STATISTICAL ANALYSIS PLAN

Study Protocol Number: E7080-M000-213

Study Protocol Title: An Open-Label, Single-Arm, Multicenter, Phase 2 Trial of Lenvatinib for the Treatment of Anaplastic Thyroid Cancer (ATC)

Date: 18 Oct 2018

Version: Final Version 1.0

1   TABLE OF CONTENTS

2   LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	anaplastic thyroid cancer
ß-hCG	beta-human chorionic gonadotropin
BOR	best overall response
BP	blood pressure
CBR	clinical benefit rate
CI	confidence interval
CR	complete response
CRF	case report form
CSR	clinical study report
CT	computerized tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCR	disease control rate
DOR	duration of response
EAS	evaluable analysis set
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
FAS	full analysis set
FDA	Food and Drug Administration
ICF	Informed consent form
INR	International Normalized Ratio
KM	Kaplan-Meier
MedDRA®	Medical Dictionary for Regulatory Activities
MRI	magnetic resonance imaging
NCI	National Cancer Institute
NYHA	New York Heart Association

Abbreviation	Term
ORR	objective response rate
OS	overall survival
PE	physical examinations
PET	position emission tomography
PFS	progression-free survival
PI	principal investigator
PR	partial response
PRES	posterior reversible encephalopathy syndrome
PT	preferred term
QD	once a day
RBC	red blood cell count
RECIST	Response Evaluation Criteria in Solid Tumors
RR-DTC	radioiodine-refractory differentiated thyroid cancer
RTK	receptor tyrosine kinase
SAE	serious adverse events
SAP	statistical analysis plan
SD	stable disease
SmPC	Summary of Product Characteristics
SOC	system organ class
SUSAR	suspected unexpected serious adverse reactions
T4	thyroxine
TEAE	treatment-emergent adverse event
TSH	thyroid stimulating hormone
TTP	time to progression
ULN	upper limit of normal
US	United States
VEGF	vascular endothelial growth factor

3   INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for Eisai Protocol E7080-M000-213, an Open-Label, Single-Arm, Multicenter, Phase 2 Trial of Lenvatinib for the Treatment of Anaplastic Thyroid Cancer (ATC). The focus of this SAP is on all the analyses specified in the final protocol of the study (i.e., Version 3.0 dated on 24May2016, Amendment 02), which covers the both interim analyses and final analyses. This SAP will be finalized before the database lock of study completion.

3.1   Study Objectives

3.1.1   Primary Objective

To evaluate objective response rate ([ORR]: complete response [CR] or partial response [PR]) by investigator review in subjects with anaplastic thyroid cancer (ATC) treated with lenvatinib.

3.1.2   Secondary Objectives

- To evaluate 12-week progression-free survival (PFS)
- To evaluate 6-month overall survival (OS)
- To evaluate median PFS and median OS
- To evaluate safety and tolerability of lenvatinib in subjects with ATC

3.1.3   Exploratory Objectives

- To explore clinical benefit rate ([CBR]: CR, PR or durable stable disease [SD] $\geq$ 23 weeks)
- To explore disease control rate ([DCR]: CR, PR or SD)
- To explore duration of response (DOR)
- To identify and explore tumor and blood biomarkers that correlate with clinical outcomes, including efficacy

3.2   Overall Study Design and Plan

This was an open-label, single arm, multicenter, Phase 2 trial of lenvatinib (24 mg/day) to assess the efficacy and safety of lenvatinib in subjects with ATC. The study was designed primarily to provide additional efficacy and safety data for single-agent lenvatinib in a larger cohort of ATC subjects. Approximately a total of 76 subjects was planned to be enrolled into this study (i.e., signed informed consent) to be sure that out of the 76 subjects there would be 57 subjects confirmed with ATC for the purpose of efficacy analyses. Also one interim analysis was planned to be performed after the first 20 evaluable subjects completed at least 2 tumor assessments or discontinue treatment due to any reason. While waiting for the

interim analysis results the enrollment would not be suspended so that no enrollment gap was in the study for the interim analysis. From the interim analysis the enrollment would be halted if the number of responders was 3 or less (ORR ≤15%, according to Response Evaluation Criteria in Solid Tumors (RECISIT) 1.1), and the safety and efficacy data would be further evaluated before making the decision of stopping the study permanently.

This study would consist of the following three phases and the corresponding procedures/assessments to each was presented as Table 4 in the final protocol of the study in details, where as shown a 28-days during the treatment period was considered as one cycle of treatment and used for study related visits/assessments:

- The **Pretreatment Phase:** including Screening period from Day -21 prior to the first study treatment of lenvatinib start. All analyses related baseline assessments would be defined as the last assessment obtained during this phase for each.

- The **Treatment Phase:** starting from the time of the first lenvatinib administration till the end of last lenvatinib administration. Subjects would be treated with lenvatinib dose of 24 mg once daily by oral administration, or dose following protocol on treatment adjustment.

- The **Posttreatment** Phase: starting after the last study dose, including End of Treatment visit (i.e., within 30 days of last dose), survival follow-ups (i.e., every 12 (±1) weeks after End of Treatment visit).

Treated subjects would discontinue treatment with lenvatinib until the time of disease progression (radiological or clinical), development of unacceptable toxicity, withdrawal of consent, lost to follow-up, pregnancy or study termination by the sponsor.

During the Follow-Up Period, in subjects who had discontinued study treatment without documented disease progression (radiological or clinical), every effort would be made to continue monitoring their disease status by radiologic imaging every 6 weeks ±1 week for the first 24 weeks and every 8 weeks ±1 week thereafter from the date of the last assessment until (1) the start of new anticancer treatment, (2) disease progression (3) death, or (4) the end of the study, whichever occurred first.

All subjects would be followed for survival until death, except where a subject withdrew consent or the sponsor chose to halt survival follow-up after completion of the primary study analysis.

4   DETERMINATION OF SAMPLE SIZE

The sample size calculation was based on the assumed objective response rate (ORR), defined in Section 5.1) of 27% in the trial as compared to 10% in the historical control. Using a binomial exact text, the power was 0.932 with 57 evaluable subjects to demonstrate statistical significance at a 1-sided alpha of 0.025. To evaluate the power in the secondary endpoint 12-week PFS (defined in Section 5.1), an exponential distribution assuming 12-week PFS being 70% was used to simulate PFS data. Censoring times were generated such that the average number of observed PFS events was approximately 45 in the 57 evaluable subjects. In the 5,000 simulation runs, the lower bound of 95% confidence interval for PFS at 12 weeks was greater than 50% (historical control), 87.7% of the time. Hence, the power was 0.877 in demonstrating statistical superiority of the 12-week PFS in single agent lenvatinib. Similarly, approximately 35 observed OS events in 57 evaluable subjects yielded a power of approximately 0.857 in demonstrating a statistically significant increase in 6-month OS over historical control. Assuming ATC confirmation rate was 75% in the enrolled subjects, approximately 76 subjects would need to be enrolled. Subjects deemed to have another diagnosis (not ATC) would be replaced for the purpose of efficacy analyses.

5   STATISTICAL METHODS

All descriptive statistics for continuous variables will be reported in mean, standard deviation (SD), median, 25 percentile (Q1), 75 percentile (Q3), minimum and maximum. Categorical variables will be summarized in number (percentage) of subjects.

5.1   Study Endpoints

All tumor related endpoints, including timepoint tumor responses and the best overall responses (i.e., complete response, partial response, stable disease, disease progression), are from investigators based on RECIST 1.1.

5.1.1   Primary Endpoint

The primary efficacy endpoint is ORR as determined by investigator review, using RECIST 1.1. ORR is the proportion of subjects who have best overall response (BOR) of CR or PR (confirmed). Confirmation of CR or PR must be performed at least 28 days following the initial achievement of the response.

5.1.2   Secondary Endpoints

- **Twelve-week PFS** is the percentage of subjects in the analysis population who remain alive and disease progression-free at 12 weeks post first study dose. This will be the Kaplan-Meier estimated rate of event (death or disease progression)-free at 12 weeks post first study dose (i.e., at the 84th day counting from the first dose date),

- **Six-month OS** is defined as the percentage of subjects in the analysis population who are alive at 6 months post first study dose. This will be the Kaplan-Meier estimated

rate of event (death)-free at 6 months post first study dose (i.e., at the 183th day counting from the first dose date, from 6 x 365.25/12),

- **Median PFS and Median OS.** These will be the estimated median event-free time from Kaplan-Meier method (in day and month). The PFS is defined as the time from the date of first dose of study drug to date of first documentation of disease progression (only from the radiological assessment) or death, whichever occurs first, and OS is defined as the time from the date of first dose of study drug until date of death from any cause.

5.1.3   Exploratory Endpoints

- **Clinical benefit rate** (CBR) is the proportion of subjects who have the best overall response (confirmed) of CR or PR or durable SD. A stable disease (SD) must be at $\geq$ 5 weeks after the first study dose. A durable SD is a subset of SD with the duration of $\geq$ 23 weeks after the first study dose.
- **Disease control rate** (DCR) is the proportion of subjects who have the best overall response (confirmed) of CR, PR or SD.
- **Duration of response** (DOR) is defined as the time from the first objective response (i.e., the initial CR or PR that was confirmed) to the first documented disease progression or death. This endpoint is just defined for the subjects who had confirmed response during study.
- Evaluate the association of tumor and blood biomarkers with clinical outcomes, including efficacy.

5.2   Study Subjects

5.2.1   Definitions of Analysis Sets

The Full Analysis Set [FAS] will include all subjects who received at least one dose of lenvatinib. The FAS will be the population for the safety analysis. Some selected efficacy endpoints will also be summarized in the FAS.

The Evaluable Analysis Set [EAS] will include FAS subjects with histological diagnosis of ATC that was confirmed by central pathology review. The EAS will be the primary population for the efficacy analysis.

5.2.2   Subject Disposition

The number (percentage) of enrolled (i.e., signed informed consent) and treated subjects will be summarized as well as reasons for discontinuation from study treatment. Survival status at the last follow-up visit, and reasons for screen failure, will also be summarized.

5.2.3   Protocol Deviations

Major protocol deviation criteria will be established and subjects with major protocol deviations will be identified and documented before the database lock for the primary analysis.

Major protocol deviations will be summarized and listed by each category. All protocol deviations identified according to the study entry criteria and during treatment will be listed.

5.2.4   Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics for both the FAS and EAS will be summarized for lenvatinib treatment, using descriptive statistics. Continuous demographic and baseline variables include age, BMI, weight, and height; categorical variables include sex, age group (<65 years, 65-75, >75 years), region, ethnicity and race. Other variables at study entry include: pregnancy test, ECOG performance status, time since disease progression, prior radiotherapy (yes/no), ATC diagnosis confirmation by central pathology review (yes/no), TNM classification (T-Primary Tumor, N-Regional Lymph Nodes, M-Distant Metastasis), Anaplastic Thyroid Cancer Stage (IVA, IVB, IVC), time from original histological diagnosis to first dose, NYHA, number of target lesions and sum of target lesion diameters.

MEDICAL HISTORY

The number (percentage) of subjects in the FAS reporting a history of any medical condition and current medical condition, as recorded on the CRF, will be summarized. A subject data listing will be provided. Medical History will be coded using MedDRA (version 21.0 or newer), and summarized on the FAS by System Organ Class and Preferred term.

5.2.5   Prior and Concomitant Therapy

All investigator terms (verbatim terms) for medications recorded in the CRF will be coded to an 11-digit code using the World Health Organization Drug Dictionary (WHO DD) (WHODDMAR2018). The number (percentage) of subjects who took prior and concomitant medications will be summarized in FAS by Anatomical Therapeutic Chemical (ATC) class, and WHO DD preferred term.

Prior medications are defined as medications that stopped before the first dose of study drug. Concomitant medications are defined as medications that (1) started before the first dose of study drug and were continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug up to 28 days after the subject's last dose. All medications will be presented in subject data listings.

Frequencies [n (%)] of previous anti-cancer procedures will be summarized similarly by system organ class and preferred term. In addition, previous anti-cancer medications will be summarized in terms of number of prior regimens, duration of last therapy, best response for last therapy, time from end of last therapy to first dose, and indication of previous medication

(neoadjuvant/adjuvant/metastatic/unknown). The following data on radiotherapy will be summarized: time from last radiotherapy to first dose, site of previous radiotherapy, progression of tumor lesion at the site since radiotherapy (yes/no), and number of all prior radiotherapy treatments.

5.2.6   Treatment Compliance

Records of treatment compliance for each subject will be kept during the study. Clinical research associates will review treatment compliance during investigational site visits and at the completion of the study. Treatment compliance will not be summarized since the data will not be entered in the clinical database.

5.3   Data Analysis General Considerations

All efficacy analyses will be summarized with available data in the clinical database at the data cutoff.

5.3.1   Pooling of Centers

Subjects from all centers will be pooled for all analyses.

5.3.2   Adjustments for Covariates

No adjustments for covariates will be performed in the study.

5.3.3   Multiple Comparisons/Multiplicity

Per protocol, the type I error calculation for the primary efficacy analysis did not factor in the descriptive interim analysis, the interim decision won't affect the overall type I error in the formal primary analysis; no multiplicity adjustments for the tests of secondary endpoints will be made in this study.

5.3.4   Examination of Subgroups

No examination of subgroups is planned.

5.3.5   Handling of Missing Data, Dropouts, and Outliers

Rules of handling missing or partial missing dates are explained in Section 8. The censoring rules used in time-to-event analyses (PFS, OS, DOR) are also explained in Section 8. Potential outlier values will be investigated, and they will be analyzed as originally reported in the locked database.

In the analyses on the endpoints described above and any sensitivity analysis, subjects with missing data will be considered as non-responders and will be included in the denominator when calculating ORR, DCR, and CBR. Non-responders will be excluded in the analysis of DOR.

5.3.6   Other Considerations

Not applicable.

5.4   Efficacy Analyses

5.4.1   Primary Efficacy Analyses

In the study design, the ORR was estimated as 27% (deemed a clinical meaningful improvement), and its historical control on recent trials was assumed to be 10%. The study primary analysis was planned to be performed for the observed ORR using the binomial exact test, on evaluable analysis set, and based on the following hypotheses

> H0: ORR=10%

> Ha: ORR $\geq$ 27%.

And in the plan if the obtained p-value from the binomial exact test was less than or equal to 0.025 (one-sided), it would be concluded that the superiority of single agent lenvatinib was demonstrated.

The 95% Clopper–Pearson confidence interval (95% CI) of ORR will be presented. The same primary efficacy analysis for ORR for evaluable analysis set will also be repeated for the full analysis set.

5.4.2   Secondary Efficacy Analyses

All analyses of secondary efficacy endpoints will be performed in FAS.

Progression-Free Survival (PFS) and Overall Survival (OS) will be analyzed using Kaplan-Meier product-limit method. The Kaplan-Meier estimated rate of event-free will be presented at 12 weeks for PFS and at 6 months for OS, together with the corresponding two-sided 95% CI. In addition, the median, first and third quartiles of the Kaplan-Meier estimated event-free time will also be presented together with the corresponding 95% Brookmeyer-Crowley (or generalized Brookmeyer-Crowley) CI.

5.4.3   Other Efficacy Analyses

All the efficacy exploratory analyses will be performed in both evaluable analysis set and/or full analysis set.

The observed clinical benefit rate and the disease control rate will be presented along with their 95% CI of exact method.

Similarly, using the same method as the analyses to PFS and OS, the Duration of Response (DOR) will be analyzed using Kaplan-Meier method and the estimated median, first and third quartiles of the DOR will be presented with the 95% CI.

5.5   Pharmacokinetic, Pharmacodynamic, Pharmacogenomic, and Other Biomarker Analyses

Not applicable.

5.6   Safety Analyses

All safety analyses will be performed based on Full Analysis Set.

5.6.1   Extent of Exposure

The analyses of extent of drug exposure will be performed for each subject on the duration of treatment, and total cumulative dose received (mg) (i.e., the sum of all doses actually taken), actual dose intensity (mg/day) received, and the relative dose intensity (%), and the definitions are:

- Duration of treatment (day) is calculated as (Date of last study drug – Date of first study drug + 1)

- Actual dose intensity (mg/day) is calculated as total cumulative dose received (mg) divided by treatment duration (day)

- Relative dose intensity (%) is calculated as actual dose intensity (mg/day) divided by 24 (mg/day).

For the data from the CRF page of Study Medication collected during treatment phase, the number (%) of subjects with dose reduced, dose interruption, and the time to the first dose reduction will be summarized.

5.6.2   Adverse Events

The adverse event (AE) verbatim descriptions (investigator terms from the CRF) will be classified into standardized medical terminology using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse events will be coded to the MedDRA (Version 21.0 or higher) lower level term (LLT) closest to the verbatim term. The linked MedDRA preferred term (PT) and primary system organ class (SOC) will also be captured in the database. All AEs will also be graded by the investigators using CTCAE v4.03.

A treatment-emergent adverse event (TEAE) is defined as an AE that emerged during treatment (up to 28 days post the last study dose), having been absent at pretreatment (Baseline) or

- Reemerged during treatment (up to 28 days post the last study dose), having been present at pretreatment (Baseline) but stopped before treatment, or

- Worsened in severity during treatment (up to 28 days post the last study dose) relative to the pretreatment state, when the AE was continuous.

Only those AEs that are treatment emergent will be included in summary tables. All AEs, treatment emergent or otherwise, will be presented in subject data listings.

The incidence of TEAEs will be reported as the number (percentage) of subjects with TEAEs by SOC and PT. A subject will be counted only once within a SOC and PT, even if the subject experienced more than 1 TEAE within a specific SOC and PT. The number (percentage) of subjects with TEAEs will also be summarized by highest CTCAE grade.

The number (percentage) of subjects with treatment-related TEAEs will also be summarized by SOC and PT. Treatment-related TEAEs include those events considered by the investigator to be related to study treatment or with missing assessment of the causal relationship.

The following TEAE summary analyses will be performed:

- An overall summary table will include the number (%) of subjects with TEAEs, with treatment-related TEAEs, with TEAEs with CTCAE grade ≥3, with severe TEAEs (grade 3 or 4), with serious TEAEs, treatment-related serious TEAEs, TEAEs leading to death, TEAEs leading to dose withdrawn/dose reduced/dose interrupted.

- TEAEs will be summarized by SOC, PT and grade in number (%) of patients. Similarly, TEAEs that are serious, treatment-related, grade 3 or above, serious and related, leading to death/dose withdrawn/dose reduced/dose interrupted will also be summarized by SOC, PT and grade.

- TEAEs will be summarized by PT in number (%) of subjects.

- Additional summary on the TEAE of special interested will be presented by special interest, preferred term, and the worst CTCAE grade.

5.6.2.1   Deaths

An overall summary of deaths will be presented: deaths within 28 days post last study drug, and deaths after 28 days post last study drug, and treatment-related deaths. A listing of all deaths will be provided.

5.6.3   Laboratory Values

Per protocol, the clinical laboratory tests in Table 1 were performed in the study. Abnormal laboratory values will be flagged and identified as those outside (above or below) the normal range, laboratory values will also be categorized by CTCAE v4.03 grades per Appendix 7, Sponsor's Grading for Laboratory Values, of the study protocol. Laboratory values will be summarized by visit and by worst post-baseline visit. Laboratory parameters will be summarized by worst grade, and laboratory parameters with CTCAE grading in both high and low directions will be summarized separately. Shift tables from baseline grade to worst value will be presented.

Table 1   Clinical Laboratory Tests

Category	Parameters
Hematology	Hematocrit, hemoglobin, platelets, RBC count, and WBC count with differential (bands, basophils, eosinophils, lymphocytes, monocytes, neutrophils), ANC, INR[b]
Chemistry	
Electrolytes	Bicarbonate, chloride, magnesium, potassium, sodium
Liver function tests	Alanine aminotransferase, alkaline phosphatase, aspartate aminotransferase, direct bilirubin, total bilirubin
Renal function tests	Blood urea/blood urea nitrogen, creatinine
Thyroid function tests[a]	TSH and free T4
Pregnancy test	Serum β-hCG (if urine not tested)
Other	Albumin, glucose[c], lactate dehydrogenase, total protein, uric acid, calcium, phosphorus
Urinalysis	Glucose[c], ketones, pH, protein, RBCs, specific gravity
Pregnancy test	Urine β-hCG (if serum not tested)

ALT = alanine aminotransferase, ANC= absolute neutrophil count, AST = aspartate aminotransferase, β hCG = beta-human chorionic gonadotropin, INR = International Normalized Ratio, RBC = red blood cells, T4 = thyroxine, TSH = thyroid stimulating hormone, WBC = white blood cells

a: Thyroid function will be assessed at the Screening Visit and then every 2 cycles (starting at C2) throughout the study.

b: INR only at screening/baseline and when clinically indicated

c: Fasting glucose at Screen, only

5.6.4   Vital Signs

Descriptive statistics for vital signs parameters (diastolic and systolic blood pressure, pulse, respiration rate, temperature, weight) and changes from baseline will be presented by visit.

5.6.5   Electrocardiograms

Descriptive statistics for ECG parameters and changes from baseline will be presented by visit.

Shift tables will present changes from baseline in ECG (categorized as normal; abnormal, not clinically significant; and abnormal, clinically significant) to the worst post-baseline (during treatment period). Other ECG values (Heart rate, QT Interval) and the changes from baseline will be summarized and presented by visit.

5.6.6   Other Safety Analyses

No other safety analyses are planned.

5.7 Other Analyses

No other analysis is planned.

5.8 Exploratory Analyses

Additional exploratory analyses may be conducted as appropriate. Any exploratory analyses to be performed will be appropriately titled/labeled as exploratory and will be clearly distinguished from planned analyses when results are reported in the Clinical Study Report.

6 INTERIM ANALYSES

Per study protocol, one interim analysis was planned to be performed after the first 20 evaluable subjects complete at least 2 tumor assessments or discontinue treatment due to any reason. From the interim analysis the enrollment would be halted if the number of responders is 3 or less (ORR ≤15%), and the safety and efficacy data would be further evaluated before making the decision of stopping the study permanently.

The study team would define the datasets for the interim analysis. The study team would assess the impact of any issues such as unsolved data queries, missing data or data entry delays on the interim results. The study statistician would perform the interim analysis, which would include calculating ORR per protocol.

7 CHANGES IN THE PLANNED ANALYSES

Treatment duration will be analyzed in months instead of nominal cycles mentioned in the protocol. The exploratory analyses of identifying and exploring tumor and blood biomarkers that correlate with clinical outcomes, including efficacy will not be performed.

8 DEFINITIONS AND CONVENTIONS FOR DATA HANDLING

Baseline: in all analyses, the baseline value is defined as the non-missing value most recently collected before the first dose.

Study day calculation: unless otherwise specified in outputs, the following convention will be used to calculate the study day used in the listings:

- Day 1 is the day of the first study drug taken

- Study Day (for an event/assessment, relative to the day of the first study drug taken) is defined as (date of assessment/event – date of the first study drug taken +1) for the event/assessment on or after the date of the first study drug, and as (date of assessment/event date of the first study drug) for the event/assessment before the first study drug.

Censoring rule for OS: Subjects who were lost to follow-up, or withdrew consent, will be censored at the last date the subject was known to be alive, otherwise the censoring date will be the data cutoff date.

Censoring rule for PFS: Table 2 shows the censoring rules for subjects without documented PD or death in the derivation of PFS based upon investigator's tumor assessment.

Censoring rule for duration of response (DOR): considering the subjects with the confirmed best overall response of CR or PR and still alive without disease progression following the objective response. The same censoring rule of PFS will be applied to the derivation of DOR for the cases applicable.

Table 2 Censoring Rules for Analysis of Progression-Free Survival

No.	Situation	Date of Progression or Censoring	Outcome
1	No baseline tumor assessments	Date of first dose	Censored
2	Progression documented between scheduled visits	Date of first radiological PD assessment	Progressed
3	No progression at the time of data cut-off or withdrawal from study	Date of last adequate radiologic assessment prior to or on date of data cut-off or withdrawal from study, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored
4	New anticancer treatment started	Date of last adequate radiologic assessment prior to or on date of new anticancer treatment, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored
5	Death before first PD assessment	Date of death	Progressed
6	Death between adequate assessment visits*	Date of death	Progressed
7	Death or progression after two or more missed visits**	Date of last adequate radiologic assessment before missed tumor assessments, or date of the first dose if no adequate post baseline tumor assessment was conducted.	Censored
8	No post-baseline tumor assessments but known alive	Date of first dose	Censored

CR = complete response, PD = progressive disease, PR = partial response, SD = stable disease,

* Adequate tumor assessment is a radiologic assessment of CR, PR, SD, non-CR/non-PD or PD as determined by investigators.

New anticancer treatment includes palliative radiotherapy. Any tumor assessments after new anticancer treatment started will be removed in the definition of PFS.

** 'More than two missed visits' occurs if the duration between the last tumor assessment and death or PD is longer than the following:
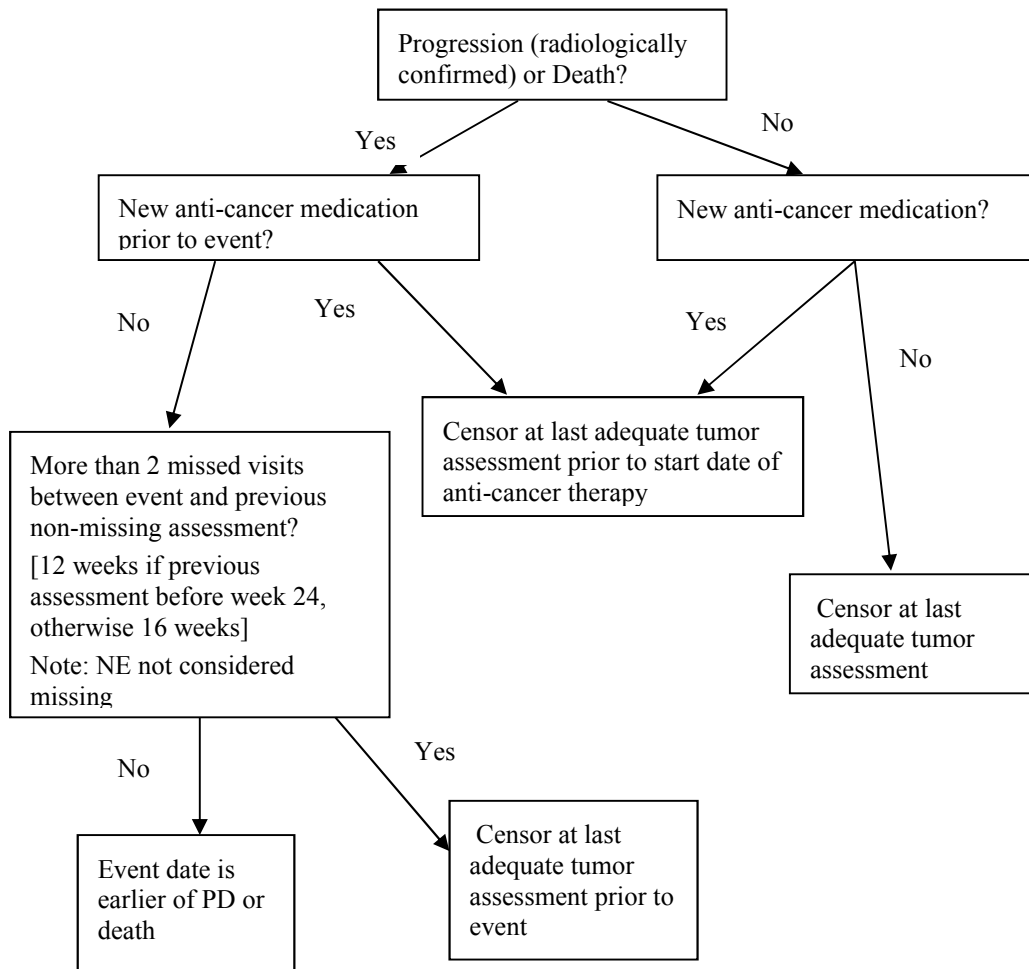
    1.   12 weeks if the last tumor assessment is prior to week 24

    2.   16 weeks if the last tumor assessment is on or after week 24

    Note: NE counts as a non-missing visit for No. 7.

Censoring rules are also described in Figure 1.

Figure 1 PFS Censoring Rules

Note: If subject is censored in chart below, censoring time will be first dose date if no adequate post-baseline assessments.

```
                          ┌────────────────────────┐
                          │ Progression (radiologically │
                          │ confirmed) or Death?       │
                          └────────────────────────┘
                    Yes   /                        \   No
                         /                          \
      ┌──────────────────────┐            ┌──────────────────────┐
      │ New anti-cancer medication │        │ New anti-cancer medication? │
      │ prior to event?            │        └──────────────────────┘
      └──────────────────────┘           Yes  /            \  No
         No  /        \  Yes                  /              \
            /          \                     /                \
  ┌──────────────────────┐  ┌─────────────────────────────┐  ┌──────────────────┐
  │ More than 2 missed     │  │ Censor at last adequate tumor │  │ Censor at last     │
  │ visits between event   │  │ assessment prior to start date │  │ adequate tumor     │
  │ and previous           │  │ of anti-cancer therapy         │  │ assessment         │
  │ non-missing assessment?│  └─────────────────────────────┘  └──────────────────┘
  │ [12 weeks if previous  │
  │ assessment before week │
  │ 24, otherwise 16 weeks]│
  │ Note: NE not considered│
  │ missing                │
  └──────────────────────┘
     No  /        \  Yes
        /          \
 ┌──────────────┐  ┌──────────────────────┐
 │ Event date is │  │ Censor at last         │
 │ earlier of PD │  │ adequate tumor         │
 │ or death      │  │ assessment prior to    │
 └──────────────┘  │ event                  │
                   └──────────────────────┘
```

Handling of Adverse Events Dates missing or partial missing:

For an AE with start date missing, if the date cannot indicate that the AE started prior to the first study treatment or after the 28 days post the last study treatment then it will be considered as treatment emergent adverse event (TEAE), and will be flagged as TEAE in the dataset. No imputation on the dates will be made in the datasets.

Unscheduled Visits:

Data from unscheduled visits (eg, vital signs or laboratory tests) will be excluded from the by-visit summary but will contribute to the worst value in required in summary tables. Listings will include all, the scheduled and unscheduled.

9   PROGRAMMING SPECIFICATIONS

The rules for programming derivations and dataset specifications are provided in separate documents.

10 STATISTICAL SOFTWARE

Statistical programming and analyses will be performed using SAS$^{®}$ (SAS Institute, Inc., Cary, NC, USA), version 9 or higher, and/or other validated statistical software as required.

11 MOCK TABLES, LISTINGS, AND GRAPHS

The study TLG shells will be provided in a separate document, which will show the content and format of all tables, listings, and graphs in detail.

12 REFERENCES

The Study Protocol of E7080-M000-213 (Version 3.0, dated 24May2016, Amendment 02) and its corresponding case report form.

Common Terminology Criteria for Adverse Events (CTCAE). Version 4.03. United States Department of Health and Human Services, National Institutes of Health, National Cancer Institute, Washington, DC, USA, May 28, 2009.

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

Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. Washington, DC: United States Food and Drug Administration; 2007.

13 APPENDICES

13.1 Sponsor's Grading for Laboratory Values

See Appendix 7 of the final protocol of this study.

SIGNATURE PAGE

Author:

PPD

_____   23 Oct 2018
                                    Date

Oncology Business Group, Eisai Inc.

Approval:

PPD

_____   23 OCT 2018
                                    Date

Oncology Business Group, Eisai Inc.

PPD

_____   30 OCT 2018
                                    Date

Oncology Business Group, Eisai Inc.