STATISTICAL ANALYSIS PLAN

Protocol Number:	SGNTUC-016
Version:	Version 3; 31JUL2023
Protocol Title:	Randomized, double-blind, phase 3 study of tucatinib or placebo in combination with ado-trastuzumab emtansine (T-DM1) for subjects with unresectable locally advanced or metastatic HER2+ breast cancer (HER2CLIMB-02)
Sponsor:	Seagen, Inc. 21823 30th Drive SE Bothell, WA 98021, USA

APPROVAL SIGNATURES

Product:	Tucatinib
Protocol Number/Amendment:	SGNTUC-016/Amendment 3
SAP Version:	Version 3
Version Date:	31JUL2023

The individuals signing below have reviewed and approved this statistical analysis plan.

[redacted]
Seagen, Inc.

Date

[redacted]
Seagen, Inc.

Date

[redacted]
Seagen, Inc.

Date

[redacted]
Seagen, Inc.

Date

TABLE OF CONTENTS

1 INTRODUCTION 7
2 STUDY OBJECTIVES 7
2.1 Primary Objectives 7
2.2 Secondary Objectives 7
2.2.1 Key Secondary Objectives 7
2.2.2 Other Secondary Objectives 7
2.3 Exploratory Objectives 8
3 STUDY ENDPOINTS 8
3.1 Primary Endpoints 8
3.2 Key Secondary Endpoints 8
3.3 Other Secondary Endpoints 8
3.4 Exploratory Endpoints 8
4 STUDY DESIGN 9
5 ANALYSIS SETS 10
5.1 Intent-to-Treat (ITT) Analysis Set 10
5.2 Safety Analysis Set 10
5.3 Patient Reported Outcomes Analysis Set 10
5.4 Pharmacokinetics (PK) Analysis Set 11
6 STATISTICAL CONSIDERATIONS 11
6.1 General Principles 11
6.2 Determination of Sample Size 11
6.3 Randomization and Blinding 12
6.3.1 Randomization 12
6.3.2 Blinding 12
6.4 Data Transformations and Derivations 13
6.5 Handling of Dropouts and Missing Data 14
6.6 Multicenter Studies 15
6.7 Multiple Comparison/Multiplicity 15
6.8 Examination of Subgroups 16
6.9 Covariates 17
6.10 Timing of Analyses 17
7 PLANNED ANALYSES 18
7.1 Disposition 18
7.2 Demographic and Baseline Characteristics 18
7.3 Protocol Deviations 18
7.4 Treatment Administration 18
7.5 Efficacy Analyses 19
7.5.1 Primary Endpoint 20
7.5.2 Key Secondary Endpoints 23
7.5.3 Other Secondary Endpoints 25
7.6 Safety Analyses 26
7.6.1 Adverse Events 26
7.6.2 Adverse Events of Special Interest and Other Adverse Events 27
7.6.3 Clinical Laboratory Results 28
7.6.4 Vital Signs 29
7.6.5 Deaths 29
7.6.6 Pregnancy 29
7.7 Pharmacokinetic Analyses 29
7.8 Health Economics and Outcomes 29
7.8.1 Health Care Resource Utilization 29
7.8.2 Patient Reported Outcomes (PRO) 30

8 INTERIM ANALYSIS 30
9 CHANGES FROM PLANNED ANALYSES 30
9.1 Changes from the Original SAP 30
10 REFERENCES 31
11 APPENDICES 32
APPENDIX A: IMPUTATION OF PARTIAL MISSING ADVERSE EVENT DATES 32
APPENDIX B: DEFINITION OF THE TERM "TREATMENT-EMERGENT" WITH RESPECT TO AE CLASSIFICATION 34
APPENDIX C: RECIST VERSION 1.1 SUMMARY 35
APPENDIX D: IMPUTATION OF PARTIAL MISSING START AND END DATE OF PRIOR SYSTEMIC THERAPY, SUBSEQUENT CANCER-RELATED THERAPY, AND HOSPITALIZATION 36
APPENDIX E: IMPUTATION OF PARTIAL MISSING DATE OF DISEASE DIAGNOSIS 37
APPENDIX F: IMPUTATION OF PARTIAL MISSING DATE OF THE END OF TUCATINIB/PLACEBO ADMINISTRATION 38
APPENDIX G: IMPUTATION OF PARTIAL MISSING DEATH DATES 39
APPENDIX I: IMPUTATION OF PARTIAL MISSING FIRST AND LAST DATES OF CORTICOSTEROID ADMINISTRATION 40
APPENDIX J: STATISTICAL ANALYSIS PLAN FOR THE CHINA POPULATION 41
J.1 RESPONSIBILITY FOR ANALYSES/IN-HOUSE BLINDING 41
J.2 HYPOTHESES/ESTIMATION 41
J.3 THE STUDY ENDPOINTS 41
J.3.1 Primary Endpoints 41
J.3.2 Secondary Endpoints 41
J.3.3 Exploratory Endpoints 42
J.4 ANALYSIS SETS 42
J.4.1 Intent-to-Treat (ITT) Analysis Set 42
J.4.2 Safety Analysis Set 42
J.4.3 Pharmacokinetics (PK) Analysis Set 42
J.5 PLANNED ANALYSES 42
J.5.1 Efficacy Analyses 42
J.5.2 Safety Analyses 43
J.5.3 Pharmacokinetic Analyses 43
J.5.4 Summaries of Baseline Characteristics, Demographics, and Other Analyses 44
J.6 INTERIM ANALYSES AND FINAL ANALYSIS 44
J.7 MULTIPLICITY 44
J.8 DETERMINATION OF SAMPLE SIZE 44

LIST OF IN-TEXT TABLES

Table 1: Summary of Planned Analyses 15
Table 2: PFS Event and Censoring Rules for Primary Analysis 20
Table 3: Search Strategy for adverse events of special interest and other adverse events (MedDRA 26.0) 28

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

AE	Adverse Event(s)
BICR	Blinded Independent Central Review
CBR	Clinical Benefit Rate
CMH	Cochran Mantel Haenszel
CNS	Central Nervous System
CR	Complete Response
CTCAE	Common Toxicity Criteria for Adverse Events
IDMC	Independent Data Monitoring Committee
DCO	Data Cutoff
DOR	Duration of response
ECOG	Eastern Cooperative Oncology Group
EOT	End of Treatment
eCRF	Electronic Case Report Form
ED	Emergency Department
HER2	Human Epidermal Growth Factor Receptor 2
HR	Hazard Ratio
HCRU	Healthcare Resource Utilization
ITT	Intent-to-Treat
LA/M	Locally-Advanced or Metastatic
MedDRA	Medical Dictionary for Regulatory Activities
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PFS	Progression-Free Survival
PK	Pharmacokinetics
PR	Partial Response
PRO	Patient-reported outcome
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SD	Stable Disease
SOC	System Organ Class
T-DM1	Ado-trastuzumab emtansine or trastuzumab emtansine
TEAE	Treatment-emergent Adverse Event
ULN	Upper Limit of Normal

1 INTRODUCTION

This document outlines the statistical methods to be implemented within the scope of Protocol SGNTUC-016: Randomized, double-blind, phase 3 study of tucatinib or placebo in combination with ado-trastuzumab emtansine (TDM1) for subjects with unresectable locally advanced or metastatic HER2+ breast cancer. Results of the proposed analyses will become the basis of the clinical study report for this protocol.

The purpose of this plan is to provide specific guidelines from which the analysis will proceed. All planned analyses specified in this document will be performed. Any changes to this plan, in the form of "post hoc" or "data driven" analyses will be identified as such in the final clinical study report. Any changes will either be reflected in amendments to this plan before the database lock or specifically documented in the clinical study report.

2 STUDY OBJECTIVES

2.1 Primary Objectives

- Compare the effect of tucatinib vs. placebo in combination with T-DM1 on progression-free survival (PFS) by investigator assessment per Response Evaluation Criteria in Solid Tumors (RECIST) v1.1.

2.2 Secondary Objectives

2.2.1 Key Secondary Objectives

- Compare overall survival (OS) between treatment arms.
- Compare PFS by investigator assessment per RECIST v1.1 in subjects with brain metastases at baseline between treatment arms.
- Compare the confirmed objective response rate (cORR) by investigator assessment per RECIST v1.1 between treatment arms.
- Compare OS in subjects with brain metastases at baseline between treatment arms.

2.2.2 Other Secondary Objectives

- Evaluate PFS by BICR per RECIST v1.1 by treatment arms.
- Evaluate PFS by BICR per RECIST v1.1 in subjects with brain metastases at baseline by treatment arms.
- Evaluate the cORR by BICR per RECIST v1.1 by treatment arms.
- Evaluate the duration of response (DOR) by investigator assessment per RECIST v1.1 by treatment arms.
- Evaluate the DOR by BICR per RECIST v1.1 by treatment arms.
- Evaluate the clinical benefit rate (CBR; stable disease [SD] or non-complete response/non-progressive disease [non-CR/non-PD] for ≥6 months or best response of

complete response [CR] or partial response [PR]) by investigator assessment per RECIST v1.1 by treatment arms.

- Evaluate the CBR by BICR per RECIST v1.1 by treatment arms.
- Evaluate the safety of tucatinib in combination with T-DM1.

2.3 Exploratory Objectives

- Evaluate the pharmacokinetics (PK) of tucatinib and DM1 following administration of tucatinib and T-DM1 in combination.
- Evaluate on-trial healthcare resource utilization (HCRU) by treatment arms.
- Evaluate patient reported outcomes (PROs) and health-related quality of life (QoL) by treatment arms.

3 STUDY ENDPOINTS

3.1 Primary Endpoints

- PFS per RECIST 1.1 by Investigator assessment

3.2 Key Secondary Endpoints

- OS
- PFS per RECIST 1.1 by investigator assessment for subjects with brain metastases at baseline
- cORR per RECIST 1.1 by Investigator assessment
- OS for subjects with brain metastases at baseline

3.3 Other Secondary Endpoints

- PFS per RECIST 1.1 by BICR
- PFS per RECIST 1.1 by BICR in subjects with brain metastases at baseline
- cORR per RECIST 1.1 by BICR
- DOR per RECIST 1.1 by investigator assessment
- DOR per RECIST 1.1 by BICR
- Clinical benefit rate (CBR) per RECIST 1.1 by investigator assessment
- CBR per RECIST 1.1 by BICR
- Type, incidence, severity, seriousness, and relatedness of adverse events (AEs)

3.4 Exploratory Endpoints

- Pharmacokinetics (PK): Plasma concentrations of tucatinib and DM1

- HCRU outcomes: cumulative incidence of health resource utilization, including length of stay, hospitalizations, and emergency department (ED) visits. An HCRU event will be counted only if it occurred after the first dose of study treatment and before 37 days past the last dose of study treatment. Day 37 is chosen because the last collection of HCRU data is at EOT follow-up visit, which occurs 30 to 37 days after the last dose of study treatment.
- PROs and health-related quality of life (QoL) assessments

4 STUDY DESIGN

This is a randomized, double-blind, placebo-controlled, international, multicenter, phase 3 study designed to evaluate the efficacy and safety of tucatinib in combination with T-DM1 in subjects with unresectable LA/M HER2+ breast cancer who have had prior treatment with a taxane and trastuzumab in any setting. Subjects will be randomized in a 1:1 manner to receive 21-day cycles of treatment in one of the following two treatment groups:

- control arm: placebo given PO BID; T-DM1 3.6 mg/kg given intravenously (IV) every 21 days,
- experimental arm: tucatinib 300 mg PO BID; T-DM1 3.6 mg/kg IV every 21 days.

Tucatinib or placebo will be dispensed to subjects in a double-blinded manner. Protocol-defined visits and cycle numbering will be determined by T-DM1 dosing date, allowing for dose holds or delays with T-DM1. In the event T-DM1 is discontinued but study treatment with tucatinib/placebo continues, protocol-defined visits and cycle numbering will proceed using a 21-day cycle regardless of dose holds or delays for tucatinib/placebo. Study treatment will continue until unacceptable toxicity, disease progression, withdrawal of consent, or study closure.

Approximately 460 subjects will be randomized to the tucatinib or placebo arms (in combination with T-DM1) in a 1:1 ratio using a centralized randomization with the following four stratification factors:

- line of treatment for metastatic disease: first line vs. other,
- hormone receptor status: negative vs. positive,
- presence or history of treated or untreated brain metastases: yes vs. no,
- Eastern Cooperative Oncology Group (ECOG) performance status: 0 vs. 1.

Disease response per RECIST v1.1 (Eisenhauer 2009) will be assessed by both investigator and BICR. Response assessments will include measurement of all known sites of unresectable LA/M disease (including at a minimum the chest, abdomen, and pelvis), preferably by high quality spiral contrast computed tomography (CT), at baseline, every 6 weeks for the first 24 weeks, and every 9 weeks thereafter, irrespective of dose interruptions. Contrast MRI of the brain will be required on this same schedule only in those subjects with prior history of brain metastases or brain metastases found at screening. Additional contrast

MRIs of the brain may also be performed in subjects without known brain metastases if there is clinical suspicion of new brain lesions.

Treatment decisions will be made based upon local assessment of radiologic scans. Response assessments for each subject will continue until a PFS event per RECIST v1.1 by investigator assessment has been documented. Follow-up for survival will continue until study closure or withdrawal of consent.

Safety assessments will include surveillance and recording of AEs, physical examination findings, and laboratory tests. Plasma concentrations of tucatinib and T-DM1will also be measured.

A detailed study assessment schedule can be found in the protocol.

An independent data monitoring committee (IDMC) will monitor the safety of subjects and provide an ongoing clinical assessment of the study treatment's evolving safety profile as the trial progresses. The IDMC will review blinded and unblinded data that include study enrollment, treatment exposure, AEs, SAEs and selected clinical lab results. The IDMC will meet on a regular basis and make recommendations to the sponsor regarding the conduct of the trial. Further details regarding the IDMC activity are described in a separate IDMC charter.

5 ANALYSIS SETS

This section defines each of the analysis sets that will be utilized. The use of each analysis set will be discussed in Section 7.

5.1 Intent-to-Treat (ITT) Analysis Set

The ITT analysis set will include all subjects who are randomized on or before the date of last-patient-in (LPI) in the global study. The date of LPI in the global study is defined as the randomization date of the last subject that is enrolled in a clinical site outside of Mainland China, or the randomization date of the first subject in Mainland China, whichever is later. Subjects will be analyzed according to the treatment arm assigned at randomization regardless of actual treatment received.

5.2 Safety Analysis Set

The safety analysis set will include all subjects who are randomized or enrolled on or before the date of LPI in the global study and receive at least one dose of study treatment (tucatinib or placebo, or T-DM1). Treatment groups will be determined using the actual treatment received. A subject will be considered as in tucatinib treatment group in the Safety Analysis Set as long as this subject received any dose of tucatinib.

5.3 Patient Reported Outcomes Analysis Set

The Patient Reported Outcomes (PRO) analysis set will include all ITT subjects who receive at least one dose of study treatment and have an evaluable (i.e., completed a PRO instrument and at least one domain or single item can be computed) baseline PRO score and at least one

evaluable post-baseline PRO assessment. Subjects will be analyzed according to the treatment arm assigned at randomization regardless of actual treatment received.

5.4 Pharmacokinetics (PK) Analysis Set

The PK analysis set will include all subjects in the safety analysis set who have at least one non-missing PK assessment. Subjects will be evaluated by the actual treatment received. The PK analysis set will be used for PK analyses.

6 STATISTICAL CONSIDERATIONS

6.1 General Principles

In general, descriptive statistics will be presented that include the number of observations, mean, standard deviation, median, minimum, and maximum for continuous variables. Frequencies and percentages will be used to summarize categorical variables.

Unless otherwise specified, confidence intervals (CIs) will be calculated at a two-sided 95% level. The two-sided 95% exact CI using Clopper-Pearson methodology will be calculated for the response rates where applicable (e.g., cORR).

For time-to-event endpoints, the median survival time will be estimated using the Kaplan-Meier method; the associated 95% CI will be calculated based on the complementary log-log transformation .

To comply with regulatory electronic submission guidelines, listings of all clinical data will be submitted as electronic data sets. To facilitate data review for the study report, only pertinent data listings will be created and attached to the appendix of the CSR.

All statistical tables, listings, and figures will be produced using SAS, version 9.4 or later. Other statistical software, if used, will be described in the CSR.

6.2 Determination of Sample Size

The sample size of this study is determined to achieve 90% power for the primary endpoint of PFS with a two-sided alpha level of 0.05.

A total of 331 PFS events is required with 90% power to detect a HR of 0.70 (median PFS of 8.57 vs. 6 months for the experimental and the control arms) at a 2-sided alpha level of 0.05 using a log-rank test. Approximately 460 subjects will be randomized in a 1:1 ratio to either the experimental or the control arm to observe 331 PFS events in approximately 30 months after the first subject is randomized, assuming 24 months of subject accrual, a 5% annual dropout rate.

It is planned that follow-up for OS will continue after the primary analysis of PFS until approximately 253 OS events have occurred. With 253 events, it will provide 80% power to detect a HR of 0.70 in OS at a 2-sided alpha level of 0.05 using a log-rank test. The final analysis of OS is estimated to take place approximately 30 months after the primary analysis of PFS assuming OS for the control arm follows an exponential distribution with a median of 29 months.

Follow-up of PFS in subjects with brain metastases at baseline (PFS.BM) will continue after the primary analysis of PFS until the second interim analysis of OS. It is anticipated that approximately 175 PFS.BM events would have occurred at the time of the second interim analysis of OS, which will provide 65% power to detect a HR of 0.7 in PFS.BM at a 2-sided significance level of 0.05 using a log-rank test assuming PFS.BM for the control arm follows an exponential distribution with a median of 5.7 months.

If the results of PFS, OS, and PFS.BM are all statistically significant, the cORR by investigator assessment will be formally tested between two arms at the 2-sided significance level of 0.05. With 460 randomized subjects, the power is 81% to detect an odds ratio of 1.72 (assuming 34% and 47% of cORR in control and treatment arm, respectively) in cORR at a 2-sided significance level of 0.05.

If the results of PFS, OS, PFS.BM and cORR are all statistically significant, the OS in subjects with brain metastases at baseline (OS.BM) will be formally tested between treatment arms at the 2-sided significance level of 0.05. It is anticipated that approximately 134 OS.BM events would have occurred at the time of final OS analysis, which will provide 83% power to detect a HR of 0.6 in OS.BM at a 2-sided significance level of 0.05 using a log-rank test assuming OS.BM for the control arm follows an exponential distribution with a median of 26 months.

Sample size and power were calculated using EAST® version 6.4, by Cytel Inc.

6.3 Randomization and Blinding

6.3.1 Randomization

Subjects will be assigned to the tucatinib or placebo arms (in combination with T-DM1) in a 1:1 ratio using a centralized randomization with the following four stratification factors:

- line of treatment for metastatic disease: first line vs. other,
- hormone receptor status: negative vs. positive,
- presence or history of treated or untreated brain metastases: yes vs. no,
- Eastern Cooperative Oncology Group (ECOG) performance status: 0 vs. 1.

6.3.2 Blinding

This is a double-blinded trial. Study subjects, site investigators and personnel, the sponsor, and all other individuals involved in the monitoring, data management, and/or conduct of the trial will be blinded. There are a few designated unblinded personnel who may access to the unblinded data. Please refer to Data Flow and Maintenance of the Blind for more details.

Unblinding a subject's treatment assignment prior to study closure must be limited to emergency circumstances where knowledge of the treatment assignment would affect decisions regarding the clinical management of the subject. In the event of such an emergency circumstance, a formal unblinding procedure, conducted by a third-party organization will be followed to allow the investigator to immediately access a subject's

treatment assignment (see Study Manual). Information on study treatment assignment should not be distributed to any other personnel involved in the clinical trial. In the event of any emergency unblinding, the sponsor is to be notified within 24 hours of the occurrence. Details regarding unblinding procedures are described in the Study Manual.

Unblinded data including deaths, discontinuations, dose reductions, adverse events (serious and non-serious) will be monitored regularly by an IDMC. The independent statistical reporting group (ISRG) preparing outputs for the IDMC will be unblinded and have access to the overall randomization scheme.

At the time of the primary analysis for the primary endpoint (PFS per investigator), sponsor personnel will be unblinded to prepare related clinical study report (CSR), however investigators and staff at clinical sites will remain blinded to individual subject treatment assignments (tucatinib/placebo) until the final analysis for the key secondary endpoint of OS.

6.4 Data Transformations and Derivations

Age: Reported age in years will be used.

Baseline: Unless otherwise specified, baseline values used in all analyses will be the most recent non-missing measurement prior to the first dose of any study treatment (for safety analyses) or randomization (for efficacy analyses).

Study Day: For subjects received treatment, study day starts on the date of the first dose (date - date of the first dose + 1) for dates on or after the first dose date. For dates prior to the first dose date, study day is calculated as date - date of the first dose. The first dose date is the earliest date of administration of any study treatment.

For subjects who did not received treatment, study day starts on the date of randomization (date – date of randomization + 1) for dates on or after randomization date. For dates prior to randomization date, study day is calculated as date – date of randomization.

Other time variables based on two dates, e.g., Start Date and End Date, will be calculated as (End Date–Start Date+1) (in days) unless otherwise specified in the planned analysis section.

The following unit conversion will be implemented unless otherwise specified:

$$\text{Months} = \text{Days}/30.4375$$

$$\text{Years} = \text{Days}/365.25$$

Study treatment: tucatinib/placebo, or T-DM1

Response assessment date: At each response assessment time point, scans to evaluate tumor lesions can be performed on multiple dates.

- The date of response for CR or PR will be recorded as the date of the last radiographic evaluation included in the series for that assessment.
- The date of response for non-CR/non-PD or SD will be recorded as the date of the earliest radiographic evaluation included in the series for that assessment.

- The date of progression (PD) will be recorded as the earliest date that PD has been documented, i.e., the earliest of the following:
    - Date of target lesion assessment when the target lesion response is PD,
    - Date of non-target lesion assessment when the lesion status is unequivocal progression,
    - Date of documenting new lesion.

  In the cases where a PD occurs due to the fact that an equivocal new lesion was continuously present and later confirmed to be an unequivocal new lesion, the PD date should be back dated to the date when the equivocal new lesion was first identified. The tumor response on the date when the equivocal new lesion was first identified will be changed to PD. If an equivocal new lesion was later absent or confirmed to be a benign lesion, then this new lesion is not considered to define a PD. In cases where PD occurs on a date after an equivocal new lesion is identified, but the progression is not due to a change of the equivocal new lesion to an unequivocal lesion, but from progression of other lesions, the PD date will not be back dated, but will be the date when definitive PD is recorded.

6.5 Handling of Dropouts and Missing Data

Except for the scenarios covered in this section, missing data will not be imputed.

Missing disease diagnosis date will be imputed for the purpose of calculating the time from diagnosis to date of randomization or date of first dose (Appendix E).

Missing AE dates will be imputed for the purpose of calculating duration of events (Appendix A).

Missing start and end date of subsequent cancer-related therapy, prior treatment date, and hospitalization dates will be imputed for the purpose of deriving the time-to-event endpoints, and other applicable analysis as applicable (Appendix D).

Missing end date of tucatinib/placebo administration will be imputed for the purpose of deriving cumulative dose of tucatinib/placebo (Appendix F).

Missing date of death will be imputed for deriving time-to-event endpoints (Appendix G).

Missing date of corticosteroid will be imputed for deriving the treatment duration (Appendix I).

Unless otherwise specified, lab values which are recorded or provided as being less than the lower limit of quantification (LLOQ), will be included in figures and summaries as LLOQ.

PK values which are recorded as being less than the lower limit of quantification (LLOQ), will be included in figures and summaries as LLOQ/2.

.

6.6 Multicenter Studies

There are multiple centers in this study, however it is not anticipated that any center will accrue enough subjects to warrant an analysis by center.

6.7 Multiple Comparison/Multiplicity

To maintain strong control of the family-wise type I error rate at 0.05 (two-sided), a fixed sequential testing procedure will be used to test the primary endpoint of PFS per investigator assessment and the key secondary endpoints of OS, PFS per investigator assessment in subjects with brain metastases at baseline, cORR per investigator and OS in subjects with brain metastases at baseline.

If the final analysis of PFS per investigator is statistically significant at a two-sided significance level of 0.05, OS will be tested three times with an overall two-sided alpha of 0.05. The first analysis of OS (OS interim analysis 1 [OS IA1]) will be performed at the same time as the primary PFS analysis when approximately 331 PFS events per investigator assessment are observed. The second OS interim analysis (OS IA2) is planned when approximately 202 (80% information fraction) OS events have occurred. The final OS analysis will occur when approximately 253 OS events are observed. The Lan-DeMets O'Brien-Fleming approximation spending function will be used for the calculation of the alpha level at each test of OS based on the actual number of OS events observed at the time of each analysis.

If both PFS and OS results are statistically significant, then PFS by investigator assessment per RECIST v1.1 in subjects with brain metastases at baseline (PFS.BM analysis) will be tested twice with an overall two-sided alpha of 0.05. The first PFS.BM analysis will use the PFS.BM data at the time of the primary endpoint analysis when approximately 331 PFS events per investigator assessment have occurred with a nominal 2-sided alpha of 0.0001 (Haybittle-Peto approach [Haybittle 1971; Peto 1976). The final PFS.BM analysis will use the PFS.BM data at the time of OS IA2 or when approximately 175 events of PFS.BM are observed if OS IA 2 is not performed, and the p-value will be compared with a 2-sided alpha of 0.05.

Table 1 shows the stopping boundaries for planned analyses for illustrative purposes. The actual stopping boundary will be determined by the number of events observed.

Table 1: Summary of Planned Analyses

Analyses	PFS Events (%)	PFS Stopping Boundary	OS Events (%)	OS Stopping Boundary [a]	PFS.BM Events (%)	PFS.BM Stopping Boundary [b]	OS.BM Events (%)	OS.BM Stopping Boundary [a]
PFS FA OS IA1 PFS.BM IA1 OS.BM IA1	331 (100)	p<0.05	116 (46)	p<0.002	145 (83)	p<0.0001	63 (47)	p<0.002
OS IA2			202 (80)	p<0.024	175 (100)	p<0.05	111 (83)	p<0.026

PFS.BM FA OS.BM IA2 [c]					
OS FA OS.BM FA [d]	253 (100)	p<0.042		134 (100)	p<0.042

FA: final analysis. IA: interim analysis.

a. Using Lan-DeMets O'Brien-Fleming approximation alpha spending function.
b. Using Haybittle-Peto alpha spending approach.
c. If OS IA2 is not performed, then PFS.BM FA and OS.BM IA2 will be conducted when approximately 175 events of PFS.BM are observed. If PFS.BM FA is not performed, then OS.BM IA2 will be conducted when approximately 111 events of OS.BM are observed.
d. If OS FA is not performed, then OS.BM FA will be conducted when approximately 134 events of OS.BM are observed.

If the results of PFS, OS, and PFS.BM are all statistically significant, the cORR by investigator assessment will be formally tested between two arms at the 2-sided significance level of 0.05. Specifically, cORR will be tested once using the cORR data at the time of PFS primary analysis.

If PFS, OS, PFS.BM and cORR results are all statistically significant, then OS in subjects with brain metastases at baseline (OS.BM analysis) will be tested three times with an overall two-sided alpha of 0.05. The first analysis of OS.BM (OS interim analysis 1 [OS.BM IA1]) will be performed at the same time as the primary PFS analysis when approximately 331 PFS events per investigator assessment are observed. The second OS.BM interim analysis (OS.BM IA2) is planned when approximately 202 (80% information fraction) OS events have occurred. The final OS.BM analysis will occur when approximately 253 OS events are observed. The Lan-DeMets O'Brien-Fleming approximation spending function will be used for the calculation of the alpha level at each test of OS.BM based on the actual number of OS.BM events observed at the time of each analysis.

6.8 Examination of Subgroups

As exploratory analyses, subgroup analyses may be conducted for PFS by investigator, PFS by investigator in subjects with brain metastases at baseline and OS. Subgroups may include but are not limited to the following:

- brain metastases at baseline (yes, no): brain metastases at baseline are defined as having either presence or history of brain metastases prior to screening/baseline MRI or having newly diagnosed or equivocal lesions found on the screening/baseline MRI. The subgroup of subjects with brain metastases at baseline will henceforth be referred to as "BrainMets" in this document. Subjects not meeting the above criteria will be assigned to the 'no' subgroup for this variable;
- line of treatment for metastatic disease (first line, Other);
- ECOG (0, 1): the latest ECOG on or before the date of the first dose of study treatment;

- hormone receptor status (negative, positive): hormone receptor is defined as positive if either estrogen receptor or progesterone receptor is positive, negative otherwise;
- geographic Region (North America, Europe/Israel, Asia-Pacific);
- age (<65, ≥65 years);
- race (white, Asian, African American, others): Asian or African American may be combined with others if any subgroup is less than 10% of the total population;
- stage at initial diagnosis (0—III, IV);
- prior use of pertuzumab (yes, no).

Subgroups will be defined using data derived from corresponding data fields in the eCRF. If the number of subjects in a subgroup is less than 10% of the total population, analysis of that subgroup will either be combined with another subgroup or will not be performed. Subgroup analyses will be conducted using stratified log-rank tests and stratified Cox proportional hazards model using randomization stratification factors. If the subgroup is a stratification factor, then the stratified models will control for the other stratification factors.

6.9 Covariates

Stratified analyses specified in Section 7 will include adjustment for the stratification factors as recorded at randomization.

6.10 Timing of Analyses

There is only one formal analysis of the primary endpoint PFS per investigator assessment, which will occur after approximately 331 PFS events in the ITT analysis set have occurred. The analysis cutoff date for this analysis will be determined once approximately 331 PFS events per investigator assessment have occurred. This is estimated to be approximately 30 months after randomization of the first subject.

The key secondary endpoint of OS will be analyzed three times. OS IA1 will be performed at the same time as the primary analysis of PFS per investigator assessment. OS IA2 is planned when approximately 202 (80%) OS events have occurred, which is estimated to be approximately 46 months after randomization of the first subject. The final analysis of OS will occur when approximately 253 OS events have occurred. The final OS analysis is estimated to occur 60 months after randomization of the first subject.

If both PFS and OS results are statistically significant, PFS.BM will be analyzed twice. The first analysis will be performed using the PFS.BM data at the same time as the primary analysis of PFS per investigator assessment. The final PFS.BM analysis will be conducted using the PFS.BM data at the time of OS IA2. The formal testing of final PFS.BM will occur at the earliest time when both PFS and OS (either OS IA2 or OS FA) results are statistically significant.

Confirmed ORR by investigator assessment will be formally tested if the results of the PFS, OS, and PFS.BM analyses are all statistically significant, using the data at time of the primary analysis of PFS.

If PFS, OS, PFS.BM and cORR results are all statistically significant, then OS.BM will be analyzed three times with an overall two-sided alpha of 0.05. OS.BM IA1 will be performed at the same time as the primary PFS analysis. OS.BM IA2 is planned at the same time as OS IA2. The final OS.BM analysis will occur at the same time as the OS final analysis.

7 PLANNED ANALYSES

7.1 Disposition

Analysis set: ITT Analysis Set

An accounting of study subjects by disposition will be tabulated by treatment arm and total. Reasons for discontinuation of treatment and study will be summarized. The number and percentage of subjects who signed informed consent and the number of subjects in each analysis set will be summarized by treatment arm and total. Number of screen failures and the percentage relative to the total number of subjects screened will be summarized. A listing of subjects who failed screening will be produced with reasons for screen failure and available demographic information.

The number of subjects enrolled in each country and at each site will be summarized by treatment arm and total. Follow up time and subsequent treatment information will be summarized.

7.2 Demographic and Baseline Characteristics

Analysis set: ITT Analysis Set

Demographics and baseline characteristics, including age, gender, ethnicity, race, baseline weight, and ECOG score will be listed and summarized; summaries will be presented for each treatment arm and total using the ITT analysis set. Disease specific characteristics, including time from diagnosis and previous cancer-related treatments will be listed and summarized for each treatment arm and the total.

7.3 Protocol Deviations

Analysis set: ITT Analysis Set

Important protocol deviations are those that represent a divergence from the protocol that could have a significant effect on the integrity of the study data, or on the subject's rights, safety, or welfare. Important protocol deviations also include exemptions to the study inclusion/exclusion criteria. Important protocol deviations will be summarized by category and treatment arm. A list of subjects with important protocol deviations will be presented.

7.4 Treatment Administration

Analysis set: Safety Analysis Set

Exposure will be summarized by treatment arm using counts and percentages for categorical variables and summary statistics (mean, median, standard deviation, and range) for continuous variables. The following information will be summarized by treatment group separately for tucatinib/placebo and T-DM1:

- total number of treatment cycles per subject,
- duration of treatment (months),
- dose modification by type (dose reduced, drug interrupted, dose held, drug withdrawn) and by reason,
- dose holds due to AE,
- duration of dose holds due to AE,
- absolute dose intensity (ADI) and relative dose intensity (RDI) for tucatinib/placebo and T-DM1.

Duration of treatment is defined, per treatment of T-DM1 and tucatinib/placebo, as the time from first dose date to the earliest of following dates:

- exposure end date:
  - for T-DM1, last infusion date + 20,
  - for tucatinib/placebo, the last dose date,
- date of death,
- end of study date,
- analysis data cutoff (DCO) date if the subject is still on study at the time of DCO.

Cumulative dose is defined as the sum of the actual dose amount that a subject received across all cycles.

Intended Dose Intensity (IDI) is defined as the intended dose of drug per unit of time, i.e., 600 mg/day for tucatinib/placebo and 3.6/21 mg/kg/day for T-DM1

Absolute Dose Intensity (ADI) is

- For tucatinib/placebo, cumulative dose / (exposure end date – first dose date + 1),
- For T-DM1, cumulative dose / (exposure end date – first infusion date +1).

Relative Dose Intensity (RDI) = ADI/IDI × 100%.

For subjects who received both tucatinib and placebo due to dispensing error, the duration of treatment and cumulative dose include all dosing of tucatinib and placebo.

7.5 Efficacy Analyses

Analysis set: ITT Analysis Set

The analysis for efficacy endpoints, including the primary endpoint of PFS, the key secondary endpoint of OS and cORR, and other secondary and exploratory efficacy endpoints will be conducted using the ITT analysis set, unless otherwise specified. Subjects will be analyzed based on their randomized treatment arm.

For the primary analysis of PFS and OS, the randomization stratification factors per Randomization and Trial Supply Management (RTSM) system will be used as strata in stratified analysis. If the total number of subjects in a stratum is less than 10% of the total population, that stratification factor will be excluded in both stratified log-rank test and stratified Cox proportional hazard regression model.

If the inconsistency between RTSM and eCRF stratification factors is larger than 5%, a sensitivity analysis will be conducted using the stratification factors derived from eCRF.

The primary analysis of the primary endpoint (PFS) and key secondary endpoints (OS, PFS.BM, cORR and OS.BM) may be repeated for each of the subgroups specified in Section 6.7. The subgroups are defined by data from eCRF.

7.5.1 Primary Endpoint

7.5.1.1 Progression Free Survival (PFS) by Investigator Assessment

Progression-free survival (PFS) is defined as the time from randomization to the first documented disease progression (as assessed by investigator per RECIST 1.1) or to death due to any cause, whichever occurs first.

Specifically,

PFS=Date of first documented PD or death or censoring-Date of randomization+1.

PFS is defined based on the table below:

Table 2: PFS Event and Censoring Rules for Primary Analysis

Scenario	Event/Censor Date	Outcome
No post-baseline tumor assessments	Date of randomization	Censored
No documented disease progression or death	Date of last tumor assessment of CR, PR, SD, or non-CR/non-PD	Censored
Subsequent cancer-related therapy (systemic, radiation, or surgery) started before PD or death observed	Date of last CR, PR, SD, or non-CR/non-PD on or prior to date of subsequent cancer-related therapy	Censored
Progressive disease (PD)	Date of PD	Event
Death before first PD assessment	Date of death	Event
Death or progression right after two or more consecutively missed tumor assessments	Date of last tumor assessment of CR, PR, SD, or non-CR/non-PD prior to the missed visits	Censored

Note: Tumor assessments are performed every 6 weeks starting at Cycle 1 Day1 through Week 24 and every 9 weeks starting at Week 24 until documented PD or death.

Disease progression includes any subject with progressive disease per RECIST 1.1, including those who have isolated CNS progression per RECIST 1.1 and continue study treatment for clinical benefit (see protocol Section 7.2.2).

The null hypothesis for the primary endpoint is that the PFS per investigator of the tucatinib arm has no difference from that of the placebo arm. To test this hypothesis, a stratified log-rank test comparing two arms in the ITT analysis set will be conducted controlling for the randomization stratification factors [i.e., line of treatment for metastatic disease (1st line vs.

other), hormone receptor status (negative vs. positive), history or presence of brain metastases (yes vs. no), and ECOG status (0 vs. 1)].

Kaplan-Meier curves and estimates of the median PFS time will be provided for each arm with the two-sided 95% confidence intervals (CI) using the complementary log-log transformation method . KM estimates of the 25th and 75th percentiles of PFS time and the observed minimum and maximum PFS time will also be reported. In addition, estimated probability of PFS will be reported at every 3 months until the time of the last PFS event.

A sample SAS code for the stratified log-rank test is provided below.

```
** Pfstime = PFS time;
** Censor = Censor variable (1 = censored);
** Trt = treatment arm
** LT = line of treatment for metastatic disease (1st line vs. other)
** HRS = hormone receptor status (negative vs. positive),
** BM = presence or history of brain metastases at baseline (Yes, No)
** ECOG (0,1)


ODS OUTPUT HomTests=chisq;

PROC LIFETEST DATA = pfsdata;
        TIME pfstime*censor (1);
        STRATA lt hrs bm ecog / GROUP=trt TEST=logrank;
RUN;
```

For the purpose of describing the treatment effect, a hazard ratio between the experimental arm and the control arm and its 95% confidence interval will be estimated using a stratified Cox proportional hazards regression model controlling for the stratification factors. A sample SAS code is provided below.

```
PROC PHREG DATA = pfsdata;
        CLASS trt(ref='B');
                STRATA lt hrs bm ecog;
                MODEL pfstime*censor (1) = trt / TIES=EFRON RL;
                HAZARDRATIO trt;
RUN;
```

7.5.1.2 Sensitivity Analyses

1. **Clinical progression:** To explore the potential impact of clinical progression on the analysis of PFS, a sensitivity analysis may be performed using the same censoring scheme and methods as described for the primary analysis of PFS with the exception that clinical progression will be counted as 'progression' in the analysis.
2. **Missing assessments of disease response:** To explore the potential impact of missing tumor assessments on PFS, two sensitivity analyses are planned:

- Ignoring the missing assessments, i.e., subjects who missed two or more consecutive scheduled assessments will not be censored and will continue to be followed for tumor assessment.
- Imputing the missing assessment, i.e., subjects who missed two or more consecutive scheduled assessments before death or PD are considered to have events at the time of the first missed scheduled assessment after the last non-missing assessment.

3. **Subsequent cancer-related therapy before PD/death:** For subjects who received subsequent cancer-related therapy before PD or death, two sensitivity analyses are planned:
   - Not to consider any subsequent cancer-related therapies (whether systemic, radiation, or other) as a censoring reason;
   - Consider any subsequent cancer-related therapies (whether systemic, radiation, or other) as an event.
4. **Mis-stratification:** In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the hazard ratio and its 95% CI will be estimated using a stratified Cox proportional hazards regression model with stratification factors derived from eCRF.
5. **Restricted mean survival time (RMST):** The proportional hazard assumption for the treatment variable will be examined using the plot of Schoenfeld residuals for the treatment variable versus time. Residuals that do not show any trend indicate the proportional hazard assumption is satisfied. Other methods may also be applied to check the PH assumption if needed.

   If the results of the above mentioned analyses suggest the proportional hazard assumption is violated, then alternative methods such as restricted mean survival time (RMST) (Royston 2011; Uno 2014) may be used to describe the treatment effect. RMST is the expected survival time within a fixed follow-up interval and will be calculated as the area under the survival curve from t=0 up to time τ months for each treatment group.

   The selection of τ should ensure that the RMST evaluation will not go beyond the maximum time point where the evaluation can be performed while also including a large period of time that is expected to provide a meaningful assessment of treatment effect. To avoid arbitrary selection of the common cutoff τ for both treatment arms, the following cutoff time will be used:

   - $\tau 1$ = min of (latest PFS event time for experimental arm, latest PFS event time for control arm),
   - $\tau 2$ = min of (latest time that ≥10% of subjects are at risk for PFS in the experimental arm, latest time that ≥10% of subjects are at risk for PFS in the control arm).

   The cutoff time $\tau 1$ and $\tau 2$ will be rounded down to full months.

   The RMST (95% CI) and the difference of RMST between treatment arms will be estimated using nonparametric method. The significance of difference in RMST between

the treatment arms will be analyzed using an ANCOVA model (Andersen 2004) adjusting for the randomization stratification factors.

7.5.2 Key Secondary Endpoints

7.5.2.1 Overall Survival (OS)

The primary analysis of overall survival (OS) will be based on the ITT analysis set. Overall survival is defined as the time from randomization to date of death due to any cause:

OS=Date of death–Date of randomization+1.

In the absence of confirmation of death, OS will be censored at the last date the subject is known to be alive. If death date or the last long-term follow-up visit (which indicates subject to be alive) date were to be after DCO, then OS will be censored at DCO.

OS analysis will be performed using similar statistical methods as to evaluate the primary endpoint of PFS in the ITT analysis set. The stratified log-rank statistics and p-value will then be calculated as described in Section 7.5.1.1. The Lan-DeMets O'Brien-Fleming approximation spending function will be used to obtain the alpha level at each time of OS analyses based on the actual number of OS events observed at the interim.

Kaplan-Meier curves and estimates of the median OS time will be provided for each arm with the two-sided 95% confidence intervals (CI) using the complementary log-log transformation method . In addition, KM estimates of the 25th and 75th percentiles of OS time and the observed minimum and maximum OS time will be reported. Estimated probability of OS will be reported at every 6 months until the last OS event. The hazard ratio and 95% confidence interval will be estimated using a stratified Cox proportional hazards regression model controlling for the randomization stratification factors.

Sensitivity Analyses

Mis-stratification: In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the hazard ratio and its 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the stratification factors derived from eCRF.

RMST: Assumption of non-proportional hazard ratio and potential RMST methods will be performed similarly as described in Section 7.5.1.2. for PFS.

7.5.2.2 PFS in BrainMets (PFS.BM) by Investigator Assessment

Progression-free survival (as defined for the primary efficacy endpoint) in the BrainMets subgroup (defined in Section 6.8), PFS.BM, by investigator assessment will be analyzed with the same statistical methods used to evaluate the primary endpoint of PFS as in Section 7.5.1.1.

Sensitivity Analyses

Mis-stratification: In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the hazard ratio and its 95% CI will be estimated using a stratified Cox

proportional hazards regression model controlling for the stratification factors derived from eCRF.

RMST: Assumption of non-proportional hazard ratio and potential RMST methods will be performed similarly as described in Section 7.5.1.2. for PFS.

7.5.2.3 Confirmed objective response rate (cORR) by Investigator Assessment

Confirmed objective response rate (cORR) per investigator assessment is defined as the proportion of subjects achieving a best overall response of a confirmed CR or a confirmed PR per RECIST 1.1 by investigator assessment. Only tumor assessments before first documented PD and subsequent cancer-related therapies will be considered. For a response to be considered as confirmed, the subsequent response needs to be at least 4 weeks after the initial response. Subjects with unconfirmed CR or PR ≥ 5 weeks after the date of randomization will be considered to have SD as their best overall response on the date of the corresponding unconfirmed CR and PR. A subject will have a best overall response of SD or non-CR/non-PD if there is at least one SD or non-CR/non-PD assessment (including the unconfirmed CR/PR) ≥ 5 weeks after the date of randomization and the subject does not qualify for CR or PR.

The analysis of cORR by investigator will be conducted for subjects in the ITT analysis set with measurable disease at baseline per investigator assessment. Comparison in cORR between treatment arms will be performed using a 2-sided Cochran-Mantel-Haenszel (CMH) test controlling for the stratification factors. The p-value from the stratified CMH test will be reported. In addition, a two-sided 95% exact confidence interval for each treatment arm will be reported using the Clopper-Pearson method (1934).

Confirmed ORR may be separately summarized for subjects with and without brain metastases at baseline (as defined in section 6.8).

Sensitivity Analyses

Mis-stratification: In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the odds ratio and its 95% CI will be estimated using a CMH test controlling for the stratification factors derived from eCRF.

7.5.2.4 OS in BrainMets (OS.BM)

Overall survival (OS) as defined in Section 7.5.2.1 in the BrainMets subgroup (defined in Section 6.8), OS.BM, will be analyzed with the same statistical methods as defined in Section 7.5.2.1.

Sensitivity Analyses

Mis-stratification: In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the hazard ratio and its 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the stratification factors derived from eCRF.

RMST: Assumption of non-proportional hazard ratio and potential RMST methods will be performed similarly as described in Section 7.5.1.2. for PFS.

7.5.3 Other Secondary Endpoints

7.5.3.1 PFS by BICR

Progression-free survival by BICR is defined as the time from the date of randomization to the date of documented disease progression as determined per RECIST 1.1 by BICR or death from any cause, whichever occurs first. Subjects who are alive and have not progressed at the time of the analysis will be censored at the time of their last tumor assessment that was a CR, PR, non-CR/non-PD, or SD. Details of the censoring scheme of PFS are described above in Table 2.

Progression-free survival by BICR will be analyzed using the same statistical methods described in Section 7.5.1.1.

The concordance between BICR and investigator assessed PFS event will be summarized.

7.5.3.2 PFS in BrainMets by BICR

Progression-free survival (as defined in Section 7.5.3.1) in the BrainMets subgroup (defined in Section 6.8) by BICR assessment will be analyzed with the same statistical methods used to evaluate the primary endpoint of PFS as in Section 7.5.1.1.

7.5.3.3 cORR by BICR

Confirmed objective response rate (cORR) by BICR is defined as the proportion of subjects achieving a best overall response of a confirmed CR or a confirmed PR per RECIST 1.1 by BICR. Only tumor assessments before first documented PD by BICR and subsequent cancer-related therapies will be considered. For a response to be considered as confirmed, the subsequent response needs to be at least 4 weeks after the initial response.

The analysis of cORR by BICR will be conducted for subjects in the ITT analysis set with measurable disease at baseline per BICR assessment. Confirmed objective response rate by BICR will be analyzed using the same statistical methods described in Section7.5.2.3.

7.5.3.4 Clinical Benefit Rate (CBR)

Clinical benefit rate is defined as the proportion of subjects achieving SD or non-CR/non-PD for ≥ 6 months (i.e., subject has been followed for at least 6 months and no documented PD or death within 6 months from date of randomization) or a best overall response of a confirmed CR or a confirmed PR per RECIST 1.1.

Only response assessments before first documented PD and subsequent cancer-related therapies will be considered. The same derivation of PD date and censoring rules as for primary PFS analysis will apply for duration of SD or non-CR/non-PD.

Clinical benefit rate by investigator and BICR assessment will be analyzed using the same statistical methods described in Section 7.5.2.3.

7.5.3.5 Duration of Response (DOR)

Duration of response is defined as the time from the date of the first objective response (confirmed CR or confirmed PR) to the date of the first documented PD per RECIST 1.1 or

death due to any cause, whichever occurs first. The same derivation of PD date and censoring rules as for primary PFS analysis will apply for DOR.

Duration of response will only be calculated for the subgroup of subjects achieving a confirmed CR or confirmed PR. Kaplan-Meier curves and estimates of median DOR will be provided for each treatment arm with the two-sided 95% CI using the complementary log-log transformation method. Estimated probability of progression-free survival will be reported at every 3 months until the time of the last event.

The analysis of DOR will be repeated based on BICR assessment and investigator assessment.

7.6 Safety Analyses

Analysis sets: Safety Analysis Set

All safety endpoints will be analyzed using the Safety analysis set.

Adverse events will be coded by standard preferred terms (PT) and system organ classifications (SOC) using Medical Dictionary for Regulatory Activities (MedDRA, version 26.0 or higher).

Laboratory values will be graded using the National Cancer Institute's Common Toxicity Criteria for Adverse Events (CTCAE, version 4.03). For creatinine increase, CTCAE v5.0 will be used.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (WHODrug, Global-B3 Mar2023 or higher).

7.6.1 Adverse Events

Adverse events will be summarized by MedDRA preferred term (PT) in descending frequency of occurrence in the experimental arm (Tucatinib + T-DM1) unless otherwise specified. For incidence reporting, if a subject reports more than one AE that was coded to the same SOC or PT, the subject will be counted only once for that specific SOC or PT. For summaries by severity, only the worst grade for an AE will be counted for a subject.

Treatment-emergent adverse event (TEAE) is defined as AE that is new or worsened after the first dose of study treatment (tucatinib/placebo or T-DM1) and up through 30 days after the last dose of study treatment (tucatinib/placebo or T-DM1, whichever is later).

Treatment-related AE is defined as AE assessed by the investigator as 'related' to tucatinib/placebo or T-DM1.

Summaries of TEAEs by treatment arm will include (but may not limit to):

- all TEAEs
- TEAE by PT
- grade 3 or higher of TEAEs by PT
- serious TEAEs by PT
- TEAEs leading to dose modification by PT

- TEAEs leading to death by PT
- TEAEs by SOC and PT
- treatment-related TEAEs by PT
- treatment-related serious TEAEs by PT
- TEAEs by SOC, PT, and maximum severity. At each SOC or PT, multiple occurrences of events within a subject are counted only once at the highest severity.

All TEAEs, grade 3 or higher TEAEs, serious TEAEs, TEAEs leading to treatment discontinuation, and TEAEs leading to death will be listed.

7.6.2 Adverse Events of Special Interest and Other Adverse Events

Potential asymptomatic left ventricular systolic dysfunction, drug-induced liver injury, and cerebral edema are considered as Adverse Events of Special Interest (AESIs). These AESIs will be assessed by investigator based on the criteria defined in the protocol, and by search strategy. Drug-induced liver injury defined by lab parameters (per study protocol) may result in a small number of AEs. Hence, hepatotoxicity by search strategy will be summarized in addition. Incident of AESIs will be summarized for each treatment arm.

Other adverse events include diarrhea.

Crude rates, treatment exposure-adjusted incidence for grade 3 or higher Adverse Events of Special Interest (AESIs), and time-at-risk exposure-adjusted incidence for grade 3 or higher AESI will be reported. Treatment exposure starts from the date of the first dose of either study drug and ends on the earliest of the following:

- last dose date of either study drug + 30 days,
- data cutoff date,
- date of end of study.

Time at risk will be calculated for each AE, starting from the date of the first dose of either study drug. The end of time at risk is:

- for subjects with the specific AE, the onset date of AE;
- for subjects without the specific AE, the earliest of the following:
    - last dose date of either study drug + 30 days,
    - data cutoff date,
    - date of death,
    - date of end of study.

Table 3: Search Strategy for adverse events of special interest and other adverse events (MedDRA 26.0)

Identified/Potential Risks	Search Strategy
Asymptomatic left ventricular systolic dysfunction	TEAEs with preferred terms from cardiomyopathy SMQ (narrow) and cardiac failure SMQ (narrow) that led to change in study treatment or discontinuation of study treatment. OR TEAEs captured on the "AESI-Asymptomatic LVEF Decline" CRF.
Cerebral edema	All events reported by investigators on the AESI-Cerebral Edema CRF
Hepatotoxicity	Drug related hepatic disorders - comprehensive search SMQ (Narrow)
Diarrhea	PT of Diarrhoea

7.6.3 Clinical Laboratory Results

Clinical laboratory results collected during the study are specified in the protocol section of 7.8.4, including serum chemistry and hematology samples. The collection schedules are specified in Appendix A in the study protocol.

Both observed value and changes from baseline will be summarized with descriptive statistics for each scheduled visit by treatment arm. Shift from baseline to maximum post-baseline NCI CTCAE grade will be summarized for each lab test by treatment. Post baseline period is defined as from the date of the first dose of study treatment to the date of last dose + 30 days.

Treatment-emergent laboratory abnormalities will also be summarized.

Laboratory results and NCI CTCAE grades for hematology and serum chemistry will be presented in data listings. Normal ranges will be documented, and out-of-range values will be flagged. A separate listing of laboratory results with CTCAE grade 3 or higher will be presented.

7.6.3.1 Incidence of Liver Abnormalities

The incidence of liver abnormalities will be summarized by treatment arm. A liver abnormality is defined as

- AST or ALT elevations that are > 3 x ULN with concurrent elevation of total bilirubin > 2 x the ULN (same day or within 21 days following AST and/or ALT elevations), or
- AST or ALT elevations that are > 20 x ULN, or
- Total bilirubin elevation of > 10 x ULN.

7.6.3.2 Ejection Fraction

The minimal post baseline cardiac ejection fraction and the maximum decrease from baseline will be summarized for each treatment group. Time to maximum decrease from baseline ejection fraction may also be tabulated.

7.6.4 Vital Signs

Vital signs (weight, body temperature, respiratory rate, heart rate, and systolic and diastolic blood pressure) will be listed by subject and visit for each treatment arm.

The frequency and percentage of subjects with post baseline clinically significant vital signs will be summarized. The clinically significant vital signs are defined as: Systolic blood pressure >=120 mmHg or diastolic blood pressure >=80 mmHg; heart rate >100 bpm; In addition, systolic blood pressure >=140 mmHg or diastolic blood pressure >=90 mmHg will also be present; Temperature >=38.0 degrees C (100.4 F); and respiratory rate >20 breaths per min.

For weight, the maximum percent decrease from baseline will also be summarized.

7.6.5 Deaths

Death information will be listed by subject.

7.6.6 Pregnancy

Positive pregnancy test will be listed by subject.

7.7 Pharmacokinetic Analyses

Analysis sets: PK Analysis Set

The analyses described in this section will be produced for the pharmacokinetics analysis set.

Tucatinib and DM1 concentrations will be summarized with descriptive statistics at each PK sampling time point. DM1 PK parameters will be determined by noncompartmental analysis and summarized with descriptive statistics. Additional PK analyses may be performed.

For the calculation of summary statistics, <LLOQ results are imputed to ½ LLOQ value. The summary statistics for a timepoint will not be calculated if more than 50% of the results are <LLOQ.

7.8 Health Economics and Outcomes

7.8.1 Health Care Resource Utilization

Analysis sets: Safety Analysis Set

Cumulative incidence of health resource utilization, including but may not be limited to, length of stay, hospitalizations, and ED visits will be summarized by treatment arm in Safety Analysis Set.

7.8.2 Patient Reported Outcomes (PRO)

Analysis sets: PRO Analysis Set

The PRO assessments defined in protocol section 7.7 will be summarized for the ITT population.

The number and percentage of subjects with an improvement of ≥10 points from baseline in global health status/QoL scale will be summarized by visit. The compliance and completion rate of the PRO assessments will be summarized for each visit. Compliance rate is defined as the proportion of subjects who completed the instrument among those who are expected to complete at a given visit (i.e., subjects started the given cycle). Completion rate is defined as the proportion of subjects who completed the instrument in the PRO analysis set.

Proportion of subjects with each score of EQ-5D-3L will be summarized by treatment arm for each sub-category of anxiety/depression, pain/discomfort, usual activities, mobility, and self-care. Summary statistics of the PRO outcomes will be provided by treatment arm and cycle.

8 INTERIM ANALYSIS

No interim analysis for the primary endpoint is planned. Two interim analyses of OS are planned if the primary analysis for PFS is statistically significant. One interim analysis of PFS.BM is planned if both analyses for PFS and OS are statistically significant. Two interim analyses of OS.BM are planned if all analyses for PFS, OS, PFS.BM and cORR are statistically significant. These interim analyses will be conducted at the time described in Section 6.2 using the ITT analysis set. The rejection boundaries of the interim and final analyses of OS, PFS.BM and OS.BM will be determined using Lan-DeMets spending functions.

9 CHANGES FROM PLANNED ANALYSES

9.1 Changes from the Original SAP

The following changes have been made:

- MedDRA version is updated from 25.0 to 26.0.
- WHO Drug Dictionary version is updated from 2019Mar B3 to 2023Mar B3.
- Table 3 is updated to be consistent with study protocol.
- Definition of liver abnormality is updated to be consistent with study protocol.
- Definition of OS censoring is updated to include survival information post DCO., i.e., if death date or the last long-term follow-up visit (which indicates subject to be alive) date were to be after DCO, then OS will be censored at DCO.

10 REFERENCES

Andersen P, Hansen M, Klein J (2004). Regression Analysis of Restricted Mean Survival Time Based on Pseudo-Observations. Lifetime Data Analysis 10:335–350.

Cheson B, Pfistner B, Juweid M, Gascoyne R, Specht L, Horning S, Coiffier B, Fisher R, Hagenbeek A, Zucca E, Rosen S, Stroobants S, Lister T, Hoppe R, Dreyling M, Tobinai K, Vose J, Connors J, Frederico M and Diehl V (2007). Revised response criteria for malignant lymphoma. Journal of Clinical Oncology 25: 579-586.

Chen LM, Ibrahim JG, Chu H (2014). Flexible stopping boundaries when changing primary endpoints after unblinded interim analyses. J Biopharm Stat 24(4): 817-33.

Collett D (1991). Statistical inference for binary data. Modelling Binary Data. London, Chapman and Hall: 23-26.

Collett D (1994). Interval-censored survival data. Modelling survival data in medical research. Boca Raton, Fla., Chapman & Hall/CRC: 237-251.

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D and Verweij J (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45: 228-47.

Haybittle, JL (1971). Repeated assessments of results in clinical trials of cancer treatment. Br. J. Radiol. 44 (526): 793–797.

(1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design". Br. J. Cancer. 34 (6): 585–612

Royston P and Parmar MK (2011). The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Stat Med 30: 2409-21.

Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, Schrag D, Takeuchi M, Uyama Y, Zhao L, Skali H, Solomon S, Jacobus S, Hughes M, Packer M and Wei LJ (2014). Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol 32: 2380-5.

11 APPENDICES

Appendix A: Imputation of Partial Missing Adverse Event Dates

For an AE with a partial start or end date, if it can be determined that the event occurred prior to the date of first dose of study treatment, the partial date will not be imputed. Otherwise, the partial date will be imputed using the rules described below. AE start dates should be imputed before imputation of AE condition end date in all cases.

Incomplete AE Start Date:

AE day only is missing.

- If the month/year is the same as the month/year of first dose of any study treatment, then AE start date will be imputed as the first dose date of any study treatment.
- If the month/year is after the month/year of first dose of any study treatment, then AE start date will be imputed as the first day of the month.

AE day and month are missing.

If the year is the same as the year of first dose of any study treatment, then AE start date will be imputed as the first dose date of any study treatment.
If the year is after the year of first dose of any study treatment, then AE start date will be imputed as January 1st.

AE day, month and year are missing.

AE start date will be imputed as the first dose date of any study treatment.

If AE condition end date is known with a full date, and the imputed start date is after the end date, the start date will be set to the AE end date.

Incomplete AE End Date:

If AE outcome is "not recovered/resolved", "unknown", or blank, then AE condition end date will not be imputed.

If AE outcome is "recovering/resolving", "recovered/resolved", "recovered/resolved with sequelae", or "fatal", then apply the following:

AE day only is missing.

AE condition end date will be imputed as the minimum of (death date, DCO date or data extraction date sans DCO, last day of the AE end date month/year, EOS date).

AE day and month are missing.

- If the year is equal to the year of the last dose date, then AE end date will be imputed as the minimum of (last dose date + 30, death date, DCO date or data extraction date sans DCO, December 31st of the end date year, EOS date).

- If the year is not equal to the year of the last dose date, then AE end date will be imputed as the minimum of (death date, DCO date or data extraction date sans DCO, December 31st of the end date year, EOS date).

AE day, month and year are missing.

AE end date will not be imputed.

Within a single record, if the imputed end date is before the start date, then the imputed end date will be set to the start date.

Example

AE Number 4: Condition/Event NAUSEA
First dose date 02APR2012

Prior to imputation

Log Line	Start date	Condition end date	Severity	Outcome
1	25APR2012	UNAPR2012	2	recovering/resolving
2	UNAPR2012	04MAY2012	1	recovered/resolved

Post imputation

Log Line	Start date	Condition end date	Severity	Outcome
1	25APR2012	**30**APR2012	2	recovering/resolving
2	**02**APR2012	04MAY2012	1	recovered/resolved

Appendix B: Definition of the Term "Treatment-Emergent" with Respect to AE Classification

A treatment-emergent adverse event (TEAE) is defined as any AE which is newly occurring or worsening in severity, with starting date on or after the first dose of any study treatment (i.e., tucatinib or placebo, or T-DM1) and before the last dose of study treatment + 30 days.

Appendix C: RECIST Version 1.1 Summary

Response Evaluation Criteria in Solid Tumors

Term	Definition
Complete response (CR)	Disappearance of all target lesions. Any pathological lymph nodes must have reduction in short axis to <10 mm.
Partial response (PR)	A ≥30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.
Progressive disease (PD)	At least a 20% increase in the sum of diameters of target lesions, 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. The appearance of one or more new lesions is also considered progression.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.
Measurable lesion	Must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of 10 mm by CT scan (CT slice thickness no greater than 5 mm).

From RECIST Version 1.1

Appendix D: Imputation of Partial Missing Start and End Date of Prior Systemic Therapy, Subsequent Cancer-Related Therapy, and Hospitalization

Partial missing date will be imputed when both month and year are present and only day is missing.

- Prior therapy start date: first day of the month
- Prior therapy end date: the earlier of
    - Last day of the month and year
    - First dose date of study drug
- Subsequent therapy and hospitalization start date: the later of
    - First day of the month and year
    - First dose date of study drug
- Subsequent therapy and hospitalization end date: the earlier of
    - Last day of the month
    - Study end date

Appendix E: Imputation of Partial Missing Date of Disease Diagnosis

Disease diagnosis date (including initial diagnosis, earliest date disease was considered unresectable locally advanced and earliest date disease was considered distant metastatic) will be imputed as the 1st of the month if both month and year are present and only day is missing, and will be imputed as 01 January if only year is present and both day and month are missing.

Appendix F: Imputation of Partial Missing Date of the End of Tucatinib/Placebo Administration

The end date of tucatinib/placebo administration, will be imputed as the later of the following dates,

- The start date of the log line where the end date is partially missing.
- The earlier of data cut-off date and the last day of the year or month/year of the partial date.

Appendix G: Imputation of Partial Missing Death Dates

Death dates are imputed if only the day is missing.

The imputation of partial missing death date depends on the last-known-alive date derived from eCRF.

If the last-known-alive date is in the same month and year of the partial missing death date, then the partial missing death date is imputed as the later of the following dates,

The last-known-alive date.
Day 15 of the month and year.

If the last-known-alive date is not in the same month and year of the partial missing death date, then the partial missing death date is imputed as day fifteen of the month and year.

Appendix I: Imputation of Partial Missing First and Last Dates of Corticosteroid Administration

Partial missing date will be imputed when both month and year are present and only day is missing.

First date of corticosteroid administration will be imputed,

- If started prior to first dose of any study treatment, as the first day of the month and year.
- If started not prior to first dose of any study treatment, as the later of the following dates
    - The date of first dose of any study treatment.
    - The first day of the month and year.

Last date of corticosteroid administration will be imputed as the earlier of the following dates,

- The last day of the month and year.
- Study end date.

Appendix J: Statistical Analysis Plan for the China Population

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

Signature Page for SGN-CLIN-025024 v1.0

eSignature Approval	[redacted]3 19:54:25 GMT+0000
eSignature Approval	[redacted]8 GMT+0000
eSignature Approval	[redacted]7 GMT+0000
eSignature Approval	[redacted]3 14:24:56 GMT+0000

Signature Page for SGN-CLIN-025024 v1.0