.1% with MEDI4736 monotherapy and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.59, with 231 PFS events (72% maturity), the trial will have 88% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 monotherapy versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.69. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PFS in PD-L1-positive$_{1\%}$ population)

With approximately 236 patients with PD-L1-positive$_{1\%}$ tumors per treatment arm (*i.e.,* 65% prevalence rate), if PFS at 12 months was 30.2% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.64, with 348 PFS events (74% maturity), the trial will have 90% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 + tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.74. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

MEDI4736 + tremelimumab versus SoC (PFS, all patients)

If PFS at 12 months was 26% with MEDI4736 + tremelimumab and 13% with SoC (with 5.8-month median PFS) and assuming the true average PFS hazard ratio (HR) is 0.71, with 551 PFS events (76% maturity),, the trial will have 89% power to demonstrate statistical significance at the 0.5% level (using a 2-sided test) for the comparison of MEDI4736 +

tremelimumab versus SoC, with the smallest treatment difference that could be statistically significant being an average HR of 0.79. With a 11-month recruitment period and a minimum follow-up period of 11 months assumed, it is anticipated that this analysis will be performed 22 months after the first patient has been recruited.

Table 1 provides a summary of the statistical assumptions.

Table 1 Summary of statistical assumptions

Endpoint	Analysis population	Events (number of patients)	Alpha (%)	Power (%)	Average hazard ratio to detect
PFS MEDI4736+tremelimumab versus SoC	PD-L1-positive$_{25\%}$	231 (320)	0.5	88	0.59
OS MEDI4736+tremelimumab versus SoC	PD-L1-positive$_{25\%}$	225 (320)	1.5[a]	86	0.62
OS MEDI4736 monotherapy versus SoC	PD-L1-positive$_{25\%}$	225 (320)	3 [a]	90	0.62

PFS Progression free survival; OS Overall Survival; PD-L1 Programmed death ligand 1.
[a] This is the overall alpha level (accounting for the 2 interim analyses)

2. ANALYSIS SETS

2.1 Definition of analysis sets

Full analysis set (Intention to treat (ITT))

The full analysis set (FAS) will include all randomized patients. The full analysis set will be used for all efficacy analyses (including PROs). Treatment groups will be compared on the basis of randomized study treatment, regardless of the treatment actually received. Patients who were randomized but did not subsequently go on to receive study treatment are included in the analysis in the treatment group to which they were randomized.

PD-L1-positive$_{25\%}$ analysis set

The PD-L1-positive$_{25\%}$ analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive$_{25\%}$ as defined by the SP236 PD-L1 IHC assay (ie, ≥25% PD-L1–membrane expression in tumoral tissue).

PD-L1- positive$_{1\%}$ analysis set

The PD-L1- positive$_{1\%}$ analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 positive$_{1\%}$as defined by the SP236 PD-L1 IHC assay (ie, ≥1% PD-L1–membrane expression in tumoral tissue).

PD-L1-low/negative analysis set

The PD-L1-low/negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 low/negative as defined by the SP236 PD-L1 IHC assay (ie, <25% PD-L1–membrane expression in tumoral tissue).

PD-L1-negative analysis set

The PD-L1-negative analysis set will include the subset of patients in the FAS whose PD-L1 status is PD-L1 negative as defined by the SP236 PD-L1 IHC assay (ie, <1% PD-L1–membrane expression in tumoral tissue).

Safety analysis set

The safety analysis set (SAS) will consist of all patients who received at least 1 dose of study treatment. Safety data will not be formally analyzed but summarized using the SAS, according to the treatment received, that is, erroneously treated patients (e.g., those randomized to treatment A but actually given treatment B) will be summarized according to the treatment they actually received.

The corresponding PD-L1-positive$_{25\%}$ safety analysis set will include the subset of patients in the SAS whose PD-L1 status is PD-L1 positive$_{25\%}$ as defined by the SP236 PD-L1 IHC assay (ie, ≥25% PD-L1–membrane expression in tumoral tissue).

The corresponding PD-L1-positive$_{1\%}$ safety analysis set will include the subset of patients in the SAS whose PD-L1 status is PD-L1 positive$_{1\%}$ as defined by the SP236 PD-L1 IHC assay (ie, ≥1% PD-L1–membrane expression in tumoral tissue).

Pharmacokinetic analysis set

All patients who received at least 1 dose of investigational product (IP) per the protocol for whom any post-dose data are available and who do not violate or deviate from the protocol in ways that would significantly affect the pharmacokinetic (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. Definitions of the analysis sets for each outcome variable are provided in Table 2.

Table 2 Summary of outcome variables and analysis populations

Outcome variable	Population
Efficacy data*	
PFS and OS*	Full analysis set (ITT population)
ORR, DoR, APF12, PFS2, PROs, and symptom endpoints*	Full analysis set (ITT population)
Demography*	Full analysis set (ITT population)
PK data	PK analysis Set
Safety Data*	
Exposure*	Safety analysis Set
AEs*	Safety analysis Set
Laboratory measurements*	Safety analysis Set
Vital signs*	Safety analysis Set
ECGs*	Safety analysis Set

Note: all the outcome variables with * will be repeated in the corresponding PD-L1-positive$_{25\%}$ analysis set and PD-L1-positive$_{1\%}$ analysis set.

2.2 Violations and deviations

The important protocol deviations will be listed and summarised by randomised treatment group. Deviation 1, below, will lead to exclusion from the Safety analysis set. Deviation 4, below, will lead to exclusion from the FAS analysis set. None of the other deviations will lead to patients being excluded from the analysis sets described in Section 2.1 (with the exception of the PK analysis set, if the deviation is considered to impact upon PK). A per-protocol analysis excluding patients with significant protocol deviations is not planned; however, a 'deviation bias' sensitivity analysis will be performed excluding patients with deviations that may affect the efficacy of the trial therapy if > 10% of patients:

- did not have the intended disease or indication or
- did not receive any randomised therapy.

The need for such a sensitivity analysis will be determined following review of the protocol deviations ahead of database lock and will be documented prior to the primary analysis being conducted.

Eligibility criteria deviations are deviations from the protocol inclusion and exclusion criteria. Post-entry deviations are deviations from the protocol that occurred after the patient was assigned to the study.

The following general categories will be considered important deviations and be listed and discussed in the CSR as appropriate for the study. If a 'deviation bias' sensitivity analysis is conducted then patients with these deviations will be excluded from the sensitivity analysis:

- Deviation 1: Patients randomised but who did not receive study treatment.
- Deviation 2: Patients who deviate from key entry criteria as per the CSP. These are inclusion criteria 3,5 and exclusion criteria 3,4, 8,17.
- Deviation 3: Baseline RECIST scan > 42 days before date of randomisation.
- Deviation 4: No baseline RECIST 1.1 assessment on or before date of first dose.
- Deviation 5: Received prohibited concomitant medications (including other anti-cancer agents). Please refer to the Clinical Study Protocol (CSP) section 7.7 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.
- Deviation 6: Patients randomised who received treatment other than that to which they were randomised to.

The categorisation of these as important deviations is not automatic and will depend on duration and the perceived effect on efficacy.

In addition to the programmatic determination of the deviations above, monitoring notes or summaries will be reviewed to determine any important post entry deviations that are not identifiable via programming, and to check that those identified via programming are correctly classified. The final classification of deviations will be made at the blinded data review meeting (BDRM) prior to database lock or data freeze. Decisions made at the BDRM will be documented and approved by AstraZeneca prior to analysis.

Misrandomisations in terms of errors in treatment dispensing, in addition to incorrect stratifications, will also be summarised and listed separately to the important protocol deviations. A misrandomisation is when a patient is not randomised or treated according to the randomisation schedule. It is envisaged that there will be 2 sub categories of this:

- Patients who receive no treatment whatsoever for a period of time due to errors in dispensing of medication. Note, this is not due to tolerability issues where patients may stop taking drug.
- The patient receives a treatment pack with a different code to their randomisation code. However, the actual treatment may still match the randomised treatment. For example, a patient is given randomisation code 0001, which according to the

randomisation schedule is MEDI4736. However, at the randomisation visit they are given treatment pack 0003, but this still contains MEDI4736.

The summary will include all patients with a dispensing error but will also include information on how many of those patients received at least one dose of the wrong treatment at any time. Patients who receive the wrong treatment at any time will be included in the safety analysis set as described in Section 2.1. During the study, decisions on how to handle misrandomisations will be made on an individual basis with written instruction from the study team leader and/or statistician.

3. PRIMARY AND SECONDARY VARIABLES

3.1 Derivation of RECIST 1.1 Visit Responses

For all patients, the RECIST version 1.1 (see further details in Appendix F in the CSP) tumour response data will be used to determine each patient's visit response. It will also be used to determine if and when a patient has progressed and also their best objective response.

The baseline assessment should be performed no more than 28 days before randomization and ideally as close as possible to the start of the assigned IP. Efficacy for all patients will be assessed by objective tumor assessments every 6 weeks for the first 48 weeks (relative to the date of randomization; Table 2 in the CSP for MEDI4736 + tremelimumab or MEDI4736 monotherapy, Table 3 in the CSP for SoC, and Table 4 in the CSP for patients who have completed/discontinued randomized treatment) then every 8 weeks thereafter, until confirmed objective disease progression per RECIST 1.1 (irrespective of the reason for stopping treatment or subsequent therapy). 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.

For patients who discontinue IP (including SoC) due to toxicity in the absence of confirmed objective progression, objective tumor assessments per the scheduled assessments should be continued every 6 weeks for 48 weeks (relative to randomization) and then every 8 weeks until confirmed objective disease progression.

A confirmatory scan is required for all patients following the initial demonstration of PD. The confirmatory scan should occur no earlier than 4 weeks after the initial assessment of PD and preferably at the next scheduled visit in the absence of clinically significant deterioration. Treatment with MEDI4736 + tremelimumab, MEDI4736 monotherapy, or SoC may continue between the initial assessment of progression and confirmation of progression. Progression would be considered confirmed per RECIST 1.1 criteria available in Appendix F in the CSP using Investigator assessments.

If a patient discontinues treatment (and/or receives a subsequent anticancer therapy) prior to progression, then the patient should still continue to be followed until confirmed objective disease progression.

Categorization of objective tumor response assessment will be based on the RECIST 1.1 criteria of response: complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). Target lesion progression will be calculated in comparison to when the tumor burden was at a minimum (ie, smallest sum of diameters previously recorded on study). In the absence of progression, tumor response (CR or PR) and SD will be calculated in comparison to the baseline tumor measurements obtained before starting treatment.

Following confirmed progression, patients should continue to be followed up for survival every 1 to 2 months as outlined in the study plan (Table 4 in the CSP). Subsequent anticancer therapy information will be collected at the timepoints indicated in Table 4 in the CSP.

Patients in the MEDI4736 + tremelimumab group who receive retreatment must have a baseline tumor assessment within 28 days of restarting treatment and additional scans every 6 weeks for the first 48 weeks relative to the date of randomization, and then every 8 weeks thereafter until disease progression. All assessments in Table 2 will be followed for patients who receive retreatment.

3.1.1 Investigator RECIST 1.1-based assessments: Target lesions (TLs)

At each visit, patients will be programmatically assigned a RECIST 1.1 visit response of CR, PR, SD, or PD depending on the status of their disease compared with baseline and previous assessments. Baseline will be assessed within the 28 days prior to randomization. If a patient has had a tumor assessment that cannot be evaluated, then the patient will be assigned a visit response of not evaluable (NE; unless there is evidence of progression in which case the response will be assigned as PD).

Measurable disease is defined as having at least one measurable lesion, not previously irradiated, which is ≥ 10 mm in the longest diameter (except lymph nodes which must have short axis ≥ 15 mm) with CT or MRI and which is suitable for accurate repeated measurements.

A patient can have a maximum of 5 measurable lesions recorded at baseline with a maximum of 2 lesions per organ (representative of all lesions involved suitable for accurate repeated measurement) and these are referred to as target lesions (TLs). If more than one baseline scan is recorded then measurements from the one that is closest and prior to the date of randomisation 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 non-target lesions (NTL) at baseline. Measurements are not required for these lesions, but their status should be followed at subsequent visits.

Measurable disease (ie at least one TL) is one of the entry criteria for the study. However, if a patient with non-measurable disease is enrolled in the study, the evaluation of overall visit responses will be based on the overall NTL assessment and the absence/presence of new

lesions. If a patient does not have measurable disease at baseline then the TL visit response will be not applicable (NA).

Table 3 TL visit responses

Visit Reponses	Description
Complete Response (CR)	Disappearance of all TLs. Any pathological lymph nodes selected as TLs must have a reduction in short axis to <10mm.
Partial response (PR)	At least a 30% decrease in the sum of diameters of TLs, taking as reference the baseline sum of diameters as long as criteria for PD are not met.
Progressive disease (PD)	A ≥ 20% increase in the sum of diameters of TLs and an absolute increase of ≥ 5mm, taking as reference the smallest sum of diameters since treatment started including the baseline sum of diameters.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD
Not Evaluable (NE)	Only relevant in certain situations (i.e. if any of the target lesions were not assessed or not evaluable or had a lesion intervention at this visit; and scaling up could not be performed for lesions with interventions). Note: If the sum of diameters meets the progressive disease criteria, progressive disease overrides not evaluable as a target lesion response
Not applicable (NA)	No TLs are recorded at baseline

Rounding of TL data

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

Missing TL data

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

- A new lesion is recorded.
- A NTL visit response of PD is recorded.

- The sum of TLs is sufficiently increased to result in a 20% increase, and an absolute increase of ≥ 5mm, from nadir even assuming the non-recorded TLs have disappeared.

Note: the nadir can only be taken from assessments where all the TLs had a lesion diameter recorded.

Lymph nodes

For lymph nodes, if the size reduces to < 10mm 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 < 10mm and all other TLs are 0mm then although the sum may be >0mm 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 a value of 5mm will be entered into the database and used in TL calculations, unless the radiologist has indicated and 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 blinded to treatment assignment.

Irradiated lesions/lesion intervention

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

Any TL (including lymph nodes), which has had intervention during the study (for example, irradiation / palliative surgery / embolisation), should be handled in the following way 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 tumours:

- 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 subject would be assigned a TL response of PD.

- Step 3: If after both steps PD has not been assigned, then, if appropriate, a scaled sum of diameters will be calculated (as long as ≤ 1/3 of the TLs have missing measurements), treating the lesion with intervention as missing, 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 also have a value of 0 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 where appropriate (as per step 2 above).

Scaling (applicable only for irradiated lesions/lesion intervention)

If > 1/3 of TL measurements are treated as missing (because of intervention) then TL response will be NE, unless the sum of diameters of non-missing TL would result in PD (ie 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 treated as 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.

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

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

$$\frac{26}{26.8} \times 29.3 = 28.4cm$$

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

Lesions that split in two

If a TL splits in 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.

3.1.2 Investigator RECIST 1.1-based assessments: Non-target lesions (NTLs) and new lesions

At each visit an overall assessment of the NTL response should be recorded by the investigator. 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 4 NTL Visit Responses

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

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

Details of any new lesions 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 tumour.

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 tumour assessments where possible until objective disease progression is observed.

3.1.3 Investigator RECIST 1.1-based assessments: Overall visit response

Table 5 defines how the previously defined TL and NTL visit responses will be combined with new lesion information to give an overall visit response.

Table 5 Overall visit responses

Target lesions	Non-target lesions	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

3.1.4 Blinded Independent Central Review (BICR) of RECIST 1.1-based assessments

The BICR will be performed on all radiological scans of all patients for the evaluation of the co-primary PFS endpoint (and all secondary endpoints determined from the tumour assessments). All images will be collected centrally. The imaging scans will be reviewed by 2 independent radiologists using RECIST 1.1 and will be adjudicated, if required. For each patient, the BICR will define the overall visit response data (CR, PR, SD, PD, or NE) and the relevant scan dates for each time point (ie, for visits where response or progression is/is not identified). 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). Endpoints (of PFS) will be derived from the overall visit response date and the scan dates.

The definitions of irCR, irPR, irSD, irPD, and irNED (ie responses according to irRC), as outlined by Wolchok et al 2009, will be outlined in the BICR charter, but a brief description of the methodology is given here. In this project irRC using a RECIST base will be implemented where the target lesions will be measured unidimensionally.

In irRC the presence of new lesions will not automatically trigger a declaration of Progressive Disease, but instead the new lesions will be measured and these measurements will be addedto the sum of diameters of the target lesions. Based on the sum of these measurements and % calculations thereof, the target lesion response assessment will be derived. The overall response assessment (irCR, irPR, irSD, irPD, irNE or irNED) will be obtained at the BICR and confirmation of irPD is required.

PFS and ORR by irRECIST 1.1 criteria using BICR assessments will also be performed for exploratory purposes.

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

3.2 Outcome Variables

The analysis of the co-primary endpoint, PFS, and the analyses of the secondary endpoints, ORR, DoR, and APF12, will be based on BICR tumor assessments according to RECIST 1.1. OS will be evaluated as a co-primary endpoint from all-cause mortality. In addition, time to secondary progression (PFS2) will be defined by local clinical practice.

All RECIST assessments, 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.

3.2.1 Co-Primary endpoints

PFS and OS are the co-primary endpoints.

3.2.1.1 Progression-free survival

Progression free survival (PFS) (per RECIST 1.1 using BICR assessments) will be defined as the time from the date of randomization until the date of objective disease progression or death (by any cause in the absence of progression) regardless of whether the patient withdraws from randomized therapy or receives another anticancer therapy prior to progression (i.e., date of PFS event or censoring – date of randomisation + 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 RECIST 1.1 assessment. If the patient has no evaluable visits or does not have baseline data, they will be

censored at study day 1 unless they die within 2 visits (2x6 weeks for tumor assessments + 2x7 days for visit window) of baseline.

A second (modified) PFS definition will also be applied to the co-primary PFS comparisons as a sensitivity analysis. This definition will be as described above, but for subjects who receive subsequent anticancer therapy prior to progression (or death), they will be censored at the time of the start date of taking the subsequent anticancer treatment.

PFS will also be obtained using the algorithm described above for the RECIST site investigator tumour data.

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:

- The date of progression will be determined based on the earliest of the scan dates of the component that triggered the progression for the adjudicated reviewer selecting PD or of either reviewer where both select PD as a time point response and there is no adjudication for central review (BICR).

- For investigational assessments, the date of progression will be determined based on the earliest of the RECIST 1.1 assessment/scan dates of the component that indicates 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 target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the target lesions, and similarly for non-target lesions, only the latest scan date is recorded out of all scans performed at that assessment for the non-target lesions.

Exploratory analyses

PFS based on RECIST 1.1 BICR data modified for confirmation of progression will be performed for exploratory purposes using the algorithm described above for the RECIST 1.1 BICR assessments, but following a modification whereby any objective disease progression must be confirmed by the next scheduled scan. The confirmatory scan must be no sooner than 4 weeks after the initial suspected progression and preferably at the next scheduled visit in the absence of clinically significant deterioration. If disease progression is confirmed (or disease progression occurs and no further scans are recorded) then the date of progression will be when it was originally observed. Patients with a single disease progression and no further tumor assessment scans will be treated as PD in the analysis. In the absence of significant clinical deterioration, the investigational site is advised to continue the patient on their randomized MEDI4736 + tremelimumab or MEDI4736 monotherapy until progression has been confirmed. If progression is not confirmed, the patient should continue their randomized

MEDI4736 + tremelimumab or MEDI4736 monotherapy treatment and on-treatment assessments. Treatment through PD in the SoC group is at the Investigator's discretion; however, a confirmatory scan is required for all patients in the SoC group, even if a subsequent treatment is started.

PFS by irRECIST 1.1 criteria using BICR assessments will also be reported.

3.2.1.2 Overall survival

OS is defined as the time from the date of randomization until death due to any cause (ie, date of death or censoring – date of randomisation + 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 (SUR_DAT, recorded within the SURVIVE module of the eCRF).

Note: Survival calls will be made in the week following the date of Data Cut Off (DCO) for the analysis (these contacts should generally occur within 7 days of the data cut off), 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 CRF fields to determine the last recorded date on which the patient was known to be alive for those patients still on treatment (since the SURVIVE module is only completed for patients off treatment if a survival sweep is not performed).

3.2.2 Secondary endpoints

3.2.2.1 Objective response rate

ORR (per RECIST 1.1 using BICR assessments) is defined as the number (%) of patients with at least 1 visit response of CR or PR. If any patients do not have measurable disease at baseline then the analysis of ORR will exclude these patients, so that the denominator is a subset of the Intent-to-Treat (ITT) population who have measurable disease at baseline. 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 will also be obtained using the algorithm described above for the RECIST 1.1 site investigator tumour data. The denominator for ORR will be all randomised patients with measurable disease at baseline per the site investigator (ie, the ITT population).

An exploratory analysis of ORR by irRECIST 1.1 criteria using BICR assessments will also be reported.

3.2.2.2 Duration of response

DoR (per RECIST 1.1 using BICR assessments) will be defined as the time from the date of first documented response 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 denominator for DoR will be defined as described for ORR (see Section 3.2.2.1).

The time of the initial response will be defined as the latest of the dates contributing towards the first visit response of CR or PR.

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.

3.2.2.3 Time from randomization to second progression (PFS2)

Time from randomisation to second progression (PFS2) is defined as the time from the date of randomisation to the earliest of the progression events (subsequent to that used for the primary variable PFS and excluding any confirmation of progression scans performed for first progression) or death (ie date of PFS2 event or censoring – date of randomisation + 1). The date of the first progression will be programmatically determined from investigator assessed data (See Section 3.2.1 for details.) The date of second progression will be recorded by the Investigator in the eCRF and defined according to local standard clinical practice and may involve any of the following: objective radiological imaging, symptomatic progression or death. RECIST assessments will not be collected for assessment of PFS2. The date of the PFS2 assessment and investigator opinion of progression status (progressed or nonprogressed) at each assessment will be recorded in the electronic case report form (eCRF).

Second progression status will be reviewed in line with scheduled follow-up (Table 4 CSP) following the progression event used for the co-primary variable PFS (the first progression) and status recorded.

The analysis of PFS2 should include all randomised patients. Patients who have a first PFS event who are alive but with no second event are censored at last known date alive. Patients who died as a first PFS event have their PFS2 event also at date of death. Patients who have a first PFS event and then die subsequently will have their PFS2 event at date of death. Patients alive without any first PFS event will be censored at their last known date alive.

3.2.2.4 Proportion of patients alive and progression free at 12 months

The proportion of patients alive and progression free at 12 months (APF12) will be defined as the Kaplan-Meier estimate of PFS (per RECIST 1.1 as assessed using BICR assessments) at 12 months.

3.2.2.5 Best objective response

Best objective response (BoR) is calculated based on the overall visit responses from each RECIST assessment, described in Appendix E in the CSP. It is the best response a patient has had following randomization but prior to starting any subsequent cancer therapy up until RECIST 1.1 progression or the last evaluable assessment in the absence of RECIST 1.1 progression, as determined by BICR.

Categorization of BoR will be based on RECIST 1.1 (Appendix E in Protocol) using the following response categories: CR, PR, SD, PD, and NE.

BoR will be determined programmatically based on RECIST 1.1 using all BICR assessments up until the first progression event, the start of any subsequent cancer therapy or the last evaluable assessment in the absence of progression. For patients whose progression event is death, BoR will be calculated based upon all evaluable RECIST 1.1 assessments prior to death.

For determination of a best response of SD, the earliest of the dates contributing towards a particular overall visit assessment will be used. SD should be recorded at least 6 weeks minus 1 week, (i.e. at least 35 days (to allow for an early assessment within the assessment window), after randomisation (i.e. study day 36). For CR/PR, the initial overall visit assessment which showed a response will use the latest of the dates contributing towards a particular overall visit assessment.

The denominator will be consistent with that used in the ORR analysis.

For patients who die with no evaluable RECIST 1.1 assessments, if the death occurs ≤90 days (ie, 2*(6 weeks ±3 days)) after randomization, 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 >90 days (ie, 2*(6 weeks ±3 days)) after the date of randomization then BoR will be assigned to the NE category.

Progression events that have been censored due to them being >90 days after the last evaluable assessment will not contribute to the BoR derivation.

3.2.2.6 Change in tumour size

For supportive purposes percentage change from baseline in tumour size will be derived at each scheduled tumour assessment visit (ie, week 6, week 12 etc hereafter referred to as week X for convenience). Best percentage change from baseline in tumour size will also be derived as the biggest decrease or the smallest increase in tumour size from baseline.

This is based on RECIST 1.1 target lesion measurements taken at baseline and at the timepoint of interest. Tumour size is defined as the sum of the longest diameters of the target lesions for the BICR data based upon RECIST 1.1 assessments. Target lesions are measurable tumour lesions. Baseline for RECIST 1.1 is defined to be the last evaluable assessment prior to starting treatment. The change in target lesion tumour size at week X will be obtained for each patient by taking the difference between the sum of the target lesions at week X and the sum of the target lesions at baseline. To obtain the percentage change in target lesion tumour size at week X the change in target lesion tumour size is divided by the sum of the target lesions at baseline and multiplied by 100 (i.e. (week X - baseline) / baseline * 100). More details on target lesions and measurements can be found in Section 3.1.

Apply a window around the week X visit: Whenever tumour size data for the week X visit (Note: or visit at which progression was documented if before week X) is available then this should be used in the analysis. A windowing rule will be applied and will follow the protocol allowed visit window; therefore any RECIST scan performed within ± 7 days of the protocol scheduled visit will be used for that visit.

The above derivations will be programmed for the BICR data based upon RECIST 1.1 assessments.

3.3 Patient-reported outcome (PRO) variables

Patient reported outcome (PRO) questionnaires will be assessed using the EORTC QLQ-C30 with the QLQ-LC13 module (HRQoL and lung cancer specific symptoms), PRO-CTCAE, PGIC, and EQ-5D-5L. All items/questionnaires will be scored according to published scoring guidelines or the developer's guidelines, if published guidelines are not available. All PRO analyses will be based on the Full Analysis Set (FAS; ITT population) and the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations..

3.3.1 EORTC QLQ-C30

The EORTC QLQ-C30 consists of 30 questions that can be combined to produce 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea/vomiting), 5 individual symptom items (dyspnea, insomnia, appetite loss, constipation, and diarrhea), and a global measure of health status/QoL. The EORTC QLQ-C30 will be scored according to the EORTC QLQ-C30 scoring manual (Fayers et al 2001). An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, each of the functional scales, and the global health status/QoL scale in the EORTC QLQ-C30 according to the EORTC QLQ-C30 Scoring Manual. Higher scores on the global health status/QoL and functioning scales indicate better health status/function, but higher scores on symptom scales/items represent greater symptom severity.

Baseline will be defined as the last non-missing assessment prior to first dose for symptoms and summaries.

The change from baseline in HRQoL will be assessed using the EORTC-QLQ-C30 global QoL scale which includes 2 items from the QLQ-C30: "How would you rate your overall

health during the past week?" (Item 29) and "How would you rate your overall QoL during the past week?" (Item 30).

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥10 for scales/items from the EORTC QLQ-C30 (Osoba et al 1998). For example, a clinically meaningful improvement in physical function (as assessed by EORTC QLQ-C30) is defined as an increase in the score from baseline of ≥10, whereas a clinically meaningful deterioration is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, the change in symptoms/functioning from baseline will be categorized as improvement, no change or deterioration as shown in Table 6.

Table 6 Mean change and assessment period response in health-related quality of life

Score	Change from baseline	Assessment period response
EORTC QLQ-C30 Global health status/QoL score	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change
EORTC QLQ-C30 symptom scales/items	≥+10	Deterioration
	≤-10	Improvement
	Otherwise	No change
EORTC QLQ-C30 functional scales	≥+10	Improvement
	≤-10	Deterioration
	Otherwise	No change

For each subscale, if <50% of the subscale items are missing, then the subscale score will be divided by the number of non-missing items and multiplied by the total number of items on the subscales (Fayers et al 2001). If at least 50% of the items are missing, then that subscale will be treated as missing. Missing single items are treated as missing. The reason for any missing questionnaire will be identified and recorded. If there is evidence that the missing data are systematic, missing values will be handled to ensure that any possible bias is minimized.

For the assessment period level summaries of Improvement/Deterioration/No change then all patients with a baseline and post-baseline score will be included thus the denominator may differ from the time to deterioration and improvement rate endpoints derived below.

3.3.1.1 Time to HRQoL/function deterioration

For the following HRQoL items of the EORTC QLQ-C30, time to deterioration will be analyzed:

The Global Health Status/ QoL scale consisting of items 29 and 30 of the EORTC QLQ C30.

- Item 29: How would you rate your overall health during the past week?
- Item 30: How would you rate your overall quality of life during the past week?

Patients are asked to rate their overall health and overall quality of life on a scale from 1 (very poor) to 7 (excellent).

Time to deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful deterioration (a decrease in the function scales or the global health status/QoL from baseline of ≥10) that is confirmed at a subsequent assessment period or death (by any cause) in the absence of a clinically meaningful deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to HRQoL/function deterioration. (ie date of HRQoL/function deterioration event or censoring-date of randomisation + 1). Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 assessment periods of the last PRO assessment where the HRQoL/function change could be evaluated.

Patients whose HRQoL (as measured by EORTC QLQ-C30) has not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the HRQoL/function could be evaluated. Also, if HRQoL deteriorates after 2 or more missed PRO assessment periods (using the same definitions for two missed assessment periods as used in the 'Time to Symptom deterioration' derivation in section 3.3.2.1) or the patient dies after 2 or more missed PRO assessment periods, the patient will be censored at the time of the last PRO assessment where HRQoL/function could be evaluated. If the patient has no evaluable assessments or does not have baseline data they will be censored at 1 day unless they die within 2 assessment periods of baseline (16 weeks plus 1 week allowing for a late assessment within the assessment period window).

The population for analysis of time to HRQoL/function deterioration will include a subset of the ITT population who have baseline scores ≥ 10. In the primary analysis, RECIST 1.1 progression will not be considered as HRQoL/function deterioration and data will not be affected by RECIST 1.1 progression. However a sensitivity analyses will be performed for time to HRQoL deterioration for global health status only where RECIST 1.1 progression is considered as an event of HRQoL deterioration. If a patient has both RECIST 1.1 progression and HRQoL deterioration, the earliest date of the two will be used.

3.3.1.2 Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥10 for EORTC QLQ-C30 symptom scales/items) in that symptom from baseline. The denominator will consist of a subset of the ITT population who have a baseline symptom score >10.

3.3.1.3 HRQoL/function improvement rate

The HRQoL/function improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (an increase from baseline score ≥10 for EORTC QLQ-C30 functional scales and global health status/HRQoL) in that scale from baseline. The denominator will consist of a subset of the ITT population who have a baseline HRQoL/function score ≤90.

3.3.2 Lung cancer module (EORTC QLQ-LC13)

The QLQ-LC13 is a lung cancer specific module from the EORTC for lung cancer comprising 13 questions to assess lung cancer symptoms (cough, hemoptysis, dyspnea, and site-specific pain), treatment-related side-effects (sore mouth, dysphagia, peripheral neuropathy, and alopecia), and pain medication.

The LC-13 incorporates symptom scales including:

- Dyspnea (multi-item scale based on 3 questions: were you short of breath when you rested; walked; climbed stairs)
- Cough: 1 item (how much did you cough?)
- Haemoptysis: 1 item (did you cough up blood?)
- Pain: 3 individual items (Have you had pain in your chest; your arm or shoulder; other parts of your body?)

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC-13 is identical in principle to that for the symptom scales/single items of QLQ-C30.

Definition of clinically meaningful changes

Changes in score compared with baseline will be evaluated. A minimum clinically meaningful change is defined as an absolute change in the score from baseline of ≥10 for scales/items from the QLQ-LC13 (Osoba et al 1998). For example, a clinically meaningful deterioration or worsening in chest pain (as assessed by the QLQ-LC13) is defined as an increase in the score from baseline of ≥10, whereas a clinically meaningful improvement is defined as a decrease in the score from baseline of ≥10. At each post-baseline assessment, the

change in symptoms from baseline will be categorized as an improvement, no change or deterioration as shown in Table 7.

Table 7 Visit response for health-related quality of life (HRQoL) and disease-related symptoms

Score	Change from baseline	Visit response
QLQ-LC13 symptom scales/items	≥10	Deterioration
	≤10	Improvement
	Otherwise	No change

QLQ-LC13 Lung Cancer Module.

For the assessment period level summaries of Improvement/Deterioration/No change then all patients with a baseline and post-baseline score will be included thus the denominator may differ from the time to deterioration and improvement rate endpoints derived below.

3.3.2.1 Time to symptom deterioration

For each of the following key symptom scales/items in the QLQ-LC13 and QLQ-C30, time to deterioration will be analyzed:

- Dyspnea (multi-item scale based on 3 questions: were you short of breath when you rested; walked; climbed stairs) from LC13
- Cough: 1 item (how much did you cough?) from LC13
- Pain (3 individual items): a) Have you had pain in your chest; b) your arm or shoulder; c) other parts of your body?) from LC-13
- Hemoptysis: 1 item (did you cough up blood?) from LC-13
- Appetite Loss (Have you lacked appetite) from C-30
- Fatigue (Have you felt weak, Did you need to rest, Were you tired) from C-30

The dyspnea scale is only used if all 3 items have been scored; otherwise the items are treated as single-item measures. The scoring approach for the QLQ-LC13 is identical in principle to that for the symptom scales/single items of the EORTC QLQ-C30.

Time to symptom deterioration will be defined as the time from the date of randomization until the date of the first clinically meaningful symptom deterioration (an increase in the score from baseline of ≥10) that is confirmed at a subsequent assessment or death (by any cause) in the absence of a clinically meaningful symptom deterioration, regardless of whether the patient withdraws from study treatment or receives another anticancer therapy prior to

symptom deterioration (ie date of symptom deterioration event or censoring – date of randomisation + 1). Patients with a single deterioration and no further assessments will be treated as deteriorated in the analysis.

Death will be included as an event only if the death occurs within 2 assessment periods of the last PRO assessment where the symptom change could be evaluated.

Patients whose symptoms (as measured by QLQ-LC13 or QLQ-C30) have not shown a clinically meaningful deterioration and who are alive at the time of the analysis will be censored at the time of their last PRO assessment where the symptom could be evaluated. Also, if symptoms deteriorate after 2 or more missed PRO assessmentsor the patient dies after 2 or more missed PRO assessments, the patient will be censored at the time of the last PRO assessment where the symptom could be evaluated. If the patient has no evaluable assessmentsor does not have baseline data they will be censored at 1 day unless they die within 2 assessments of baseline (126 days (i.e. 16 weeks x7 days) plus 2x7 days allowing for a late assessment within the assessment window, for a QLQ-C30 assessment, and 70 days (i.e. 8 weeks x7 days) plus 2x7 days allowing for a late assessment within the assessment window, for a QLQ-LC13 assessment).

Given the scheduled PRO assessment scheme (ie every 4 weeks for first 8 weeks and then every 8 weeks thereafter for QLQ-C30 and every 2 weeks for first 8 weeks and then every 4 weeks thereafter for QLQ-LC13) the definition of 2 missed assessments will change. (1) For QLQ-C30, if the previous PRO assessment is less than study day 70 (ie week 10) then two missing assessments will equate to 10 weeks since the previous PRO assessment, allowing for early and late assessments. If the two missed assessments occur over the period when the scheduled frequency of PRO assessments changes from 4-weekly to 8-weekly this will equate to 14 weeks. The time period for the previous PRO assessment will be from study days 70 to 119 (ie week 10 to week 17). From week 17 onwards, two missing assessments will equate to 18 weeks. If the patient has no evaluable assessments or does not have baseline data they will be censored at day 1 unless they die within 2 assessments of baseline (8 weeks plus 1 week allowing for a late assessment within the assessment window). (2) for QLQ-LC13 If the previous PRO assessment is less than study day 42 (ie week 6) then two missing assessments will equate to 6 weeks since the previous PRO assessment, allowing for early and late assessments. If the two missed assessments occur over the period when the scheduled frequency of PRO assessments changes from 2-weekly to 4-weekly this will equate to 8 weeks. The time period for the previous PRO assessment will be from study days 70 to 91 (ie week 10 to week 13). From week 13 onwards, two missing assessments will equate to 10 weeks. If the patient has no evaluable assessments or does not have baseline data they will be censored at day 1 unless they die within 2 assessments of baseline (4 weeks plus 1 week allowing for a late assessment within the assessment window for QLQ-LC13 and 8 weeks plus 1 week allowing for a late assessment within the assessment window for QLQ-C30).

The population for analysis of time to symptom deterioration will include a subset of the ITT population who have baseline scores ≤ 90.

In the primary analysis, RECIST 1.1 progression will not be considered as symptom deterioration and data will not be affected by RECIST 1.1 progression. However a sensitivity analyses will be performed for QLQ-LC13 time to symptom deterioration for each of chest pain, arm/shoulder pain, and other pain where RECIST 1.1 progression is considered as an event of symptom deterioration.

If a patient has both RECIST 1.1 progression and symptom deterioration, the earliest date of the two will be used.

3.3.2.2 Symptom improvement rate

The symptom improvement rate will be defined as the number (%) of patients with 2 consecutive assessments at least 14 days apart that show a clinically meaningful improvement (a decrease from baseline score ≥10 for QLQ-LC13 symptom scales/items) in that symptom from baseline. The denominator will consist of a subset of the ITT population who have a baseline symptom score >10.

3.3.3 Healthy state utility (EQ-5D-5L)

The EQ-5D is a standardised measure of health status developed by the EuroQol Group in order to provide a simple, generic measure of health for clinical and economic appraisal. Applicable to a wide range of health conditions and treatments, it provides a simple descriptive profile and a single index value for health status that can be used in the clinical and economic evaluation of health care.

The EQ-5D-5L index comprises 5 dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression). For each dimension, respondents select which statement best describes their health on that day from a possible 5 options of increasing levels of severity (no problems, slight problems, moderate problems, severe problems, and unable to/extreme problems). A unique EQ-5D health state is referred to by a 5-digit code allowing for a total of 3125 health states. For example, state 11111 indicates no problems on any of the 5 dimensions. These data will be converted into a weighted health state index by applying scores from EQ-5D value sets elicited from general population samples (the base case will be the United Kingdom valuation set, with other country value sets applied in scenario analyses). Where values sets are not available, the EQ-5D-5L to EQ-5D-3L crosswalk will be applied (Oemar and Jansen 2013). In addition to the descriptive system, respondents also assess their health on the day of assessment on a visual analogue scale, ranging from 0 (worst imaginable health) to 100 (best imaginable health). This score is reported separately.

The evaluable population will comprise a subset of the ITT analysis set who have a baseline EQ-5D-5L assessment.

3.3.4 PRO Compliance Rates

Summary measures of overall compliance and compliance over time will be derived for the EORTC-QLQ-C30, LC13 and EQ-5D-5L respectively. These will be based upon:

- Received questionnaire = a questionnaire that has been received and has a completion date and at least one individual item completed.

- Expected questionnaire = a questionnaire that is expected to be completed at a scheduled assessment time e.g. a questionnaire from a patient who has not withdrawn from the study at the scheduled assessment time but excluding patients in countries with no available translation. For patients that have progressed, the latest of progression and safety follow-up will be used to assess whether the patient is still under HRQoL follow-up at the specified assessment time. Date of study discontinuation will be mapped to the nearest visit date to define the number of expected forms.

- Evaluable questionnaire = a questionnaire with a completion date and at least one subscale that is non-missing.

- Overall PRO compliance rate is defined as: Total number of evaluable questionnaires across all time points, divided by total number of questionnaires expected to be received across all time points multiplied by 100.

- Overall patient compliance rate is defined for each randomised treatment group as: Total number of patients with an evaluable baseline and at least one evaluable follow-up questionnaire (as defined above), divided by the total number of patients expected to have completed at least a baseline questionnaire multiplied by 100.

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

3.4 Safety

Safety and tolerability will be assessed in terms of adverse events (AEs) (including serious adverse events [SAEs]), deaths, laboratory data, vital signs, electrocardiograms (ECGs) and exposure. These will be collected for all patients. Safety data will be summarised from the treatment period for the immunotherapy agents alongside the SOC agents. Safety data from the re-treatment period for the MEDI4736 + tremelimumab group may also be summarised via a small set of headline summaries should there be sufficient number of patients re-treated to warrant this. Any safety summaries representing the re-treatment period will be based upon a subset of the safety analysis set representing patients who have had at least one dose of study treatment in the re-treatment period.

Data from the treatment period on the immunotherapy agents (MEDI4736 or MEDI4736+tremelimumab) will be compared against SoC in the main presentations of safety data and safety data from the retreatment period for the MEDI4736 + tremelimumab group

may also be summarised separately (see Section 4.1). 'On treatment' will be defined as assessments between date of start dose and 90 days following last dose of the immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) on each period of treatment and between date of start dose and 30 days following last dose of the Standard of Care agents. Note that for one version of the safety outputs the period of time after the administration of subsequent therapy will not be considered 'on treatment' (see further Section 4.2.10).

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

3.4.1 Adverse events (AEs)

AEs and SAEs will be collected throughout the study, from date of first dose and 90 days after the last dose of immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) and between date of first dose and 30 days following last dose of the Standard of Care agents. The Medical Dictionary for Regulatory Activities (MedDRA) (using the latest or current MedDRA version) will be used to code the AEs. AEs will be graded according to the National Cancer Institute Common Terminology Criteria for AEs (CTCAE Version 4.03). A treatment emergent adverse event (TEAE) is an AE with an onset date or a pre-existing AE worsening following the first dose of study treatment through to 90 days after the last dose of immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) or 30 days after the last dose of the Standard of Care agents. For the MEDI4736+tremelimumab group, in the unlikely event of the components being administered separately then date of first dose/last dose will be considered as the earliest/latest dosing date of either component.

Other significant adverse events

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

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

AEs of special interest

An adverse event of special interest (AESI) is one of scientific and medical interest specific to understanding of the Investigational Product and may require close monitoring and rapid communication by the investigator to the sponsor. An AESI may be serious or non-serious.

The rapid reporting of AESIs allows ongoing surveillance of these events in order to characterize and understand them in association with the use of this investigational product.

AESIs for MEDI4736 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 AESIs are being closely monitored in clinical studies with MEDI4736 monotherapy and combination therapy. An immune-related adverse event (irAE) is defined as an adverse event that is associated with drug exposure and is consistent with an immune-mediated mechanism of action and where there is no clear alternate aetiology. Serologic, immunologic, and histologic (biopsy) data, as appropriate, should be used to support an irAE diagnosis. Appropriate efforts should be made to rule out neoplastic, infectious, metabolic, toxin, or other etiologic causes of the irAE.

Some clinical concepts (including some selected individual preferred terms and higher level terms) have been considered "AEs of special interest" (AESI) to the MEDI4736 program. These AESIs have been identified as Pneumonitis, Colitis, Hepatitis, Hypothyroidism, Hyperthyroidism, Hypnophysitis, Adrenal Insufficiency, Dermatitis, Nephritis /Acute Renal Failure, Pancreatitis, Neuropathy, Infusion-related Reactions and Hypersitivity / Anaphylactic Reactions. Other categories may be added or existing terms may be merged as necessary. An AstraZeneca medically qualified expert after consultation with the Global Patient Safety Physician has reviewed the AEs of interest and identified which preferred terms contribute to each AESI. A further review will take place prior to Database lock (DBL) to ensure any further terms not already included are captured within the categories.

3.4.2 Treatment exposure

Exposure will be defined separately for the initial treatment period and for the re-treatment period for the MEDI4736 + tremelimumab group as follows:.

Total (or intended) exposure of MEDI4736 (combination)

- Total (or intended) exposure is defined as the total treatment period from first dose date of study drug to the earliest of "last dose date of study drug + 27 days" or death date or DCO or start of re-treatment (applies to initial treatment period only).

Total (or intended) exposure of tremelimumab (combination)

- Total (or intended) exposure is defined as the total treatment period from first dose date of study drug to the earliest of "last dose date of study drug + 27 days" or death date or DCO or start of re-treatment (applies to initial treatment period only).

Actual exposure of MEDI4736/tremelimumab

- Actual exposure is defined as above, but excluding total duration of dose delays

The total (or intended) exposure for each SoC treatment will be calculated using the same principle as above, according to the dose schedule required for each SoC. The total (or intended) exposure will also be summarised by combining the SoC treatments together. Actual exposure will not be calculated for SoC.

The total (or intended) exposure for each SoC is defined as follows:

Total (or intended) exposure of Paclitaxel / Carboplatin / Cisplatin

- Total (or intended) exposure is defined as the total treatment period from first dose date of study drug to the earliest of "last dose date of study drug + 20 days" or death date or DCO.

Total (or intended) exposure of Gemcitabine

- Total (or intended) exposure is defined as the total treatment period from first dose date of study drug to the earliest of "last dose date of study drug + 6 days (if last dose is day 1 of cycle) or + 13 days (if last dose is day 8 of cycle)" or death date or DCO.

Dose reductions are not permitted per Section 6.7 of the CSP for the immunotherapy agents (MEDI4736 or MEDI4736+tremelimumab). The actual exposure calculation makes no adjustment for any dose reductions that may have occurred.

Exposure will also be measured by the number of cycles received. For all five choices of SoC regimen, a cycle corresponds to a period of 21 days, but for each immunotherapy agent a cycle corresponds to one dose of treatment. If a cycle is prolonged due to toxicity, this should still be counted as one cycle. A cycle will be counted if treatment is started even if the full dose is not delivered.

Calculation of duration of dose delays (for actual exposure), MEDI4736 and tremelimumab:

Duration of dose delays = Sum of (Date of the dose - Date of previous dose – 28 days)

Patients who permanently discontinue during a dose delay

If a decision is made to permanently discontinue study treatment in-between cycles or during a dose delay then the date of last administration of study medication recorded will be used in the programming.

3.4.3 Dose intensity

Dose intensity will be derived separately for the initial treatment period and the re-treatment period for the MEDI4736+tremelimumab group. It will also be derived for the SOC agents. Relative dose intensity (RDI) is the percentage of the actual dose intensity delivered relative to the intended dose intensity through to treatment discontinuation.

Relative dose intensity (RDI) will be defined as follows for MEDI4736, tremelimumab and all Standard of Care therapy:

- RDI = 100% * d/D, where d is the actual cumulative dose delivered up to the actual last day of dosing and D is the intended cumulative dose up to the actual last day of dosing. D is the total dose that would be delivered, if there were no modification to dose or schedule. When accounting for the calculation of intended cumulative dose 3 days should be added to the date of last dose to reflect the protocol allowed window for dosing.

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

3.4.4 Laboratory data

Laboratory data will be collected throughout the study, from screening to the follow-up visits as described in the CSP. Blood and urine samples for determination of haematology, clinical chemistry, and urinalysis will be collected as described in Section 5.2.1 of the CSP. For the definition of baseline and the derivation of post baseline visit values considering visit window and how to handle multiple records, derivation rules as described in Section 3.4.8 below will be used.

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

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

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

Absolute values will be compared to the project reference range 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-dose value at any time.

Local reference ranges will be used for the primary interpretation of laboratory data at the local laboratory. Project reference ranges will be used throughout for reporting purposes. The denominator used in laboratory summaries of CTC 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 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 CTCAE criterion does not consider changes from baseline, to be evaluable the patient need only have 1 post dose-value recorded.

3.4.5 Time to first subsequent therapy from discontinuation of study treatment

Time to subsequent therapy from date of last dose is defined as the time from the date of discontinuation of study treatment to the start date of the first subsequent therapy after discontinuation of treatment. Any patient not known to have had a first subsequent therapy will not have this calculation performed.

3.4.6 ECGs

ECG data obtained up until the 30 days from date of last dose of study treatment will be used for reporting. For derivation of post baseline visit values considering visit window and to handle multiple records present in any visit window, derivation rules as described in Section 3.4.7 below will be used.

At each time point the Investigator's assessment of the ECG will be collected locally. Heart rate, duration of QRS complex, RR, PR and QT intervals will be collected centrally via a digital read. This digital copy of all ECGs will be held centrally by a central ECG provider, and the data from this review will be stored for analysis if necessary at the end of the study. If it is necessary to analyse this data then QTcF (Fridericia) will be calculated programmatically using the reported ECG values (RR and QT).

QTcF = QT/RR^ (1/3) where RR is in seconds

For triplicate ECGs, the mean of the three ECG assessments will be used to determine the value at that time point.

3.4.7 Vital signs

Vital signs data obtained up until the 30 days from date of last dose of study treatment will be used for reporting. Change from baseline in vital signs variables will be calculated for each post-dose visit on treatment. For derivation of post baseline visit values considering visit window and to handle multiple records, derivation rules as described in Section 3.4.8 below will be used.

The denominator in vital signs data should include only those patients with recorded data.

3.4.8 General considerations for safety assessments

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

- The time windows should be exhaustive so that data recorded at any time point has the potential to be summarised. 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 half way 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 for MEDI4736 + tremelimumab combination therapy and MEDI4736 monotherapy (with 4 weeks between scheduled assessments) are:

  Day 29, visit window 2 – 42

  Day 57, visit window 43 – 70

  Day 85, visit window 71 – 98

  Day 113, visit window 99 – 126

  Day 141, visit window 127 – 154

  Day 169, visit window 155 – 182

  Day 197, visit window 183 – 210

  Day 225, visit window 211 – 238

  Day 253, visit window 239 – 266

  Day 281, visit window 267 – 294

  Day 309, visit window 295 – 322

  Day 337, visit window 323 – 350

  Note: Due to the differing assessment schedules the visit windows will be different for the different study treatments and endpoints.

- 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 summarised, 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 post baseline values collected are used including those collected at unscheduled visits regardless of whether or not the value is closest to the scheduled visit date.
  - To prevent very large tables or plots being produced that contain many cells with meaningless data, for each treatment group visit data should only be summarised if the number of observations is greater than the minimum of 20 and > 1/3 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.
- Baseline will be defined as the last non-missing measurement prior to dosing with study treatment. For the re-treatment period for the MEDI4736 + tremelimumab group then baseline is similarly defined as the last non-missing measurement prior to the first dose on the re-treatment period. For laboratory data, any assessments made on day 1 will be considered pre-dose. Alternatively, 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 with no washout or other intervention in the screening period), the average can be taken as a baseline value. For non-numeric laboratory tests (ie 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 day 1, one with time recorded and the other without time recorded, the one with time recorded would be selected as baseline. Where safety data are summarised over time, study day will be calculated in relation to date of first treatment.

Missing safety data will generally not be imputed. However, safety assessment values of the form of "< x" (i.e., below the lower limit of quantification) or > x (i.e., above the upper limit of quantification) will be imputed as "x" in the calculation of summary statistics but displayed as "< x" or "> x" in the listings. Additionally, adverse events that have missing causality (after data querying) will be assumed to be related to study drug.

3.5 Biomarker Variables

PD-L1 expression status (positive$_{25\%}$, and positive$_{1\%,}$ low-negative and negative) is defined according to following criteria:

- Positive$_{25\%}$:- ≥ 25% PD-L1 membrane-expression in tumoral tissue .
- Positive$_{1\%}$:- ≥ 1% PD-L1 membrane-expression in tumoral tissue .
- Low/Negative:- < 25% PD-L1 membrane-expression in tumoral tissue.
- Negative:- < 1% PD-L1 membrane-expression in tumoral tissue

3.6 Pharmacokinetic and Immunogenicity variables

Pharmacokinetic analyses will be performed on the full PK analysis set and also on the subset of patients in the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ sub-populations as deemed necessary to support the submission.

3.6.1 Population pharmacokinetics and exposure-response/safety analysis

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. The results of such an analysis will be reported separately from the main CSR. The PK, pharmacodynamic (PDx), 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-PDx methods.

3.6.2 Pharmacokinetic non-compartmental analysis

. The actual sampling times will be used in the PK calculations. PK concentration data and summary statistics will be tabulated. Individual and mean blood concentration-time profiles will be generated. PK parameters will be determined using standard non-compartmental methods. The following PK parameters will be determined after the first and steady-state doses: peak and trough concentration (as data allow). Samples below the lower limit of quantification will be treated as missing in the analyses.

3.6.3 Immunogenicity analysis

Immunogenicity results will be analyzed descriptively by summarizing the number and percentage of patients who develop detectable ADAs against MEDI4736 and tremelimumab. The immunogenicity titer and presence of neutralizing ADAs will be reported for samples confirmed positive for the presence of ADAs. Summaries will be based upon all patients from the safety population. The effect of immunogenicity on PK, pharmacodynamics, efficacy, and safety will be evaluated, if the data allow, but will be reported in a separate report.

Additional summaries of immunogenicity will be produced based on patients in the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ as deemed necessary to support the submission.

3.7 Health Resource Use

Health resource use outcome variables include the following:

- Length of hospital stay
- Reasons for hospitalisation
- Length of any time spent in an intensive care unit (ICU)

The length of hospital stay will be calculated as the difference between the date of hospital discharge (or death date) and the start date of hospitalisation or start of study drug if the start of study drug is after start date of hospitalisation (length of hospital stay = end date of hospitalisation – start date of hospitalisation + 1). Patients with missing discharge dates will be calculated as the difference between the last day with available data and the start date of hospitalisation.

Sum of total duration of hospital stay will be considered for analysis if >20% of patients who were hospitalised were admitted to hospital more than one time during study period.

The length of ICU stay will be calculated in the using the same method as detailed above for the length of hospital stay.

4. ANALYSIS METHODS

4.1 General principles

The formal statistical analysis will be performed to test the main hypotheses:

- H0: No difference between MEDI4736 + tremelimumab and SoC
- H1: Difference between MEDI4736 + tremelimumab and SoC

The co-primary endpoints are PFS and OS in patients with PD-L1-positive$_{25\%}$ tumours (with PFS using BICR assessments per RECIST 1.1). The study has been sized to characterize the PFS and OS benefits of MEDI4736 + tremelimumab versus SoC.

The primary PFS analysis and the first interim OS analysis will be performed when both of the following conditions have been met:

- Approximately 231 PFS events have occurred across the MEDI4736+tremelimumab and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (73% maturity) AND
- Approximately 44 weeks follow up from last patient randomized to the

The final (primary) analysis of OS will be performed when the following conditions have been met:

- Approximately 225 OS events have occurred across the MEDI4736+tremelimumab and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (70% maturity) AND

- Approximately 225 OS events have occurred across the MEDI4736 monotherapy and SoC treatment groups in patients with PD-L1-positive$_{25\%}$ tumours (70% maturity)

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

- Descriptive statistics will be used for all variables, as appropriate, and will be presented by treatment group. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, minimum, and maximum. For log transformed data it is more appropriate to present geometric mean, coefficient of variation (CV), median, minimum and maximum. Categorical variables will be summarized by frequency counts and percentages for each category.

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

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

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

- SAS® version 9.2 will be used for all analyses.

Baseline will be the last assessment of the variable under consideration prior to the intake of the first dose of IP, except for efficacy variables. For efficacy variables, baseline is defined as the last visit prior to randomization.

All data collected will be listed. Study population, demography data, efficacy and PRO data will be summarized and analyzed based on the FAS and the PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations. PK data will be summarized and analyzed based on the PK Analysis Set. Safety and treatment exposure data will be summarized on the Safety Analysis Set and, for the subsets of the Safety Analysis Set for patients within the positive$_{25\%}$ and positive$_{1\%}$ populations.

All outputs will be summarized by treatment group for all randomized patients (ITT),all randomized patients within the positive$_{25\%}$ and positive$_{1\%}$ populations, and Safety Analysis Set, SAS for patients within the positive$_{25\%}$ and positive$_{1\%}$ populations, correspondingly.

Safety data will be summarised from the treatment period for the immunotherapy agents alongside the SOC agents. Safety data from the re-treatment period for the MEDI4736 + tremelimumab group may also be summarised via a small set of headline summaries should there be sufficient number of patients re-treated to warrant this. Any safety summaries representing the re-treatment period will be based upon a subset of the safety analysis set representing patients who have had at least one dose of study treatment in the re-treatment period.

4.2 Analysis methods

Results of all statistical analysis will be presented using a appropriately sized confidence intervals (CI) and 2-sided p-value, unless otherwise stated.

The following table (Table 8) details which endpoints are to be subjected to formal statistical analysis, together with pre-planned sensitivity analyses, making it clear which analysis is regarded as primary for that endpoint.

Table 8 Pre-planned statistical and sensitivity analyses to be conducted

Endpoints analyzed	Notes

Progression-free survival	Stratified log-rank tests for:
	• Co-Primary analysis using BICR RECIST 1.1 assessments:
	  – MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{25\%}$ population (stratified only for histology)
	• Secondary analysis using BICR RECIST 1.1 assessments:
	  – MEDI4736 + tremelimumab versus SoC (ITT population)
	  – MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{1\%}$ population
	  – MEDI4736 monotherapy versus SoC for PD-L1-positive$_{25\%}$ population (stratified only for histology)
	  – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population)
	  – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$ population (stratified only for histology)
	  – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{1\%}$ population
	• Sensitivity analyses using Investigator assessments (RECIST 1.1)
	• Sensitivity analysis based on RECIST 1.1 modified for confirmation of progression using BICR assessments
	• Additional analysis using Cox proportional hazards models to determine the effect of covariates on the HR estimate
	• Additional analysis using Cox proportional hazards models to determine the consistency of treatment effect between stratification subgroups via the approach of Gail and Simon 1985.
	• Subgroup analysis using Cox proportional hazards model
	• Exploratory analysis using BICR data for irRECIST 1.1

Endpoints analyzed	Notes
Overall survival	Stratified log-rank test for: • Co-Primary analysis – MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{25\%}$ population (stratified only for histology) – MEDI4736 monotherapy versus SoC for PD-L1-positive$_{25\%}$ population (stratified only for histology) • Secondary analysis: – MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{1\%}$ population – MEDI4736 + tremelimumab versus SoC (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$ population (stratified only for histology) – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{1\%}$ population • Additional analysis using Cox proportional hazards models to determine the effect of covariates on the HR estimate • Additional analysis using Cox proportional hazards models to determine the consistency of treatment effect between stratification subgroups via the approach of Gail and Simon 1985 • Subgroup analysis using Cox proportional hazards model
Objective response rate	Logistic regression for: • Secondary analysis using BICR RECIST 1.1 assessments – MEDI4736 + tremelimumab versus SoC for the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations – MEDI4736 monotherapy versus SoC for the PD-L1-positive$_{25\%}$ population – MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations • Sensitivity analysis using Investigator RECIST 1.1 assessments • Subgroup analysis using logistic regression model • Exploratory analysis using BICR data for irRECIST 1.1
Duration of response	Analysis methods as described by Ellis et al 2008 for: • Secondary analysis using BICR assessments (RECIST 1.1) – MEDI4736 + tremelimumab versus SoC for the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations – MEDI4736 monotherapy versus SoC for the PD-L1-positive$_{25\%}$ population

Endpoints analyzed	Notes
Proportion of patients alive and progression free at 12 months	Hazard ratio using the Kaplan Meier estimates of progression free survival at 12 months (following method described by Klein et al 2007) • Secondary analysis: – MEDI4736 + tremelimumab versus SoC for the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations – MEDI4736 monotherapy versus SoC for the PD-L1-positive$_{25\%}$ population
Time from randomization to second progression	Stratified log-rank test • Secondary analysis: – MEDI4736 + tremelimumab versus SoC for the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations – MEDI4736 monotherapy versus SoC for the PD-L1-positive$_{25\%}$ population
Time to deterioration (EORTC QLQ-C30 and QLQ-LC13 endpoints)	Stratified log-rank test

4.2.1 Multiple testing strategy

In order to strongly control the type I error at 5% 2-sided, a multiple testing procedure (MTP) with gatekeeping strategy will be used across the co-primary endpoints (PFS, OS), analysis populations (ITT, PD-L1-positive$_{25\%}$, and PD-L1-positive$_{1\%}$ populations), and treatment regimens (MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC). If the higher level hypothesis in the MTP is rejected for superiority, the following hypothesis will then be tested as shown in Figure 3.

Hypotheses will be tested using a multiple testing procedure with an alpha-exhaustive recycling strategy (Burman et al 2009). With this approach, hypotheses will be tested in a pre-defined order as outlined in Figure 3. According to alpha (test mass) splitting and alpha recycling, the test mass that becomes available after each rejected hypothesis is recycled to secondary hypotheses not yet rejected. Since OS is tested at multiple timepoints (ie, 2 interim analyses and final analysis), the OS tests that for the same comparison/population (ie, shown in 1 box in the MTP) will be considered as 1 test family. As long as one test in the family can be rejected, the family is rejected thus the assigned total alpha to the family can be recycled to next MTP level. This testing procedure stops when the entire test mass is allocated to non-rejected hypotheses. Implementation of this pre-defined ordered testing procedure, including recycling, will strongly control type I error at 5% (2-sided), among all key hypotheses. Figure 3 shows the multiple testing framework.

The details on how the alpha will be spent / controlled in all the possible scenarios are outlined below.

1. Test $H_{0,1}$, $H_{0,2}$, and $H_{0,3}$ at level 0.5%, 3% and 1.5% respectively.

(a) If none of the 3 tests is significant, accept $H_{0,1,}$ $H_{0,2,}$ and $H_{0,3}$ and stop procedure.

(b) For the PFS tests in the 1st column, if $H_{0,1}$ is statistically significant at 0.5% level, then reject $H_{0,1}$ (but accept $H_{0,2,}$ and $H_{0,3}$) and continue testing $H_{0,4}$ at 0.5% level.

(i) If $H_{0,4}$ is not statistically significant at 0.5% level, accept $H_{0,4}$ and stop the procedure.

(ii) If $H_{0,4}$ is statistically significant at 0.5% level, then reject $H_{0,4}$ and continue testing $H_{0,6}$ at 0.5% level

(iii) If $H_{0,6}$ is not statistically significant at 0.5% level, accept $H_{0,6}$ and stop the procedure.

(iv) If $H_{0,6}$ is statistically significant at 0.5% level, then reject $H_{0,6}$ and continue testing $H_{0,8}$ at 0.5% level

(v) If $H_{0,8}$ is not statistically significant at 0.5% level, accept $H_{0,8}$ and stop the procedure.

(vi) If $H_{0,8}$ is statistically significant at 0.5% level, then reject $H_{0,8}$ and the 0.5% alpha will be recycled to $H_{0,3.}$

(c) For the OS tests in the 2nd and 3rd columns of the MTP, if $H_{0,2}$ is statistically significant at 3% level, and $H_{0,3}$ is not statistically significant at 1.5% level, then reject $H_{0,2}$ and accept $H_{0,3,}$ and continue testing $H_{0,5}$ at 3% level.

(i) If $H_{0,5}$ is not statistically significant at 3% level, accept $H_{0,5}$ and stop the procedure.

(ii) If $H_{0,5}$ is statistically significant at 3% level, then reject $H_{0,5}$ and continue testing $H_{0,7}$ at 3% level.

(iii) If $H_{0,7}$ is not statistically significant at 3% level, accept $H_{0,7}$ and stop the procedure.

(iv) If $H_{0,7}$ is statistically significant at 3% level, then reject $H_{0,7}$ and the 3% alpha will be recycled to $H_{0,3.}$

(d) If $H_{0,3}$ is statistically significant at 1.5% level, and $H_{0,2}$ is not statistically significant at 3% level, then reject $H_{0,3}$ and accept $H_{0,2,}$ and continue testing $H_{0,6}$ at 1.5% level.

(v) If $H_{0,5}$ is not statistically significant at 1.5% level, accept $H_{0,5}$ and stop the procedure.

(vi) If $H_{0,5}$ is statistically significant at 1.5% level, then reject $H_{0,5}$ and continue testing $H_{0,7}$ at 1.5% level.

(vii) If $H_{0,7}$ is not statistically significant at 1.5% level, accept $H_{0,7}$ and stop the procedure.

(viii) If $H_{0,7}$ is statistically significant at 1.5% level, then reject $H_{0,7}$ and the 1.5% alpha will be recycled to $H_{0,2.}$

(e) If both the two hypotheses $H_{0,2}$ and $H_{0,3}$ are statistically significant at 3% and 1.5% levels, respectively, then reject both $H_{0,2}$ and $H_{0,3}$ and recycle 4.5% alpha from tests of $H_{0,2}$ and $H_{0,3}$ to test of $H_{0,5.}$

(ix) If $H_{0,5}$ is not statistically significant at 4.5% level, accept $H_{0,5}$ and stop the procedure.

(x) If $H_{0,5}$ is statistically significant at 4.5% level, reject $H_{0,5}$ and continue testing $H_{0,7}$ at 4.5% level.

(xi) If $H_{0,7}$ is not statistically significant at 4.5% level, accept $H_{0,7}$ and stop the procedure..

(xii) If $H_{0,7}$ is statistically significant at 4.5% level, reject $H_{0,7}$

The co-primary endpoint OS is tested at 2 interim and a final timepoint. The OS tests for the same comparison/population (e.g., $H_{0,2}$) will be considered as 1 test family. The alpha level allocated to one OS test family will be controlled at the interim and primary time points by using the Lan DeMets (Lan and DeMets 1983) spending function that approximates an O'Brien Fleming approach, where the alpha level applied at the interim depends upon the proportion of information available. The first OS interim analysis for superiority will occur at the primary PFS analysis, when it is expected that approximately 68% of the target death events may occur.

The second OS interim will subsequently be performed at approximately 80% of the target death events, with the primary OS analysis performed when 225 deaths have accumulated in the PD-L1-positive$_{25\%}$ population. If exactly 68% and 80% of the target events in the PD-L1-positive$_{25\%}$ population are available at the time of the first and second interims analyses, respectively (ie, 152/225 and 180/225 deaths have occurred), with an overall 2-sided alpha level 0.015 and 0.03 respectively for the comparisons of MEDI4736 + tremelimumab versus SoC and MEDI4736 monotherapy versus SoC, the 2-sided alpha to be applied for the interim and final analyses would be 0.0023, 0.0049 and 0.0132 for the comparisons of MEDI4736 + tremelimumab versus SoC and 0.0062, 0.0113 and 0.0258 for the comparisons of MEDI4736 monotherapy versus SoC.

If the interim or final analyses indicate superiority in OS for either monotherapy or combination therapy in PD-L1 positive$_{25\%}$, then subsequent analyses of secondary OS

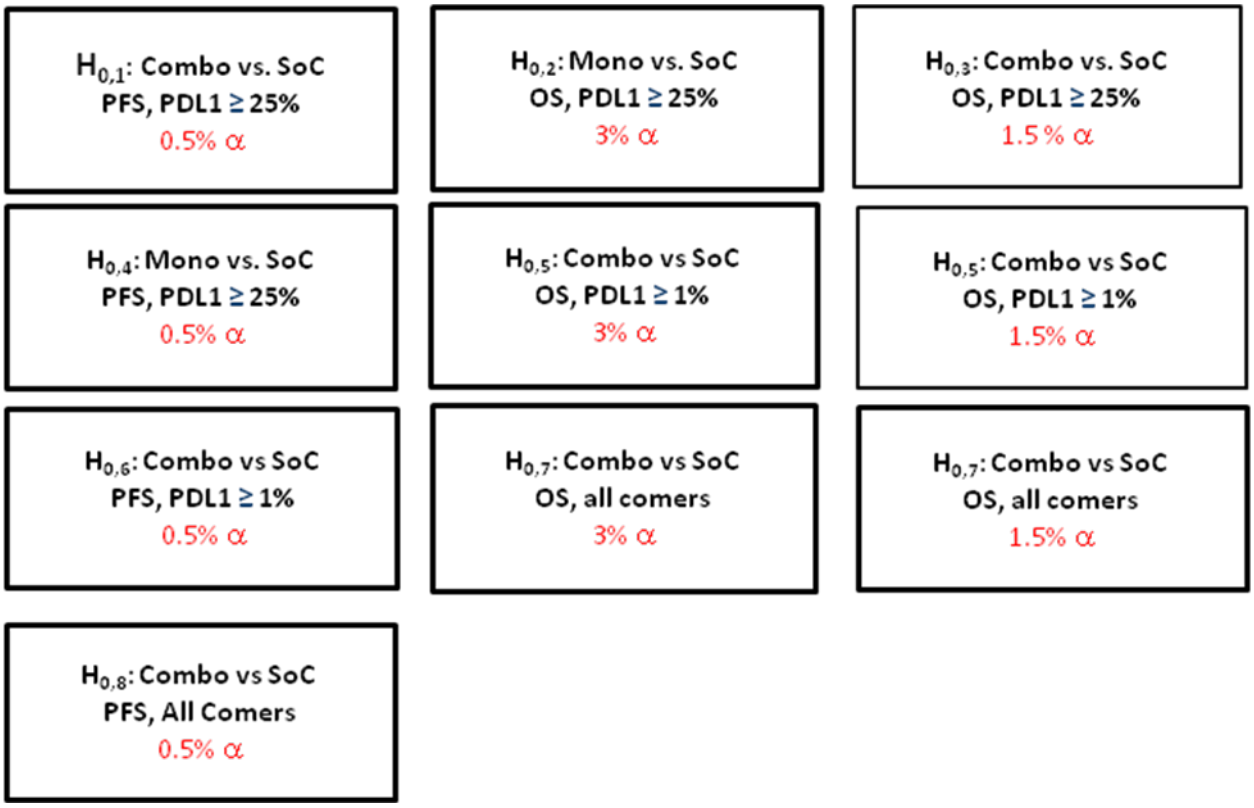
endpoints will be performed in accordance with the hierarchical testing strategy. A separate Lan DeMets (O'Brien Fleming) spending function will be used to determine the alpha levels at the interim and final analyses for testing the PD-L1 positive$_{1\%}$ and the all-comers hypotheses.

The timing of analysis for the secondary endpoint of PD-L1-positive$_{1\%}$ population will be defined by meeting the target number of events for the primary endpoint in the PD-L1-positive$_{25\%}$ population. The sample size of N=236 patients per arm in the PD-L1-positive$_{1\%}$ population (as specified in Protocol Section 8.2 Sample size estimate) assumes that the prevalence in this population will be 65%. The final target number of events of 337 deaths across the MEDI4736 + tremelimumab combination and standard of care arm then assumes that the final analysis will be conducted after approximately 71% of the PD-L1-positive$_{1\%}$ patients have had a death event. The alpha levels at the interim and final analyses are subsequently computed using the Lan DeMets (O'Brien Fleming) spending function, based on the anticipated 337 death events in the PD-L1-positive$_{1\%}$ population.

Although a 65% prevalence of PD-L1-positive$_{1\%}$ population has been assumed in the protocol to calculate the sample size and subsequently the target number of events, it is possible that the actual prevalence of PD-L1-positive$_{1\%}$ population may be higher (or lower) than what has been assumed. If that is the case, the total number of PD-L1-positive$_{1\%}$ patients will be larger (or smaller) than N=236 patients per arm. Therefore, the target number of 337 events may no longer reflect a maturity of 71% at the time of final analysis, as intended for this analysis.

To avoid this issue, the alpha levels for the two interim analyses and the final analysis for OS in the PD-L1-positive$_{1\%}$ population will then be computed based on the assumption that approximately 71% of the actual randomized PD-L1-positive$_{1\%}$ patients between the MEDI4736 + tremelimumab combination and standard of care arms will have a death event at the final OS analysis. If the interim results do not meet the criterion of stopping for superiority for a given hypothesis, then follow-up will continue until the final target number of OS events for that comparison has been observed, following which the hypothesis will be re-tested. If the hypothesis is then rejected, subsequent testing will continue hierarchically. The above testing procedure will ensure strong control of the family-wise error rate (Glimm et al, 2010).

Figure 3 Multiple testing procedures for controlling the type 1 error rate



Combo MEDI4736 + tremelimumab combination therapy; Mono MEDI4736 monotherapy; SoC Standard of care.

Note: The alpha values shown in Figure 3 assume all the tests in the MTP higher levels are rejected successfully. The actual alpha for the lower level tests in the MTP depends on how many higher level tests are rejected. The general approach of alpha splitting shown in Figure 3 will be applied to the actual alpha after considering the higher level tests results.

4.2.2 Co-Primary endpoints

4.2.2.1 Progression-free survival

The co-primary PFS analyses will be based on the programmatically derived RECIST 1.1 using the BICR tumor assessments. The co-primary analysis will be performed in the PD-L1-positive$_{25\%}$ population using a stratified log-rank test adjusting for histology (squamous versus non-squamous) only. The effect of MEDI4736 + tremelimumab versus SoC treatment will be estimated by the HR together with its corresponding 99.5% CI and p-value.

The covariates in the statistical modeling will be based on the values entered into IVRS at randomization, even if it is subsequently discovered that these values were incorrect.

The HR and its CI can be estimated from the Cox proportional hazards model (Cox 1972):

Secondary analyses will be performed using the same methodology as for the primary analyses described above.

Kaplan-Meier plots of PFS will be presented by treatment group, and by treatment group and PD-L1 tumor status subgroup, where appropriate. Summaries of the number and percentage of patients experiencing a PFS event and the type of event (RECIST 1.1 or death) will be provided along with median PFS for each treatment.

The assumption of proportionality will be assessed, initially only with regards to the primary treatment comparison. Proportional hazards will be tested firstly by examining plots of complementary log-log (event times) versus log (time) and, if these raise concerns, by fitting a time-dependent covariate to assess the extent to which this represents random variation. If a lack of proportionality is evident, the variation in treatment effect will be described by presenting piecewise HR calculated over distinct time-periods. In such circumstances, the HR can still be meaningfully interpreted as an average HR over time unless there is extensive crossing of the survival curves. If lack of proportionality is found, this may be a result of treatment-by-covariate interactions, which will be investigated. In addition, the KM curve along with landmark analyses (e.g., one year PFS rate) will also help in understanding the treatment benefit.

The treatment status at progression of patients at the time of analysis will be summarised. 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. This will also provide distribution of number of days prior to progression for the patients who have discontinued treatment.

Additional supportive summaries/graphs

In addition, the number of patients prematurely censored will be summarised by treatment group together with baseline prognostic factors of the prematurely censored patients. A patient would be defined as prematurely censored if they had not progressed (or died in the absence of progression) and the latest scan prior to DCO was more than one scheduled tumour assessment interval plus 2 weeks (8 weeks if time period between randomisation and DCO for that patient is 48 weeks or less; 10 weeks otherwise) prior to the DCO date.

Additionally, summary statistics will be given for the number of days from censoring to data cut-off for all censored patients.

A summary of the duration of follow-up will be summarised using median time from randomisation to date of censoring (date last known to be non-progressor) in censored (not progressed) patients only, presented by treatment group.

Additionally, summary statistics for the number of weeks between the time of progression and the last evaluable RECIST 1.1 assessment prior to progression will be presented for each treatment group.

Summaries of the number and percentage of patients who miss two or more consecutive RECIST assessments will be presented for each treatment group.

All of the collected RECIST 1.1 data will be listed for all randomised patients. In addition, a summary of new lesions (i.e., sites of new lesions) will be produced.

Sensitivity Analyses

The following sensitivity analyses will only be performed for the primary treatment comparison.

- Evaluation-Time bias

Sensitivity analyses will be performed to assess possible evaluation-time bias that may be introduced if scans are not performed at the protocol-scheduled time points. The midpoint between the time of progression and the previous evaluable RECIST 1.1 assessment will be analyzed using a log-rank test. For patients whose death was treated as PFS event, the date of death will be used to derive the PFS time used in the analysis. This approach has been shown to be robust even in highly asymmetric assessment schedules (Sun and Chen 2010).

- Attrition bias

Attrition bias will be assessed by repeating the PFS analysis except that the actual PFS event times, rather than the censored times, of patients who progressed or died in the absence of progression immediately following 2 or more non-evaluable tumor assessments will be included. In addition, patients who take subsequent therapy prior to progression or death will be censored at their last evaluable assessment prior to taking the subsequent therapy. This analysis will be supported by a Kaplan-Meier plot of the time to censoring where the censoring indicator of the PFS analysis is reversed.

- Ascertainment bias

Ascertainment bias will be assessed by analyzing the site investigator data. The stratified log-rank test will be repeated on the programmatically derived PFS using the site investigator data based upon RECIST. The HR and CI will be presented.

If there is an important discrepancy between the primary analysis using the BICR assessments and this sensitivity analysis using investigator assessments, then the proportion of patients with site but no central confirmation of progression will be summarised; such patients have the potential to introduce bias in the central review due to informative censoring. An approach that imputes an event at the next visit in the central review analysis may help inform the most likely HR value (Fehrenbacher et al 2016, Fleischer et al 2011), but only if an important discrepancy exists.

Disagreements between investigator and central reviews of RECIST 1.1 progression will be presented for each treatment group. The summary will include the early discrepancy rate

which is the frequency of central review declared progressions before the investigator review as a proportion of all central review progressions and the late discrepancy rate which is the frequency of central review declared progressions after the investigator review as a proportion of all discrepancies.

An additional sensitivity analysis will be performed for the primary treatment comparisons which will use the CRF values for the covariate histology in the statistical modelling rather than the values from IVRS.

Secondary Analysis

Secondary analyses of PFS based on the programmatically derived RECIST 1.1 using the BICR tumor assessments will be performed for the following treatment comparisons (in accordance with the multiple testing procedure):

- MEDI4736 + tremelimumab versus SoC (ITT population)
- MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{1\%}$ population
- MEDI4736 monotherapy versus SoC for PD-L1-positive$_{25\%}$ population (stratified only for histology)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$ population (stratified only for histology)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{1\%}$ population

These analyses will be performed using a stratified log-rank test. The effect of treatment will be estimated by the HR together with its corresponding 95% CI and p-value, The HR and CI will be estimated using the same approach as specified above for the primary analysis of PFS.

Exploratory Analyses

An exploratory analysis of PFS using BICR assessment based on RECIST 1.1 modified for confirmation of progression as well as PFS based on BICR assessments according to irRECIST 1.1 criteria will be performed. The stratified log-rank test used for the primary analysis of PFS will be repeated. The HR and CI will be presented.

A forest plot illustrating the hazard ratio and 95% confidence interval will be provided to compare the primary and sensitivity analyses of progression free survival.

Subgroup analyses and a forest plot will be generated comparing PFS between treatments.

No adjustment to the significance level for testing will be made since all these subgroup and sensitivity analyses will be considered supportive of the primary analysis of PFS.

The effect of covariates upon the HR estimate and the consistency of treatment effect between subgroups will be analysed for PFS.

Subgroup Analyses

Subgroup analyses will be conducted comparing PFS (per RECIST 1.1) between MEDI4736 + tremelimumab versus SoC and between MEDI4736 monotherapy versus SoC in the following subgroups of the PD-L1-positive$_{25\%,}$ PD-L1-positive$_{1\%}$ analysis sets and the FAS as deemed appropriate (but not limited to):

- Sex (male versus female)
- Age at randomization (<65 versus ≥65 years of age)
  - This will be determined from the date of birth (BIRTHDAT in the DEM module) and date of randomisation (RND_DAT in the CRIT1 module) on the eCRF at screening. Patients with a partial date of birth (ie for those countries where year of birth only is given) will have an assumed date of birth of 1st Jan [given year]). Patients with a missing age value will be included using the mean age (overall FAS) and categorised accordingly.
- PD-L1 status (<25% versus ≥25%)
- PD-L1 status using additional cutpoints of 1%, 10% and 50% tumour expression: (<1% versus ≥1%,<10% versus ≥10% and <50% versus ≥50%), plus 25% immune cell expression (<25% versus ≥25%)
- Histology (squamous versus non-squamous)
- Smoking (smoker versus non-smoker [never smoker])
  - This will be determined from the response to 'Nicotine Use Occurrence' (SU module) on the eCRF at screening. Patients with a missing smoking status will be included in the 'smoker' category.
- Race (Asian versus non-Asian)

The subgroup analyses for the stratification factors will be based on the values entered into the IVRS, all other factors will be based on values recorded on the eCRF as indicated above.

Other baseline variables may also be assessed if there is clinical justification or an imbalance is observed between the treatment groups. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic and/or predictive factors.

No adjustment to the significance level for testing of the subgroup and sensitivity analyses will be made since all these analyses will be considered supportive of the primary analysis of PFS and OS.

For each subgroup, the HR and 95% CI will be calculated from an unstratified Cox proportional hazards model with treatment as the only covariate. The Cox models will be fitted using SAS® PROC PHREG with the Efron method to control for ties, using the by statement to obtain HR and 95% CI for each subgroup level separately.

These hazard ratios and associated two-sided 95% CIs will be summarised and presented on a forest plot, along with the results of the overall primary analysis.

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less than 20 events in a subgroup), the relationship between that subgroup and PFS will not be formally analysed. In this case, only descriptive summaries will be provided.

Unless there is a marked difference between the results of the statistical analyses of the PFS from the BICR data and that of the site Investigator tumor data, these subgroup analyses will only be performed on the PFS endpoint using the BICR data.

Effect of covariates on HR estimate

Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. A model will be constructed, containing treatment and the stratification factors alone, to ensure that any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

The result from the initial model and the model containing additional covariates will be presented.

Additional covariates for this model will include age at randomization, sex, smoking status, race, and number target lesions. Same as the other efficacy analyses, this multivariate Cox model will be performed in FAS, PD-L1-positive$_{25\%}$ population, and PD-L1-positive$_{1\%}$ population.The stratification factor PD-L1 status (<25% versus ≥25%) will not be included in the analysis for PD-L1-positive$_{25\%}$ population. This analysis evaluates the treatment effect adjusting for any potential imbalances in baseline prognostic factors.

The model will include the effect regardless of whether the inclusion of effect significantly improves the fit of the model providing there is enough data to make them meaningful.

Consistency of treatment effect between subgroups

Interactions between treatment and stratification factors will also be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

4.2.2.2 Overall survival

OS in the PD-L1-positive$_{25\%}$ population will be analyzed using a stratified log-rank tests, using the same methodology as described for the primary PFS endpoint. The effects of MEDI4736 + tremelimumab versus SoC and MEDI4736 monotherapy versus SoC will be

estimated by the HR together with its corresponding CI (98.5% for MEDI4736 + tremelimumab versus SoC and 97% for and MEDI4736 monotherapy versus SoC, adjusted for the interim analyses) and p-value. Kaplan-Meier plots will be presented by treatment group. 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 for each treatment.

The assumption of proportionality will be assessed in the same way as for PFS.

OS will be analysed at the time of the primary PFS analysis (OS interim analysis 1), at a second interim analysis (when 80% of OS target events have occurred) and at a final analysis (when target of 225 OS events have occurred in the PD-L1-positive$_{25\%}$ population)

The alpha will be split between the final and the 2 interim analyses using the hierarchical testing strategy as already described in Section 4.2.1. The boundaries (ie, adjusted alpha levels) for the treatment comparison at the interim and final analyses for OS will be derived based upon the exact number of OS events using the Lan and DeMets approach that approximates the O'Brien Fleming spending function (see Section 5.1)

A sensitivity analysis will be performed for the primary treatment comparisons which will use the CRF values for the covariate histology in the statistical modelling rather than the values from IVRS.

Secondary Analysis

Secondary analyses of OS will be performed for the following treatment comparisons (in accordance with the multiple testing procedure):

- MEDI4736 + tremelimumab versus SoC for PD-L1-positive$_{1\%}$ population
- MEDI4736 + tremelimumab versus SoC (ITT population)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy (ITT population)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$ population (stratified only for histology)
- MEDI4736 + tremelimumab versus MEDI4736 monotherapy for PD-L1-positive$_{25\%}$ population (stratified only for histology)

Subgroup Analyses

Subgroup analyses will be conducted comparing OS between MEDI4736 + tremelimumab versus SoC and between MEDI4736 monotherapy versus SoC in the same subgroups as specified for the PFS subgroup analyses.

Effect of covariates on HR estimate

Cox proportional hazards modeling will be employed to assess the effect of covariates on the HR estimate. A model will be constructed, containing treatment and the stratification factors

alone, to ensure that any output from the Cox modeling is likely to be consistent with the results of the stratified log-rank test.

The result from the initial model and the model containing additional covariates will be presented.

Additional covariates for this model will include age at randomization, sex, smoking status, race, and number target lesions. Same as the other efficacy analyses, this multivariate Cox model will be performed in FAS, PD-L1-positive$_{25\%}$ population, and PD-L1-positive$_{1\%}$ population.The stratification factor PD-L1 status (<25% versus ≥25%) will not be included in the analysis for PD-L1-positive$_{25\%}$ population. This analysis evaluates the treatment effect adjusting for any potential imbalances in baseline prognostic factors.

The model will include the effect regardless of whether the inclusion of effect significantly improves the fit of the model providing there is enough data to make them meaningful.

Consistency of treatment effect between subgroups

Interactions between treatment and stratification factors will also be tested to rule out any qualitative interaction using the approach of Gail and Simon (Gail and Simon 1985).

Impact of switching (crossover outside of this study) to immunotherapies (or other potentially active investigational agents) on OS analyses

Exploratory analyses of OS adjusting for the impact of subsequent immunotherapy or other investigational treatment may be performed, if a sufficient proportion of patients switch. Methods such as Rank Preserving Structural Failure Time (Robins and Tsiatis 1991), Inverse Probability of Censoring Weighting (Robins 1993) and other methods in development will be explored. The decision to adjust and the final choice of methods will be based on a blinded review of the data and the plausibility of the underlying assumptions. Baseline and time-dependent characteristics will be explored, and summaries of baseline characteristics will be summarised by treatment group, splitting between those that have and haven't switched at the time of the analyses. Further detail will be provided in the Payer Analysis Plan. These analyses are intended to support reimbursement appraisals.

Subsequent therapies received after discontinuation of treatment will be summarised and listed by treatment group. Patients who subsequently received an immunotherapy agent or entered an immunotherapy trial will be summarised and listed by treatment group according to line of subsequent therapy, i.e. immediately after immunotherapy or as a later line.

4.2.3 Objective response rate

The ORR will be based on the programmatically derived RECIST 1.1 using the BICR tumor data. The ORR will be compared between MEDI4736 + tremelimumab versus SoC using logistic regression models adjusting for the same factors as the primary endpoint (PD-L1 tumor expression and histology). The results of the analysis will be presented in terms of an odds ratio (an odds ratio greater than 1 will favour MEDI4736+tremelimumab) together with

its associated profile likelihood CI (e.g. using the option 'LRCI' in SAS procedure GENMOD) and p-value (based on twice the change in log-likelihood resulting from the addition of a treatment factor to the model). This analysis will be performed in the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations. The analysis of the PD-L1-positive$_{25\%}$ population will be performed using a logistic regression model adjusting for only histology.

Additional analyses of ORR comparing MEDI4736 monotherapy versus SoC in the PD-L1-positive$_{25\%}$ population and comparing MEDI4736 + tremelimumab versus MEDI4736 monotherapy versus SoC in the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations will be performed.

This analysis of ORR will be repeated using the results of the programmatically derived ORR using the site investigator tumour data based upon RECIST 1.1 as a sensitivity analysis to confirm the results of the primary analysis derived from the CRFs.

ORR by irRECIST 1.1 criteria using BICR assessments will also be reported in the ITT population.

If there are not enough responses for a meaningful analysis using logistic regression then a Fisher's exact test using mid p-values will be presented.

The mid-p-value modification of the Fisher exact test amounts to subtracting half of the probability of the observed table from Fisher's p-value.

Fisher's exact test mid p-value = Two-sided p-value – (Table probability / 2)

Summaries will be produced that present the number and percentage of patients with a tumor response (CR/PR). Overall visit response data will be listed for all patients (ie, the FAS). For each treatment group, best objective response (BoR) will be summarized by n (%) for each category (CR, PR, SD, PD, and NE). No formal statistical analyses are planned for BoR.

Subgroup Analyses

Subgroup analyses will be conducted comparing ORR between MEDI4736 + tremelimumab versus SoC and between MEDI4736 monotherapy versus SoC in the same subgroups as specified for the PFS subgroup analyses.

For each subgroup, the odds ratio and 95% CI will be calculated from a logistic regression model with treatment as the only covariate. These odds ratios and associated two-sided 95% CIs will be summarised and presented on a forest plot, along with the results of the overall primary analysis.

If there are too few responses available for a meaningful analysis of a particular subgroup, the relationship between that subgroup and ORR will not be formally analysed. In this case, only descriptive summaries will be provided.

4.2.4 Duration of response

In order to analyze the DoR between MEDI4736 + tremelimumab and SoC, the expected duration of response (EDoR) will be derived for each treatment group (Ellis et al 2008) using the BICR tumor data. The EDoR is the product of the proportion of patients responding to treatment and the mean DoR in responding patients and provides an estimate based on all randomized patients. Treatments will be compared by calculating the ratio of EDoRs, using the weibull distribution (unless the data suggests otherwise) for DoR in responding patients. Additionally, descriptive data will be provided for the DoR in responding patients (i.e. median duration of response and 95% CIs) by treatment group, including the associated Kaplan-Meier curves (without any formal comparison of treatment groups or p-value attached).

This analysis will be performed on the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations.

Additional analyses of DoR comparing MEDI4736 monotherapy versus SoC in the PD-L1-positive$_{25\%}$ population will be performed.

4.2.5 Proportion of patients alive and progression free at 12 months

The APF12 (where 12 months equates to study day 366) will be summarized (using the Kaplan-Meier curve) and presented by treatment group. APF12 will be compared between MEDI4736 + tremelimumab and SoC by using the Kaplan-Meier estimator of PFS at 12 months for each treatment to obtain the HR. The HR and CI will be presented using the following approach (Klein et al 2007):

- The *HR(group1:group2)* is estimated as $\frac{\ln \hat{S}_1(t)}{\ln \hat{S}_2(t)}$
- The variance for *ln(HR)* is estimated as $\frac{\hat{\sigma}_1(t)^2}{\ln^2 S_1(t)} + \frac{\hat{\sigma}_2(t)^2}{\ln^2 S_2(t)}$

where $\hat{\sigma}_i(t)^2 = \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}$ is the variance for ln{S(t) derived from greenwood's formula for the variance of *S(t)* and can be estimated from standard software packages, where d_i and n_i refer to the number of events and patients at risk for each risk set.

The *ln(HR)* and its variance in each strata will be estimated and combined by weighting inversely proportionately according to each within stratum variance (Whitehead and Whitehead 1991).

This analysis will be performed in the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations.

Additional analysis of APF12 comparing MEDI4736 monotherapy versus SoC in the PD-L1-positive$_{25\%}$ population will be performed.

4.2.6 Time from randomization to second progression

PFS2 is defined as the time from the date of randomization to the earliest of the progression event subsequent to that used for the PFS endpoint or death. PFS2 in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint. The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding 95% CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

This analysis will be performed in the PD-L1-positive$_{25\%}$, PD-L1-positive$_{1\%}$ and ITT populations.

Additional analysis of PFS2 comparing MEDI4736 monotherapy versus SoC in the PD-L1-positive$_{25\%}$ population will be performed.

For supportive purposes, the time to the start of subsequent therapy will be analyzed using the same methodology and model. The HR for the treatment effect together with its 95% CI will be presented. In addition, a Kaplan-Meier plot of the time to the start of subsequent therapy will be presented by treatment group and the time between progression and starting subsequent therapy will be assessed. No multiplicity adjustment will be applied as these are viewed as supportive endpoints.

A summary table of first subsequent therapies by treatment group will be provided,.

4.2.7 Change in tumour size

The absolute values and percentage change in target lesion tumour size from baseline will be summarized using descriptive statistics and presented at each timepoint for each treatment group. The best change in target lesion tumour size from baseline, (where best change in target lesion size is the maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction) will also be summarised and presented for each treatment group.

Tumour size will also be presented graphically using waterfall plots for each treatment group, to present each subject's best percentage change in tumour 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 tumour size levels will be added to the plots, which correspond with the definitions of progression and 'partial' response respectively. On each of the waterfall plots the histology classification (Squamous versus All other) of each patient will be indicated. Additional waterfall plots showing percentage change in tumour size at specific timepoints may be produced if it is felt that these are warranted to provide greater clarity.

The above outputs will be programmed for the BICR RECIST 1.1 assessments.

4.2.8 Patient reported outcomes

Health related quality of life will be assessed using the EORTC QLQ-C30, Lung cancer symptoms will be assessed with the QLQ-LC13 module.

Treatment efficacy will be evaluated primarily on what patients and clinicians consider the primary symptoms of lung cancer (NSCLC working group material, presentation at ASCO 2016): cough, dyspnea, pain (in the chest, pain in other parts of the body) as well as fatigue and appetite loss. The assessments of cough,,dyspnea (breathlessness) as well as pain as assessed by the EORTC QLQ LC13 and fatigue and appetite loss from EORTC QLQ C30 will be used as secondary efficacy endpoints.

For these secondary efficacy endpoints, the overall type I error (5% 2 sided) will be controlled across the five primary PRO measures of cough, dyspnoea and pain as assessed by the EORTC QLQ-LC13 and fatigue and appetite loss as assessed by the EORTC QLQ-C30 using the Bonferroni-Holm procedure (Holm 1979).

The physical functioning and global health status/QoL domains of the EORTC QLQ C30 are furthermore pre-specified endpoints of interest

4.2.8.1 EORTC QLQ-C30

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 3 symptom scales and the 5 individual symptom items will be produced. Similarly, a summary of improvement rate for each of the 5 function scales (physical, role, emotional, cognitive, and social) and global health status/QoL will be produced.

Summaries of original and change from baseline values of each symptom scale/item, the global health status/QoL score and each functional domain will be reported by assessment period for each treatment group. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each assessment for each ordinal item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration) will also be produced for each treatment group.

4.2.8.2 EORTC QLQ-LC13

Time to deterioration in the ITT population will be analyzed using a stratified log-rank test, using the same methodology as described for the primary PFS endpoint.

The effect of MEDI4736 + tremelimumab versus SoC will be estimated by the HR together with its corresponding CI and p-value. Kaplan-Meier plots will be presented by treatment group. Summaries of the number and percentage of patients who have an event as well as who were censored will be provided along with the medians for each treatment.

A summary of the symptom improvement rate for each of the 6 individual symptom items will be produced.

Summaries of original and change from baseline values of each symptom (dyspnea, cough, hemoptysis, chest pain, arm/shoulder pain, other pain) and each treatment-related side effect (sore mouth, dysphagia, peripheral neuropathy and alopecia) will be reported by assessment periodfor each treatment group. Graphical presentations may also be produced as appropriate. Summaries of the number and percentage of patients in each response category at each assessmentfor each ordinal symptom item (in terms of the proportion of patients in the categories of improvement, no change, and deterioration) will also be produced for each treatment group.

4.2.8.3 Mixed models repeated measures of change from baseline in PRO symptoms

In addition to the time to deterioration endpoints listed above the following longitudinal endpoints are of interest: fatigue and appetite loss; LC13 dyspnoea, cough and chest pain. These are not part of the main multiple testing procedure and are considered a separate set of PRO endpoints from the time to deterioration endpoints listed above. A Bonferroni adjustment to the significance level will be applied to the tests described below to control the overall Type I error at the 5% level.

Change from baseline in these pre-specified the PRO symptom scores of dyspnea, cough, chest pain, fatigue and appetite loss, will be analysed using a mixed model for repeated measures (MMRM) analysis with use of all data from baseline up to 12 months. The analysis will be to compare the average treatment effect from the point of randomisation until PD or 12 months (whichever is earlier) unless there is excessive missing data (defined as >75% missing data). It is acknowledged that patients will discontinue treatment at different timepoints during the study and that this is an important time with regards to symptoms and HRQoL data collection. To account for this, and in order to include the discontinuation and follow up assessments, a generic assessment time variable will be derived for each subject in order that the average treatment effect can be analyzed using the above method. Each visit will be assigned a sequential number. The time from randomization to each of these will be derived in order to select only those assessment times occurring within the first 12 months of randomization or until PD.

As an example, say a patient X has data collected at the first 4 scheduled assessments of a 4-weekly schedule and then discontinues treatment, whilst patient Y discontinues treatment after the first scheduled assessment, the first 6 generic assessment times would be as follows:

Generic assessment time	Study Day

	Patient X	Patient Y
Baseline	Baseline	Baseline
1	29	28
2	57	50 (discontinuation)
3	85	85
4	113	113
5	130 (discontinuation)	141
6	169	169

The MMRM model will include treatment, age at randomisation (<65 vs ≥65 years of age), sex (male vs female), smoking history (smoker vs non-smoker), assessment time and the interaction between treatment and assessment time interaction as fixed factors, baseline as a covariate and further adjusted for the interaction between baseline and assessment time. Restricted maximum likelihood (REML) estimation will be used. An overall adjusted mean estimate will be derived that will estimate the average treatment effect over assessments giving each assessment equal weight. For this overall treatment comparison, adjusted mean estimates per treatment group and corresponding 95% CIs will be presented along with an estimate of the treatment difference, 95% CI and p-value

An unstructured covariance matrix will be used to model the within-subject error and the Kenward-Roger approximation will be used to estimate the degrees of freedom. If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be tried in order until convergence is reached: toeplitz with heterogeneity, autoregressive with heterogeneity, toeplitz, autoregressive and compound symmetry.

Multiple imputation techniques for missing values may be considered to explore the robustness of any treatment effect.

These analyses will be performed in the FAS, and PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations as deemed appropriate.

An effect size estimate to interpret the magnitude of the effect and potential therapeutic benefit will be further specified in the PAP.

4.2.8.4 PRO-CTCAE

Data from the PRO-CTCAE will be summarized using FAS, and PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations as deemed appropriate. The number (%) of patients with each level

of response for each CTCAE item at baseline and over time will be summarized.. Further summaries to explore the data (i.e. the severity of symptoms) may be produced.

4.2.8.5 Patients' Global Impression of Change

PGIC data will be presented using summaries and descriptive statistics based on the FAS, and PD-L1-positive$_{25\%}$ and PD-L1-positive$_{1\%}$ populations as deemed appropriate.

4.2.8.6 EQ-5D-5L

Descriptive statistics, graphs, and listings will be reported for health state utility values and the visual analogue scale byassessment, as well as the change in these scores from baseline. To support future economic evaluations, additional appropriate analyses may be undertaken, for example, mean health state utility pre- and post-treatment, and pre- and post-progression, and will be outlined in the payer analysis plan.

4.2.9 Healthcare resource use

An exploratory health economic analysis of hospital episodes including type of contact (hospitalization, outpatient, or day case), reason, length of stay by ward type (including intensive care unit), procedures, and tests may be undertaken to examine the impact of disease and treatment on resource use to primarily support the economic evaluation of MEDI4736+tremelimumab in comparison to SoC, and will be outlined in the Payer Analysis Plan. This would include providing descriptive statistics as appropriate, including means, median, and ranges.

4.2.10 Safety data

Safety and tolerability data will be presented by treatment group using the full safety population and the subset of patients in the PD-L1-positive$_{25\%}$ population and PD-L1-positive$_{1\%}$ population. Safety data will be summarized only. No formal statistical analyses will be performed on the safety data. Data from all cycles of treatment will be combined in the presentation of safety data. AEs (both in terms of MedDRA preferred terms and CTCAE grade) will be listed individually by patient. The number of patients experiencing each AE will be summarized by treatment group and CTCAE grade. Additionally, data presentations of the rate of AEs per person-years at risk may be produced. Any safety summaries examining re-treatment with MEDI4736+tremelimumab will be produced separately, if required.

Other safety data will be assessed in terms of physical examination, clinical chemistry, hematology, vital signs, and ECGs. Exposure to MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC will be summarized. Time on study, MEDI4736 + tremelimumab, MEDI4736 monotherapy, and SoC, dose delays/interruptions and dose reductions will also be summarized. At the end of the study, appropriate summaries of all safety data will be produced.

The following sections describe the planned safety summaries. However, additional safety tables may be required to aid interpretation of the safety data. For example, if an imbalance is

seen in AEs or laboratory abnormalities that could be due to the differential follow-up periods (showing up more of the background/disease related AEs/abnormalities), additional summaries may be produced using a 30 day follow up period for both treatment arms to further explore / explain.

Adverse Events

All AEs, both in terms of current MedDRA preferred term and CTCAE grade, will be summarised descriptively by count (n) and percentage (%) for each treatment group. The current MedDRA dictionary will be used for coding. The majority of the AE summaries, unless stated otherwise, will be based on TEAEs. Any AE occurring before study treatment (i.e. before the administration of the first dose on Study Day 1) will be included in the AE listings, but will not be included in the summary tables (unless otherwise stated). These will be referred to as 'pre-treatment'. However, any AE occurring before the administration of the first dose on Study Day 1 that increases in severity after the first dose will be regarded as treatment emergent and thus will be included in the majority of summary tables.

AEs observed up until 90 days following discontinuation of the immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) / 30 days following discontinuation of the Standard of Care agent or until the initiation of the first subsequent anti-cancer therapy following discontinuation of treatment (whichever occurs first) will be used for reporting of all of the AE summary tables. This will more accurately depict AEs attributable to study treatment only as a number of AEs up to 90 days following discontinuation of the immunotherapy agents / 30 days following discontinuation of the Standard of Care agent are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, all of the AE summaries will also be produced containing AEs observed up until 90 days following discontinuation of the immunotherapy agents / 30 days following discontinuation of the Standard of Care agent (ie without taking subsequent anti-cancer therapy into account).

A selection of AE summaries may also be produced containing AEs (by system organ class and preferred term) observed from the initiation of the first subsequent anti-cancer therapy following discontinuation of study treatment until 90 days following discontinuation of immunotherapy agents / 30 days following discontinuation of the Standard of Care agent (ie summarising those AEs experienced by patients taking subsequent therapy during the AE collection follow-up window post discontinuation of study treatment). These outputs will only be produced if the number of AEs observed warrant the inclusion of such outputs for interpretational purposes. Any data post 90 days last dose for immunotherapy agents / 30 days following discontinuation of Standard of Care agents will be presented in a separate summary that presents any events that occur prior to dosing or starting more than 90/30 days after discontinuing treatment.

All reported AEs will be listed along with the date of onset, date of resolution (if AE is resolved) and investigator's assessment of severity and relationship to study drug. Frequencies and percentages of patients reporting each preferred term will be presented (i.e. multiple

events per patient will not be accounted for apart from any on episode level summaries which may be produced).

Summary information (the number and percent of patients by system organ class and preferred term separated by treatment group) will be tabulated for:

- All AEs
- All AEs causally related to study medication (as determined by the reporting investigator)
- AEs with CTCAE grade 3 or 4
- AEs with CTCAE grade 3 or 4 , causally related to study medication (as determined by the reporting investigator)
- AEs with outcome of death
- AEs with outcome of death causally related to study medication (as determined by the reporting investigator)
- AEs by outcome
- All SAEs
- All SAEs causally related to study medication (as determined by the reporting investigator)
- SAEs leading to discontinuation of study medication
- AEs leading to discontinuation of study medication
- AEs leading to discontinuation of study medication, causally related to study medication (as determined by the reporting investigator)
- AEs leading to hospitalization
- AEs leading to dose delay of study medication
- Other significant AEs
- Other significant AEs causally related to study medication (as determined by the reporting investigator)
- Immune mediated AEs (as determined by the reporting investigator)
- Infusion reaction AEs (as determined by the reporting investigator)

An overall summary of the number and percentage of patients in each category will be presented, as will an overall summary of the number of episodes in each category. In addition, a truncated AE table of most common AEs and another table showing most common AEs with CTCAE grade 3 or 4, showing all events that occur in at least 5% of patients overall will be summarised by preferred term, by decreasing frequency. This cut-off may be modified after review of the data. When applying a cut-off (ie, x %), the raw percentage should be compared to the cut-off, no rounding should be applied first (i.e., an AE with frequency of 4.9% will not appear if a cut-off is 5%). Summary statistics showing the time to onset and the duration of the first AE may also be presented as appropriate.

Each AE event rate (per 100 patient years) will also be summarised by preferred term within each system organ class for the output summarising all AEs. For each preferred term, the event rate is defined as the number of patients with that AE divided by the total drug exposure of patients at risk of AE. The denominator is calculated as the total over each patient of days from first dose to the earlier of the date of onset of the event or the last day of study medication.

Summaries of the number and percentage of patients will be provided by maximum reported CTCAE grade, system organ class, preferred term and treatment group.

Fluctuations observed in CTCAE grades during study will be listed for those AEs which are CTCAE ≥ 3.

In addition, all AEs will be listed.

Deaths

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

- Total number of deaths (regardless of the date of death)
- Death related to disease under investigation ONLY, as determined by investigator (regardless of the date of death)
- TEAE with outcome of death ONLY and onset date prior to initiation of subsequent anti-cancer therapy
- AE with outcome of death ONLY and onset date falling after 90 days following the date of last dose of immunotherapy /30 days following the date of last dose of SoC or initiation of subsequent anti-cancer therapy (whichever is earlier)
- Death related to disease under investigation, as determined by the investigator, and with TEAE with outcome of death and onset date prior to initiation of subsequent anti-cancer therapy

- Death related to disease under investigation, as determined by the investigator, and with AE with outcome of death and onset date falling after 90 days following the date of last dose of immunotherapy /30 days following the date of last dose of SoC or initiation of subsequent anti-cancer therapy (whichever is earlier)
- Death occurred 90 days after the date of last dose of immunotherapy/30 days following the date of last dose of SoC or after initiation of subsequent anti-cancer therapy (whichever is earlier), and unrelated to AE or disease under investigation
- Patients with unknown reason for death.
- Other deaths

These summaries will be produced twice; firstly accounting for subsequent therapy and, secondly, without taking subsequent therapy into consideration

Adverse events of special interest

Preferred terms used to identify adverse events of special interest (as defined in section 3.4.1) will be listed before DBL and documented in the Study Master File. Grouped summary tables of certain MedDRA preferred terms will be produced and may also show the individual preferred terms which constitute each AESI grouping.. Groupings will be based on preferred terms provided by the medical team prior to DBL, and a listing of the preferred terms in each grouping will be provided.

Summaries of the above-mentioned grouped AE categories will include number (%) of patients who have:

- At least adverse event of special interest presented by outcome
- At least one adverse event of special interest causally related to study medication
- At least one adverse event of special interest leading to discontinuation of study medication

A summary of total duration (days) of AESI will be provided for events which have an end date and this will be supported by summaries of ongoing AESIs at death and, separately, at data cut-off.

Additionally, there will be several summaries of AESIs requiring concomitant treatment, and particularly the relationship of AESIs to the use of immunosuppressive agents (ie, depicting which AESI triggered immunosuppressive use) and, separately, to the use of immunosuppressive agents at high doses.

Summary of long term tolerability

To assess long term tolerability, provided that there are a sufficient number of patients with events to warrant it, prevalence plots, life table plots and cumulative incidence plots will be presented for each of the AESI grouped terms and any other events considered important after review of the safety data, provided there are ≥ 10 events.

A prevalence plot provides information on the extent to which the events may be an ongoing burden to patients. The prevalence at time t after first dose of study treatment is calculated as the number of patients experiencing the event divided by the number of patients receiving study treatment or in safety follow-up at time t; generally, t is categorised by each day after dosing. The prevalence will be plotted over time and presented for each treatment group separately. Multiple occurrences of the same event are considered for each patient but a patient is only counted in the numerator whilst they are experiencing one of the occurrences of the event. These plots will only be produced for AESIs that have ≥10 events.

For each AE, median time to first onset of the AE from the date of first dose will be presented in patients in the safety analysis set by treatment group. Patients who did not experience the AE will be censored at the end of their safety follow-up. Summary tables of time to first onset for each AE will also be produced (e.g. 1-28 days, 29-56 days, 57-84 days, 85-112 days, >112 days). Median duration of the AE will be presented in patients who experienced each AE.

A life table plot can be used to describe the time to onset of the event and specifically when patients are at most risk of first experiencing the event. The hazard, or in other words, the probability of having an AE in a specified time period (e.g. 0-1 months, 1-3 months, 3-6 months, etc.) given that the patient reaches that time period without having an event is plotted for each time period. These plots will only be produced for AESIs that have ≥10 events.

A cumulative incidence plot is a plot of the raw cumulative incidence and cumulative incidence function over time with the treatment groups presented on separate plots. The raw cumulative incidence is the actual probability that a patient will have experienced their first occurrence of the event by a given time point. The cumulative incidence function estimates the cumulative incidence if the data cut-off had not been imposed and all patients had completed safety follow-up (Pintilie 2006). These plots will only be produced for AESIs that have ≥10 events.

Laboratory assessments

Data obtained up until the 90 days following discontinuation of immunotherapy agents (ie, the last dose of MEDI4736 or MEDI4736+tremelimumab) or 30 days following discontinuation of the Standard of Care agent or until the initiation of the first subsequent therapy following discontinuation of treatment (whichever occurs first) will be used for reporting. This will more accurately depict laboratory toxicities attributable to study treatment only as a number of toxicities up to 90 days following discontinuation of immunotherapy agents or 30 days following discontinuation of the Standard of Care agent are likely to be attributable to subsequent therapy.

However, to assess the longer term toxicity profile, summaries of laboratory data will also be produced containing data collected up until 90 days following discontinuation of the immunotherapy agents or up until 30 days following discontinuation of the Standard of Care agent (ie, without taking subsequent therapy into account).

A small selection of summaries of laboratory data may also be produced containing data from initiation of the first subsequent therapy following discontinuation of study treatment until 90 days following discontinuation of immunotherapy agents or until 30 days following discontinuation of the Standard of Care agent (ie summarising the laboratory data collected on patients taking subsequent therapy during the safety collection follow-up window post discontinuation of study treatment). These outputs will only be produced if the number of laboratory toxicities observed warrant the inclusion of such outputs for interpretational purposes. Any data post 90 days last dose for immunotherapy agents or post 30 days last dose for Standard of Care agents will not be summarised.

Data summaries will be provided in International System (SI) of units.

Scatter plots (shift plots) of baseline to maximum value / minimum value (ass appropriate) on treatment (i.e. on-treatment is defined as data collected between the start of treatment and the relevant follow-up period following the last dose of study treatment) may be produced for certain parameters if warranted after data review.

Box-plots of absolute values by week, and box-plots of change from baseline by week, may be presented for certain parameters if warranted after data.

For continuous laboratory assessments, absolute value and change from baseline will be summarised using descriptive statistics at each scheduled assessment time by actual treatment group

Shift tables for laboratory values by worst CTC 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 CTC grade shift outputs will be produced are:

- Haematology: Haemoglobin, Leukocytes, Lymphocytes, absolute count, Neutrophils, absolute count, Platelets
- Clinical chemistry: ALT, AST, ALP, Total bilirubin, Albumin, Magnesium – hypo and – hyper, Sodium – hypo and – hyper, Potassium – hypo and – hyper, Corrected calcium – hypo and – hyper, Glucose – hypo and – hyper, Creatinine

For the parameters with no CTCAE grading that are listed in the CSP, shift tables from baseline to worst value on-treatment will be provided. Additional summaries will include a shift table for urinalysis (Bilirubin, Blood, Glucose, Ketones, Protein) comparing baseline value to maximum on-treatment value.

Shift tables showing baseline to maximum and baseline to minimum will be produced for TSH, T3 and T4.

Liver Enzyme Elevations and Hy's law

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

- Elevated ALT, AST, and Total bilirubin during the study
  - ALT ≥ 3x –≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x and >20x Upper Limit of Normal (ULN) during the study
  - AST ≥ 3x- –≤ 5x, > 5x – ≤8x, > 8x - ≤ 10x, >10x - ≤ 20x,and >20x ULN during the study
  - Total bilirubin ≥2x-≤3x, >3x-≤5x, >5x ULN during the study
  - ALT or AST ≥3x - ≤5x, >5x - ≤8x, >8x - ≤ 10x, >10x - ≤ 20x, and >20x ULN during the study
  - ALT or AST ≥3x ULN and Total bilirubin ≥2x ULN during the study (Potential Hy's law): 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 ≥ 3x ULN plus Total bilirubin ≥ 2x ULN or AST ≥ 3x ULN plus Total bilirubin ≥ 2x ULN at any visit.

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

Plots of ALT and AST vs. Total bilirubin by treatment group 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.

Assessment of Thyrotoxicity

After the discontinuation of the study medication, the thyroid function tests, TSH, T3 and T4, were evaluated at 30 days after last dose, hence, the analysis of thyroid function tests will be based on data up to 30 days after the last dose of study medication or date of initiation of subsequent therapy (whichever occurs first).

Absolute value and change from baseline will be summarised using descriptive statistics at each scheduled assessment time.

Shift tables showing baseline to maximum and baseline to minimum will also be produced for TSH, T3 and T4, as deemed necessary.

ECGs

ECG data obtained up until the safety follow-up will be included in the summary tables. Overall evaluation of ECG is collected at each visit in terms of normal or abnormal, and the relevance of the abnormality is termed as "clinically significant" or "not clinically significant". A shift table of baseline evaluation to worst evaluation will be produced.

Vital signs

Vital signs data obtained up until the 30 day safety follow-up visit will be included in the summary tables.

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

Vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), pulse rate, temperature, respiratory rate and weight) will be summarised over time in terms of absolute values and changes from baseline at each scheduled measurement by actual treatment group.

Time to Subsequent Therapy from discontinuation of study treatment

A descriptive summary will be produced for time to subsequent therapy from discontinuation of study treatment. This summary is supportive of the Adverse Event and Laboratory data outputs.

Physical examination

All individual physical examination data will be listed only.

Other Safety Data

Data from positive pregnancy tests will be listed only.

4.2.11 WHO performance status

All WHO performance status will be summarised over time for the ITT population.

4.2.12 PK data (MEDI4736 monotherapy and MEDI4736+tremelimumab groups only)

Pharmacokinetic concentration data will be listed for each patient and each dosing day, and a summary provided for all evaluable patients in the PK analysis population..

4.2.13 Immunogenicity analysis

Immunogenicity results will be listed by patient and a summary will be provided of the number and percentage of patients who develop detectable anti-MEDI4736 and anti-tremelimumab antibodies based on the safety population. The immunogenicity titre and neutralizing ADA data will be listed for samples confirmed positive for the presence of anti-MEDI4736 antibodies and/or anti-tremelimumab antibodies.

The effect of immunogenicity on PK, PDx, efficacy and safety will be evaluated if data allow..

4.2.14 PK/PDx relationships (MEDI4736 monotherapy and MEDI4736+tremelimumab)

If the data are suitable, the relationship between PK exposure and efficacy/safety parameters may be investigated graphically or using an appropriate data modelling approach. These outputs will be produced by AstraZeneca/MedImmune Clinical Pharmacology group or designee.

4.2.15 Biomarker data

If applicable, the relationship of exploratory biomarkers to OS, PFS, ORR and DoR will be presented for a subset of patients in the ITT population who are evaluable for each biomarker.

This will be assessed using similar summary and graphical representations to those that are outlined for the efficacy outputs.

PD-L1 expression determined by immunohistochemistry 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.

4.2.16 Demographic and baseline characteristics data

The following will be summarised for all patients in the FAS, PD-L1-positive$_{25\%}$ analysis set and PD-L1-positive$_{1\%}$ analysis set by treatment group:-

- Patient disposition (including screening failures and reason for screening failure)
- Important protocol deviations
- Inclusion in analysis populations
- Demographics (age, age group[<50, >=50-< 65, ≥ 65 - <75 years and ≥ 75 years], sex, race and ethnicity)
- Patient characteristics at baseline (height, weight, weight group, body mass index (BMI) and body mass index group)
- Patient recruitment by country and centre

- Previous disease-related treatment modalities
- Number of regimens of previous chemotherapy at baseline
- Previous lung cancer therapy
- Disease characteristics at baseline / diagnosis (WHO performance status, primary tumour location, histology type, tumour grade and overall disease classification, best response to previous therapy)
- Extent of disease at baseline
- TNM classification
- Disease related medical history (past and current)
- Relevant surgical history
- Physical examination at baseline
- Time from most recent disease progression to start of study treatment
- Disallowed concomitant medications
- Allowed concomitant medications
- Post-discontinuation cancer therapy
- Nicotine use, categorised (never, current, former)
- Stratification factors as per IVRS and eCRF data

The AZ drug dictionary (AZDD) will be used for concomitant medication coding.

Patient disposition data will also be summarised at the time of OS analysis.

4.2.17 Treatment exposure

The following summaries related to study treatment will be produced for the safety analysis set, PD-L1-positive$_{25\%}$ safety analysis set and PD-L1-positive$_{1\%}$ safety analysis set by randomised treatment group:

- Total exposure of each treatment group.
- Actual exposure of each treatment group.
- Total number of cycles received.

- Reasons for dose delays and infusion interruptions of MEDI4736 and tremelimumab and reasons for dose delays/interruptions, dose reductions and dose modifications for the Standard of Care agents. Dose interruptions will be based on investigator initiated dosing decisions.

- Number of dose delays and duration of delays of MEDI4736. In addition, delays due to AEs and due to reasons other than AEs will be summarized separately.

- Number of infusions received.

- RDI (relative dose intensity) of MEDI4736, tremelimumab and Standard of Care agents.

- Exposure over time will be plotted.

For patients on study treatment at the time of the PFS and OS analysis, the DCO date will be used to calculate exposure.

4.2.18 Subsequent Therapy

Subsequent therapies received after discontinuation of study treatment will have summaries produced by treatment group, together with number of regimens received.

5. INTERIM ANALYSES

5.1 Analysis Methods

Interim safety monitoring will be conducted by an IDMC. Interim analyses will be performed for efficacy as described below:

Two OS interim analyses will be performed for superiority; the first one at the time of the primary PFS analysis and the second one when approximately 80% of the final number of deaths has been reached. These analyses will be performed by an IDMC.

The Lan DeMets spending function that approximates an O'Brien Fleming approach will be used to account for multiplicity introduced by including the 2 interim analyses for superiority (Lan and DeMets 1983).

The criterion for superiority in the OS co-primary endpoint is a statistically significant improvement in OS at the interim analysis. The alpha level that is to be spent at the interim and final analyses for the OS analyses in the PD-L1-positive$_{25\%}$ population will be calculated using the Lan DeMets spending function separately. If exactly 68% and 80% of the target events in the PD-L1-positive$_{25\%}$ patients are available at the time of the first and second interim analyses, respectively, that is, 152/225 and 180/225 events have occurred, with overall 2-sided alpha levels of 0.015 and 0.03 respectively for the comparisons of MEDI4736 +tremelimumab versus SoC and MEDI4736 monotherapy versus SoC the 2-sided alpha level

to be applied for the interim and the final analysis would be 0.0023, 0.0049 and 0.0132 for the comparison of MEDI4736 +tremelimumab versus SoC and 0.0062, 0.0113 and 0.0258 for the comparison of MEDI4736 monotherapy versus SoC,

If the interim analyses indicate superiority in the PD-L1-positive$_{25\%}$ population, then subsequent analyses of the further secondary endpoints will be performed in accordance with the hierarchical multiple testing strategy.

If the interim results do not meet the criterion of stopping for superiority in the PD-L1-positive$_{25\%}$ population, then follow-up will continue until at least 225 deaths have occurred in the PD-L1-positive$_{25\%}$ population. OS will be retested in both the PD-L1-positive$_{25\%}$ population at the final analysis. Similarly if the criteria for statistical significance at the interim is met for PD-L1-positive$_{25\%}$ but not PD-L1-positive$_{1\%}$ then follow-up will continue until 337 events have been reached in this population after which the OS will be retested.

The recommendations from the IDMC will not reveal the results of the analyses but will take the form of "Continue/Modify/Recommend early submission/Stop."

Details of the IDMC plan and communication process is provided in the IDMC Charter

5.2 Independent Data Monitoring Committee

This study will use an external IDMC to assess ongoing safety analyses, and to perform the formal interim analyses of OS. The committee will meet approximately 6 months after the study has started or after the first 30 patients have been randomized, whichever occurs first, to review the safety data from the study. The IDMC will meet at least every 6 months thereafter. Following each meeting, the IDMC will report to the sponsor and may recommend changes in the conduct of the study.

This committee will be composed of therapeutic area experts and biostatisticians, who are not employed by AstraZeneca/MedImmune and do not have any major conflict of interest.

Following the reviews, the IDMC will recommend whether the study should continue unchanged, be stopped, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to AstraZeneca/MedImmune. The report will include the recommendation and any potential protocol amendments, and will not contain any unblinding information.

The final decision to modify or stop the study will sit with the sponsor. The sponsor or IDMC may call additional meetings if at any time there is concern about the safety of the study.

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

The safety of all AstraZeneca/MedImmune clinical studies is closely monitored on an ongoing basis by AstraZeneca/MedImmune representatives in consultation with the Patient Safety Department. Issues identified will be addressed; this could involve, for instance, amendments to the clinical study protocol and letters to investigators.

6. CHANGES OF ANALYSIS FROM PROTOCOL

None from protocol version 7.

7. REFERENCES

Burman et al 2009
Burman CF, Sonesson C, Guilbaud O. A recycling framework for the construction of Bonferroni-based multiple tests. Stat Med 2009;28:739-61.

Ciuleanu et al 2009
Ciuleanu T, Brodowicz T, Zielinski C, Kim JH, Krzakowski M, Laack E, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomized, double-blind, phase 3 study. Lancet 2009;374(9699):1432-40.

Cox 1972
Cox DR. Regression models and life-tables. J Royal Stat Society 1972;Series B 34(2):187-220.

Ellis et al 2008
Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials 2008;29(4):456-65.

Fayers et al 2001
Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A; EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd ed. Brussels: European Organisation for Research and Treatment of Cancer: 2001.

Fehrenbacher et al 2016
Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, Phase 2, randomised, controlled trial. Lancet 2016; published online 09 March 2016.

Fleischer et al 2011
Fleischer F, Gaschler-Markefski B, Bluhmki E. How is retrospective independent review influenced by investigator-introduced informative censoring: a quantitative approach. Stat Med 2011;30(29):3373-86

Gail and Simon 1985
Gail M, Simon R. Tests for qualitative interactions between treatment effects and patient subsets. Biometrics 1985;41(2):361–72.

Holm 1979
Holm S. A simple sequentially rejective multiple test procedure. Scand J Statistics 1979;6:65-70.

Klein et al 2007
Klein JP, Logan B, Harhoff M, Andersen PK. Analyzing survival curves at a fixed point in time. Stat Med 2007;26(24):4505-19.

Lan and DeMets 1983
Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983; 70 (3):659-663

Oemar and Janssen 2013
Oemar M, Oppe M. EQ-5D-3L User Guide: Basic information on how to use the EQ-5D-3L instrument V5.0 (October 2013). Available from: URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2013_v5.0_October_2013.pdf. Accessed 07 January 2014.

Osoba et al 1998
Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16(1):139-44.

Paz-Ares et al 2013
Paz-Ares LG, de Marinis F, Dediu M, Thomas M, Pujol JL, Bidoli P, et al. PARAMOUNT: Final overall survival results of the phase III study of maintenance pemetrexed versus placebo immediately after induction treatment with pemetrexed plus cisplatin for advanced nonsquamous non-small-cell lung cancer. J Clin Oncol 2013;31(23):2895-902.

Pintilie 2006.
Competing risks: A practical perspective. Wiley, 2006.

Robins 1993
Robins JM. Information Recovery and Bias Adjustment in Proportional Hazards Regression Analysis of Randomized Trials Using Surrogate Markers. Proceedings of the Biopharmaceutical Section, American Statistical Association 1993; 24-33.

Robins and Tsiatis 1991
Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics-Theory and Methods 1991; 20(8):2609-31.

Scagliotti et al 2008
Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008;26(21)3543-51.

Sun and Chen 2010
Sun X, Chen C. Comparison of Finkelstein's Method with the conventional approach for interval-censored data analysis. Stat Biopharm Res 2010;2(1):97-108.

Whitehead and Whitehead 1991
Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10(11):1665-77.

8. APPENDIX

None.

Changes to planned analyses

Key details of change	Reason for change	Responsible for change
Changes made before database lock		
Study protocol amendments and program decisions		
Clarifications in the objectives section to note inclusion of multiple primary endpoints, use of BICR for PFS, and updates to set of secondary objectives	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Study design section updated to reflect increase in sample size and allowing treatment to progression for both combination and monotherapy treatment groups	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Changes in the Primary and Secondary variable section to reflect use of multiple primary endpoints and use of BICR	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Sample size section updated to reflect increase in sample size based on changes to the treatment comparisons of interest and inclusion of 2 formal interim analysis for OS endpoint	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Update of the Analysis Methods section to reflect use of multiple primary endpoints, the updates to secondary endpoints, and requisite updates to the multiple testing strategy	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Changes to the Analysis Methods section to note use of Cox proportional hazards for calculation of HR and confidence interval and addition of additional analysis set for PD-L1 status (using 10% cut-off for positive/negative split)	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Interim analysis for Overall Survival was added	Alignment with protocol amendments, versions 3 to 5	AstraZeneca
Changes introduced in SAP Edition 3[a]		
Clarifications in Objectives section to note inclusion of additional primary endpoint in OS and additional populations for the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations in the primary and secondary objectives	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Study design section updated to reflect the introduction of the additional populations for the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations in the revised primary and secondary objectives	Alignment with protocol amendments, versions 6 and 7	AstraZeneca

Key details of change	Reason for change	Responsible for change
Update to the definitions of the analysis sets due to introduction of the additional populations for the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations in the revised primary and secondary objectives	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Confirmation in Biomarker Variables section on the definition of the various PD-L1 expression analysis sets	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Update of Analysis Methods section to reflect use of the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations for the revised primary and secondary objectives/endpoints and requisite updates to the multiple testing strategy	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Update to analysis of primary endpoint PFS to reflect the primary and secondary endpoints based on PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations and the changes to relevant confidence intervals due to the revised multiple testing strategy	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Analysis sets updated to include PD-L1 tumor cell expression for the additional cut-off points of 1%, 10%, and 50%, plus inclusion of 25% cut-off for immune cell expression	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Update to analysis of primary endpoint OS to reflect the primary and secondary endpoints based on PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations and the changes to relevant confidence intervals due to the revised multiple testing strategy	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Update to analysis of all secondary endpoints to reflect the secondary objectives based on PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Analyses of safety data updated to include appropriate analyses based on the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations	Alignment with protocol amendments, versions 6 and 7	AstraZeneca
Interim analysis details updated to reflect the change in primary objectives now based on the PD-L1 (TC ≥25%) population as primary		
Changes introduced in SAP Edition 4[a]		
Update to the definition of corresponding safety analysis set in PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations	Consistent with program decisions	AstraZeneca
Update to Table 2 of SAP on the analysis of the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations	Consistent with program decisions	AstraZeneca
Updates to study population and demographic data will be summarized in Full Analysis Set and Full Analysis Set for patients within the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations, correspondingly	Consistent with program decisions	AstraZeneca

Key details of change	Reason for change	Responsible for change
Updates to safety analyses will be performed in the safety analysis set and safety analysis sets for patients within the PD-L1 (TC ≥25%) and PD-L1 (TC ≥1%) populations, correspondingly	Consistent with program decisions	AstraZeneca
Updates to PFS/OS/ORR additional analyses	Consistent with program decisions	AstraZeneca
Clarification that PFS subgroup analysis will also be performed for the comparison between durvalumab monotherapy versus SoC chemotherapy	Consistent with program decisions	AstraZeneca
Updates to the PFS multivariate Cox regression model to evaluate the treatment effect adjusting for any potential imbalances in baseline prognostic factors	Consistent with program decisions	AstraZeneca
Clarification that OS subgroup analysis will also be performed for the comparison between durvalumab monotherapy versus SoC chemotherapy	Consistent with program decisions	AstraZeneca
Updates to additional analyses using Cox proportional hazards models to determine the effect of covariates on the OS HR estimate	Consistent with program decisions	AstraZeneca
Updates to additional analyses using Cox proportional hazards models to determine the consistency of OS treatment effect	Consistent with program decisions	AstraZeneca
Updates to ORR subgroup analysis	Consistent with program decisions	AstraZeneca
Changes introduced in SAP Edition 5[a]		
Clarification of the interim efficacy boundary calculation for the OS comparison of durvalumab + tremelimumab versus SoC in PD-L1 (TC ≥1%) population	Consistent with program decisions	AstraZeneca
Changes introduced in SAP Edition 5[a] Errata		
Clarification that FAS includes all randomized patients, regardless of whether a patient received study therapy or had any important protocol deviation, consistent with the intent-to-treat principle	To correct errors noted in the SAP	AstraZenecac
Important protocol deviation 2 to also include inclusion criteria 4	To correct errors noted in the SAP	AstraZeneca
Terminology for causally related AEs updated to “possibly related” AEs	To correct errors noted in the SAP	AstraZeneca

Key details of change	Reason for change	Responsible for change
Changes made after database lock		
Exploratory analyses were performed on blood TMB and tissue TMB status.	To align with emerging interest in TMB as a biomarker and assess the effect of efficacy and safety by TMB status	

[a]

BICR Blinded Independent Central Review; HR hazard ratio; ORR overall response rate; OS overall survival; PD-L1 programmed cell death ligand 1; PFS progression-free survival; SAP Statistical Analysis Plan; SoC standard of care; TMB tumor mutational burden.