Official Title	AN OPEN-LABEL, PHASE 2 BASKET STUDY OF NERATINIB IN PATIENTS WITH SOLID TUMORS WITH SOMATIC ACTIVATING *HER* MUTATIONS
NCT Number	NCT01953926
Date	13-DEC-2022

STATISTICAL ANALYSIS PLAN

Protocol Title:	AN OPEN-LABEL, PHASE 2 BASKET STUDY OF NERATINIB IN PATIENTS WITH SOLID TUMORS WITH SOMATIC ACTIVATING *HER* MUTATIONS
Study Protocol No.	PUMA-NER-5201
Disease Condition	Solid tumors harboring somatic activating *HER* mutations
Sponsor	Puma Biotechnology, Inc. 10880 Wilshire Blvd, Suite 2150 Los Angeles, CA 90024 USA Phone: +1 424.248.6500 Fax: +1 424.248.6501
Date of SAP	V1.0 13-DEC-2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
ATC	Anatomical Therapeutic Chemical [Classification System]
CB	clinical benefit
CBR	clinical benefit rate
CR	complete response
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of response
ECHO	echocardiogram
ECOG	Eastern Cooperative Oncology Group
ECG	electrocardiogram
ERBB2	human epidermal growth factor receptor 2
EOS	end of study
EOT	end of treatment
GLP	good laboratory practice
HER2	human epidermal growth factor receptor 2
HER3	human epidermal growth factor receptor 3
IDMC	independent data monitoring committee
ITT	intent to treat
LVEF	left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
mPERCIST	modified PERCIST
MUGA	multiple-gated acquisition scan
ORR	overall objective response rate
ORR_{first}	objective response rate at the first tumor assessment
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PR	partial response

Abbreviation	Definition
PT	preferred term
PERCIST	PERCIST
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SOC	system organ class
TEAE	treatment-emergent adverse event
WHODrug	World Health Organization Drug Reference List

1. PURPOSE OF THE ANALYSIS

This statistical analysis plan (SAP) provides a comprehensive and detailed description of the strategy and statistical methodology to be used for analysis of data for the 5201 study. This study is sponsored by Puma Biotechnology, Inc. The purpose of the SAP is to ensure the credibility of the study findings by pre-specifying the statistical approaches to the analysis of study data prior to database lock. This analysis plan is meant to supplement the study protocol. Any deviations from this plan will be described in the Clinical Study Report (CSR).

2. PROTOCOL SUMMARY

2.1. Study Objectives

The objectives of this study, applicable to each cohort, are:

For the randomized hormone receptor positive (HR+), *HER2* negative metastatic breast cancer:

Primary:

- To determine the confirmed objective response rate (ORR) by independent central review according to Response Evaluation Criteria in Solid Tumors (RECIST) v1.1

Secondary:

- To determine the confirmed ORR by investigator
- To determine the duration of response (DOR) by both independent central review and investigator
- To determine the clinical benefit rate (CBR) by both independent central review and investigator
- To determine the progression-free survival (PFS) by both independent central review and investigator

For the metastatic cervical cancer cohort:

Primary:

- To determine the confirmed ORR by independent central review according to RECIST v1.1

Secondary:

- To determine the confirmed ORR by investigator
- To determine the DOR by both independent central review and investigator
- To determine the CBR by both independent central review and investigator
- To determine the PFS by both independent central review and investigator
- To determine overall survival (OS)

For all other cohorts:

Primary:

- To determine the first objective response rate (ORR_{first}) by investigator at the first post-baseline tumor assessment

Secondary:

- To determine the confirmed ORR by investigator

- To determine the DOR by investigator
- To determine the CBR by investigator
- To determine the PFS by investigator
- To determine OS

Safety Objectives:

For all cohorts, including the randomized HR+, *HER2* negative metastatic breast cancer and metastatic cervical cancer cohorts:

- To assess the safety profile and tolerability of study treatments
- To assess Patient Reported Outcomes (PRO)

Exploratory Objectives:

- To collect and retrospectively evaluate somatic mutations or gene aberrations using next-generation sequencing (NGS) in the most recent pretreatment tumor biopsy or fresh tumor tissue biopsies at a central laboratory.
- To explore genetic modifiers of sensitivity and/or resistance to neratinib using molecular profiling techniques in pretreatment archival and/or fresh tumor specimens and paired normal whole blood.
- To evaluate cell-free DNA (cfDNA) from plasma specimens collected at baseline/screening, during the course of treatment, and upon disease progression to identify *HER* mutations and other gene aberrations and to assess any potential associations with neratinib sensitivity and/or primary/acquired resistance to neratinib or neratinib-containing therapy.
- To evaluate potential genes or protein biomarkers that may be reported to confer neratinib sensitivity and/or primary/acquired resistance to neratinib or neratinib-containing therapy from *optional* fresh core tumor biopsies during time of treatment and/or at the time of treatment discontinuation or disease progression.

2.2. Overall Study Design and Plan

2.2.1. Study Design

This is an open-label, multicenter, multinational, Phase 2 study exploring the efficacy and safety of neratinib as monotherapy or in combination with other therapies in patients with *HER* (*EGFR*, *HER2*) mutation-positive solid tumors. Patients with tumors harboring somatic mutations in *HER* will be identified through previously documented mutation testing performed prior to screening. The presence of human *HER* mutations (*EGFR*, *HER2*) will be retrospectively

confirmed by central testing via next generation sequencing (NGS). The study has a basket design and includes several cohorts, either defined by an actionable somatic mutation or by actionable mutation and tumor histology (for example *HER2* mutant cervical cancer). In the course of the study, enrollment in certain cohorts can be completed and enrollment in new cohorts initiated to test the anticancer effect of neratinib in other histologies, specific molecular abnormalities, and/or in combination with other drugs.

Patients with *HER2* mutant, *HER2* negative breast cancer will be divided into different treatment cohorts on the basis of HR status: HR negative (Triple Negative Breast Cancer [TNBC]) and HR+.

- Patients in the HR negative cohort (TNBC) will receive neratinib in combination with trastuzumab.
- Patients in the HR+, *HER2* negative, *HER2* mutant positive cohort with RECIST measurable tumors and who have been previously been treated with CDK4/6 inhibitors (CDK4/6i) will be randomized to receive single agent fulvestrant, fulvestrant in combination with trastuzumab, or neratinib in combination with trastuzumab and fulvestrant with a randomization ratio of 1:1:1. Randomization will be stratified by the number of lines of prior therapy for metastatic disease (≤ 2 and >2 lines) and by prior fulvestrant therapy.

  The randomized HR-positive (HR +), *HER2* negative, *HER2* mutant metastatic breast cancer cohort is designed to investigate the individual contribution of neratinib and/or trastuzumab to fulvestrant via Simon's 2 stage optimal design. The decision to carry over on enrollment of each arm in the cohort will be based on stage I and II analysis of Simon's 2 stage in consultation with the Independent Data Monitoring Committee (IDMC).

  - Patients in the HR+ cohort who receive single agent fulvestrant or fulvestrant plus trastuzumab will be eligible for triplet therapy (neratinib, fulvestrant, trastuzumab) upon progression. Efficacy will be determined based on response to initial regimen only.
- Patients in the HR+ cohort who have not received prior CDK4/6i therapy (e.g. CDK4/6i naive) will receive neratinib in combination with trastuzumab and fulvestrant as part of an open label cohort.

The trial will consist of a screening period, a treatment period, and a follow-up period after the study therapy is discontinued for any reason. An end of treatment (EOT) assessment is

performed 28 days (+14 days) after the last dose of investigational product(s) and adverse events are collected 28 days after the last dose of investigational product(s).

Neratinib will be administered orally with food once daily (recommended to be taken in the morning), on a continuous basis. Dose delays and modifications will be handled as per instructions in the package insert. All patients taking neratinib will maintain a patient diary for the study to record each dose of neratinib taken and while receiving antidiarrheal prophylaxis with loperamide taken for the first two cycles of treatment. For cohorts receiving combination treatment that includes trastuzumab, post-treatment radiographic evaluation of their disease will be conducted every 3 cycles. For all other cohorts, post-treatment disease assessment will be conducted every 2 cycles. Treatment should be administered until there is unequivocal evidence of disease progression as documented by imaging, clinical examination, or disease-related symptoms. Tumor markers or circulating tumor cells should not be used as the sole criteria for determining progression. At the time of decision regarding disease progression, radiological exam should be performed. Patients will continue study treatment until disease progression, unacceptable toxicity, patient withdrawal of consent, or death. Patients who develop disease progression, but in the opinion of the Investigator would still benefit from continuing study, may continue per-protocol therapy if approved by the Sponsor. Survival follow-up will be every 12 weeks after treatment discontinuation. Enrollment will continue as dictated by the Simon's 2-stage design in all the histology and mutation specific cohorts and/or up to 30 patients per cohort in multi-cancer, mutation specific NOS cohorts. Enrollment in the randomized breast cancer, non-randomized breast cancer and cervical cancer cohorts will also continue as dictated by the Simon's 2-stage design up to 50 patients per cohort.

Tumor Cohorts Open to Enrollment in Amendment 7

Randomized Breast Cancer Cohort

Tumor Cohort	Mutation	Randomized Treatment	Sample Size (# subjects)
Breast HR Positive (with prior CDK4/6i)	*HER2* mutant	• Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter Or • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks Or • Neratinib: 240 mg daily • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 18 Up to 18 Up to 50

Non-randomized Breast Cancer Cohorts

Tumor Cohort	Mutation	Assigned Treatment	Cohort Sample Size (# subjects)
Breast HR Positive (CDK4/6i naïve)	*HER2* mutant	• Neratinib: 240 mg daily • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 50
Breast TNBC (HR Negative)	*HER2* mutant	• Neratinib: 240 mg daily • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 50

Cervical Cancer Cohort

Tumor Cohort	Mutation	Assigned Treatment	Cohort Sample Size (# subjects)
Cervical	*HER2* mutant	• Neratinib: 240 mg daily	50

Cohorts Not Receiving Trastuzumab

Tumor Cohort	Mutation	Assigned Treatment
Salivary Gland	*HER2* mutant	• Neratinib: 240 mg daily
Non-small Cell Lung	*EGFR* exon 18 mutations	• Neratinib: 240 mg daily

Tumor Cohorts Closed to Enrollment in Amendment 7

The following cohorts were closed to enrollment in Amendment 7: solid tumors (NOS) *HER2* mutant monotherapy and bladder/urinary *HER2* mutant combination therapy of neratinib + paclitaxel.

Tumor Cohorts Open to Enrollment in Amendment 6

Randomized Breast Cancer Cohort

Tumor Cohort	Mutation	Randomized Treatment	Sample Size (# patients)
Breast HR Positive (with prior CDK4/6i)	*HER2* mutant	• Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter Or • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks Or • Neratinib: 240 mg daily • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 18 Up to 18 Up to 50

Non-randomized Breast Cancer Cohorts

Tumor Cohort	Mutation	Assigned Treatment	Cohort Sample Size (# patients)
Breast HR Positive (CDK4/6i naive)	*HER2* mutant	• Neratinib: 240 mg daily • Fulvestrant: 500 mg on Study Day 1, 15, and 29; once every 28 days thereafter • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 50
Breast TNBC (HR Negative)	*HER2* mutant	• Neratinib: 240 mg daily • Trastuzumab: 8 mg/kg IV followed by 6 mg/kg IV every 3 weeks	Up to 50

Cervical Cancer Cohort

Tumor Cohort	Mutation	Assigned Treatment	Cohort Sample Size (# patients)
Cervical	*HER2* mutant	• Neratinib: 240 mg daily	50

Cohorts Not Receiving Trastuzumab

Tumor Cohort	Mutation	Assigned Treatment
Salivary Gland	*HER2* mutant	• Neratinib: 240 mg daily
Lung	*EGFR* exon 18 mutations	• Neratinib: 240 mg daily
Bladder/ Urinary Tract	*HER2* mutant	• Neratinib: 240 mg daily • Paclitaxel: 80 mg/m^2 IV on Days 1, 8, and 15 of every 4-week cycle
Solid tumors (NOS)	*HER2* mutant	• Neratinib: 240 mg daily

Tumor Cohorts Closed to Enrollment in Amendment 6

The following cohorts were closed to enrollment in Amendment 6: colorectal cancer combination therapy with neratinib + trastuzumab, lung cancer *HER2* mutant combination therapy with neratinib+ trastuzumab, gastroesophageal cancer monotherapy, biliary cancer monotherapy, ovarian cancer monotherapy, and solid tumor (NOS) *HER4* mutant.

If a tumor has more than one qualifying mutation, the patient will be assigned to the appropriate tumor-specified cohort if one exists upon consultation with the Investigator and Sponsor.

Tumor Cohorts Closed to Enrollment in Other Amendments

Refer to the protocol for cohorts open or closed in prior amendments.

2.2.2. Investigational Product, Dose, and Administration:

- Neratinib: Patients will receive six 40-mg tablets (total daily dose 240 mg) administered orally, once daily with food (recommended to be taken in the morning), continuously.
- Fulvestrant: Patients will receive 500 mg total dose administered as two 5 mL injections, by intramuscular injection, one in each buttock on Days 1, 15, and 29; then once every 4 weeks thereafter; see fulvestrant package insert and Schedule of Procedures.
- Trastuzumab: Patients will receive an initial dose of 8 mg/kg of trastuzumab intravenously (IV) administered on Cycle 1 Day 1 (C1D1), followed by 6 mg/kg IV once every 3 weeks thereafter; see Herceptin (trastuzumab) package insert and Schedule of Procedures.
- Paclitaxel: Patients in the bladder/urinary tract cancer cohort treated with combination therapy will receive 80 mg/m2 administered IV on Days 1, 8, and 15 of every 4-week cycle.

3. SAMPLE SIZE DETERMINATION

3.1. Sample Size Justification for the Randomized HR+ Breast Cancer Study Cohort:

A Simon's 2-stage optimal design (Simon, 1989) will be used to determine whether there is sufficient activity to warrant further development of the therapy and to minimize the number of patients exposed to therapy if ineffective. For each treatment arm, using Simon's 2-stage optimal design (with significance level 10% and power of 80%), an ORR (confirmed) of 10% or less per RECIST by independent assessment will be considered unacceptable (null hypothesis) whereas an ORR (confirmed) of 30% per RECIST by independent assessment will merit further study (alternative hypothesis). In the first stage, 7 patients are enrolled in each treatment arm. If at least 1 response is observed in the first stage, the second stage will be opened. In the second stage, 11 additional response evaluable patients will be accrued and randomized for a total of 18 patients in the treatment arm. The null hypothesis will be rejected (for each arm separately) if at least 4 responses are observed in Stage 2 for each arm.

Once the Simon's 2-stage criteria are met, enrollment of the neratinib arm may continue until up to 50 patients have been enrolled.

Prior to protocol amendment 6, there were approximately 10 HR+, HER2- patients in the N+F+T cohort that received prior CDK4/6i therapy; therefore, data can be pooled with the data from 50 patients randomized in N+F+T arm in protocol amendment 6. The 2-sided Clopper-Pearson 95% CIs for the possible ORRs are summarized in the following table for total of 60 patients:

Table 1: Two-sided Clopper-Pearson 95% CI for ORR

ORR	95% CI
25%	(14.7%, 37.9%)
30%	(18.8%, 43.2%)
35%	(23.1%, 48.4%)
40%	(27.6%, 53.5%)
45%	(32.1%, 58.4%)
50%	(36.8%, 63.2%)

3.2. Sample Size Justification for All Other Cohorts:

A Simon's 2-stage optimal design (Simon, 1989) will be used to determine whether neratinib monotherapy has sufficient activity to warrant further development in the following cohorts: cervical, salivary gland, NOS, and lung cancers. A similar Simon's 2-stage design will be used to determine whether neratinib combination therapy has sufficient activity to warrant further development in the following cohorts: HR+ and CDK4/6i naive, TNBC, and bladder/urinary tract cancers. Early study termination will be permitted if data at the first stage indicate that the

treatment is ineffective. For each cohort, using Simon's optimal 2-stage design (with significance level 10% and power of 80%), an ORR_{first} of 10% or less will be considered unacceptable (null hypothesis) whereas an ORR_{first} of 30% will merit further study (alternative hypothesis). In the first stage, enrollment will continue until 7 patients received at least one dose of study treatment and completed the first tumor assessment by the investigator (response evaluable). If no responses are observed, the second stage for the cohort must be discontinued. Otherwise, 11 additional response evaluable patients will be accrued for a total of 18 patients in the cohort. The null hypothesis will be rejected (for each cohort separately) if at least 4 responses are observed in Stage 2 for each cohort.

Enrollment to the salivary gland, lung (*EGFR* exon 18), bladder/urinary tract and *HER2*-mutant NOS cohort may enroll up to 30 patients.

Once the Simon's 2-stage criteria are met, enrollment into these cohorts may continue until up to 50 patients have been enrolled. A new cohort may also be opened separately at any time per Sponsor discretion and follow the Simon's 2-stage criteria. Cohorts may close prior to planned enrollment.

4. GENERAL ANALYSIS AND REPORTING CONVENTIONS

In general, efficacy analyses in this study are meant to be cohort-specific analyses. Safety analyses will be summarized across all monotherapy and combination therapy where appropriate. For patients who develop disease progression and start combination therapy, the safety on combination therapy will be summarized separately starting from the start time of combination therapy.

Categorical variables will be summarized using counts and percentages. Percentages will be displayed to 1 place after the decimal point (xx.x), with the exception of 100%, which will be displayed without additional decimal places, and with the exception of 0%, which will not be displayed. Continuous variables will be summarized using number of patients, mean, median, standard deviation, and interquartile range (25^{th} – 75^{th} percentile), minimum, and maximum.

In general, the baseline value will be considered the last measurement observed prior to taking the first dose of study treatment. In the case more than one measurement on the same day is available for a patient/assessment, the average will be considered as baseline.

SAS statistical software, version 9.4 or later, will be used for all analyses.

If departures from these general conventions are present in the specific evaluations section of this SAP, then those conventions will take precedence over these general conventions.

4.1. Date Imputations

4.1.1. Missing/Partial Dates in Adverse Events

Adverse events with start dates that are completely or partially missing will be analyzed as follows:

1. If the start date has a month and year but the day is missing, the event will be considered treatment emergent if the month and year of the start date of the event are:
    a. On or after the month and year of the date of the first dose of study drug
2. If the start date has a year, but the day and month are missing, the event will be considered treatment emergent if the year of the start date of the event is:
    a. On or after the year of the date of the first dose of study drug and
    b. On or before the year of the date of the last dose of study drug plus 28 days
3. If the start date of an event is completely missing then the event is assumed to be treatment emergent.

4.1.2. Missing/Partial Dates in Concomitant Therapies

Concomitant therapies with start dates that are completely or partially missing will be analyzed as follows:

1. If the start date has a month and year but the day is missing, the event will be considered concomitant if the month and year of the start date of the event are:
    - On or after the month and year of the date of the first dose of study drug and
    - On or before the month and year of the date of the last dose of study drug plus 28 days
2. If the start date has a year, but the day and month are missing, the event will be considered concomitant if the year of the start date of the event is:
    - On or after the year of the date of the first dose of study drug
3. If the start date of an event is completely missing then the event is assumed to be concomitant.

When the start date is complete and is before the first dose, and the concomitant medication is not ongoing but the end date is missing completely or partially, a similar algorithm should be used to assess whether the end date is before the last dose of study drug plus 28 days to be included.

4.1.3. Missing/Partial Dates in Subsequent Therapies

Subsequent therapies with start dates that are completely or partially missing will be analyzed as follows:

1. When the month and year are present and the day is missing:
    a. If the onset month and year are the same as the month and the year of last dose of study drug, the day of the last dose + 1 will be imputed.
    b. If the onset month and year are not the same as the month and year of the last dose of study drug, the first day of the month will be imputed.
2. When only a year is present:
    a. If the onset year is the same as the year of the last dose of study drug, the date of the last dose + 1 will be imputed.
    b. If the onset year is not the same as the year of the last dose of study drug, the first day of the year will be imputed.

3. If no components of the onset date are present the date of the last dose + 1 will be imputed.

5. ANALYSIS POPULATIONS

For the purposes of patient disposition, intent-to-treat population (ITT) is defined as all patients who are enrolled into the study. The safety population are all patients who received at least one dose of study drug.

The Primary Analysis Population is defined as all patients who received at least 1 dose of study treatment. This population will be used for all efficacy and safety analysis, if not otherwise specified.

6. STUDY PATIENTS

6.1. Disposition of Patients

The number and percentage of patients entering and completing the study will be presented. Reasons for withdrawal of treatment and study will also be summarized by treatment group.

6.2. Protocol Deviations

Protocol deviations will be classified and monitored regularly during the duration of the study. Among other reasons, failure to meet any of the protocol inclusion or exclusion criteria will be considered a protocol deviation. All protocol deviations will be listed and summarized by type.

7. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographic data, medical history, and concomitant medication will be summarized by means of descriptive statistics or frequency tables. For categorical variables, the number and percentage of patients in each category will be presented. For continuous variables, summaries will include the number of patients with data, mean, median, standard deviation, minimum, and maximum. All summaries will be based on the safety population.

7.1. Demographic and baseline characteristics

The following demographic and baseline variables will be summarized by treatment group:

- age (years)
- age group (<65 years, ≥ 65 years)
- sex
- race and ethnicity (per CRF)
- height (cm)
- weight (kg)
- body mass index (BMI)

Body mass index (BMI) will be calculated as: $BMI\left(\frac{kg}{m^2}\right) = \frac{Weight\ (kg)}{(Height(cm)*0.01)^2}$

7.2. Medical History

Medical history data including: chronic conditions, relevant surgical procedures, symptoms experienced during the previous 30 days, symptoms ongoing at the time of screening, any medical conditions that require medication and cancer history will be collected at screening, within 14 days prior to Cycle 1/ Day1 in accordance with the Schedule of Procedures included in the protocol.

Cancer history variables include date of first diagnosis, nodal status, histology, tumor stage at diagnosis, previous chemotherapy/biotherapy/ immunotherapy, previous adjuvant therapy, previous radiation, and prior cancer related surgical therapies.

Medical history and cancer history data will be summarized and listed; tabulations will be by treatment group.

7.3. Prior and Concomitant Medications

Concomitant medications will be defined as medications documented on the Concomitant Medications CRF. Concomitant medications will be coded using the World Health Organization (WHODrug) dictionary, v2022-09, and summarized in a table and a data listing.

8. MEASUREMENTS OF TREATMENT COMPLIANCE

Patient dose, exposure, and duration of neratinib will be tabulated. Exposure to other study drugs will be listed.

Duration: number of days from the first dose of study drug to the last dose of study drug

Cumulative actual dose (cumdose) : total dose taken, in mg.

Dose intensity (DI): the cumulative dose divided by the duration, or

$$DI = \frac{cumdose}{duration}$$

Relative dose intensity (RDI) is the cumulative dose divided by the expected dose, which is 240 mg*duration, or

$$RDI = \frac{cumdose}{240 * duration} * 100\% .$$

9. EFFICACY EVALUATION

For cohorts receiving combination treatment that includes trastuzumab, post-treatment disease assessment by CT or magnetic resonance imaging (MRI) will be conducted at the beginning of Week 10 (ORR_{first}), and every 3 cycles (± 7 days) thereafter (see Schedule of Procedures, Appendix 1 Table A1.1). For all other cohorts not receiving trastuzumab, post-treatment disease assessment will be conducted at the beginning of Week 9 (ORR_{first}), and every 2 cycles (± 7 days) thereafter (see Schedule of Procedures, Appendix 1 Table A1.2). Complete or partial response (CR or PR) must be confirmed with a repeat scan performed no sooner than 4 weeks after the criteria for response are first met. In cases where the subject discontinues treatment for reasons other than progressive disease (e.g., adverse event, patient choice, noncompliance, etc.) and the response is either CR or PR, then a confirmation scan is required no sooner than 4 weeks after the criteria of response is met. Radiological response is assessed by RECIST v1.1. Following one year on therapy, scans may move to every 4 cycles for trastuzumab-containing regimens and every 3 cycles for non-trastuzumab containing regimens. Following two years on therapy, scans may move to every 5 cycles for trastuzumab-containing regimens and every 4 cycles for non-trastuzumab-containing regimens. Tumor assessments will be performed by the study investigators for all patients, and by the independent central review for the patients in the HR+, HER2- cohort of patients who received a prior CDK4/6 inhibitor, the randomized HR+, HER2- mBC cohort, and mCC cohorts.

Table 2 gives an overview of the efficacy variables and the analysis methods that will be implemented.

Table 2: Efficacy Variables and Analysis Methods

Efficacy Variables	Analysis Methods
Objective Response Rate at the first tumor assessment (ORR_{first})	Estimate with associated 2-sided Clopper-Pearson 95% confidence intervals
Confirmed Overall Objective Response Rate (ORR)	Estimate with associated 2-sided Clopper-Pearson 95% confidence intervals
Clinical Benefit Rate (CBR)	Estimate with associated 2-sided Clopper-Pearson 95% confidence intervals
Progression-Free Survival (PFS)	Median PFS estimated via Kaplan-Meier with associated 2-sided 95% confidence intervals
Overall Survival (OS) [a]	Median OS estimated via Kaplan-Meier with associated 2-sided 95% confidence intervals

Duration of Response (DOR)	Median DOR estimated via Kaplan-Meier with associated 2-sided 95% confidence intervals

[a]Long-term follow up was discontinued January 4, 2022 and patients will not be followed for OS. This analysis will not be conducted.

9.1. Definitions of Efficacy Endpoints

9.1.1. Objective Response Rate at the First Tumor Assessment (ORR_{first})

ORR_{first} is defined as the proportion of patients who achieve complete responses (CR) or partial responses (PR), per RECIST (v1.1) or other defined response criteria at the first scheduled tumor assessment.

9.1.2. Overall Response Rate (ORR)

ORR is defined as the proportion of patients who achieve confirmed complete responses (CR) or partial responses (PR), per RECIST (v1.1) or other defined response criteria, as their best overall response. Complete or partial responses are confirmed with repeat tumor evaluation using same criteria after 4 weeks.

9.1.3. Clinical Benefit Rate (CBR)

The CBR is defined as the proportion of patients who achieve overall tumor response (CR or PR) or SD for at least 16 weeks (at least 24 weeks for the breast cancer cohorts). To allow for visits occurring earlier than scheduled, 105 days and 161 days define 16 weeks and 24 weeks, respectively.

9.1.4. Progression-Free Survival (PFS)

Progression-free survival (PFS) is defined as the time interval from the date of C1D1 until the first date on which recurrence, progression or death due to any cause is documented, censored at the last assessable evaluation, or at the initiation of new anticancer therapy. It is not necessary to confirm disease progression.

PFS is censored on the date of the last tumor assessment on study for patients who did not have PD and who did not die while on study or who started a new anti-cancer therapy prior to documented PD. Additionally, patients lacking an evaluation of tumor response after enrollment have their PFS time censored on the date of enrollment with duration of 1 day. Finally, those patients who have not

progressed by the time of database lock will have their PFS time censored at the time of their last assessment. The table describes the various censoring mechanisms for the various cohorts.

Table 3: Handling of Missing Response Assessment and Censoring for Progression-Free Survival Primary Analysis

Situation	Date of Progression or Censoring	Outcome
No baseline and/or no post baseline assessment	1. Randomization for the randomized cohorts. 2. Treatment start date for other cohorts.	Censored
Disease progression documented between scheduled visits	Date of disease progression	PFS event
No documented disease progression or death	Date of last adequate assessment	Censored
Progressive disease or death immediately after more than one consecutively missed tumor assessment visit	Date of last adequate assessment	Censored (see note below)
Alternate antineoplastic therapy started prior to disease progression	Date of last adequate assessment prior to the start of subsequent antineoplastic therapy	Censored
Death before first assessment	Date of death	PFS event
Death between adequate assessment visits	Date of death	PFS event

Note: Two consecutive missed tumor assessments are defined by the following criteria:

- For all patients enrolled before amendment 6, or if the latter assessment occurs during year 1 (day ≤ 365) and the patient is enrolled under amendment 6:
    - For patients receiving trastuzumab, the number of days between the 2 visits is greater than 133 days
    - For patients not receiving trastuzumab, the number of days between the 2 visits is greater than 119 days
- For patients enrolled under amendment 6 or 7
    - if the latter assessment occurs between day 365 and 730, the number of days between the 2 visits is greater than 175 days
    - if the latter assessment occurs after 730, the number of days between the 2 visits is greater than 231 days

9.1.5. Duration of Response (DOR)

Duration of response is measured from the time at which measurement criteria are first met for CR or PR (whichever status is recorded first) until the first date of recurrence or progressive disease (PD) or death is objectively documented. Censoring rules are similar to those described in section 9.1.4.

9.2. Analysis Methods

Estimates of ORR_{first} , ORR and CBR will be determined and their associated 2-sided Clopper-Pearson 95% confidence intervals will be derived for each cohort. Patients who drop out prior to the first tumor assessment will be considered non-responders.

Kaplan-Meier methods will be used to estimate PFS and DOR. Median PFS and DOR will be derived from the resulting product limit estimator. The associated 2-sided 95% confidence intervals for median PFS and median DOR will be determined via methods outlined by Brookmeyer and Crowley (1982). Summary tables, by specific tumor type, of the efficacy data will be generated.

In various cohorts, tumor data were collected for patients in accordance with RECIST, (modified) PERCIST, or both. For those patients whose assessments were made with RECIST, then RECIST criteria were used to define the efficacy endpoints. For those patients whose assessments were made with

(modified) PERCIST only then the PERCIST assessments were the criteria for response, progression, and benefit.

9.3. Examination of Subgroups

Analyses of subgroups are not currently planned for this study, excepting that efficacy analyses will be generated by cohorts.

10. SAFETY EVALUATION

10.1. Overview of Safety Analysis Methods

All safety analysis will be performed for all patients in the Primary Analysis Population. The following assessments will be used to evaluate the safety of neratinib:

- Adverse events (AEs)
- Vital sign measurements
- Physical examination findings
- Electrocardiogram (ECG)
- LVEF results from multiple-gated acquisition (MUGA) scan or echocardiogram (ECHO)
- Laboratory assessments.

All safety endpoints will be summarized by treatment, and all treated patients when appropriate.

10.2. Extent of Exposure

Extent of exposure to neratinib will be summarized by total dose, dose intensity, and duration of therapy for each treatment group. Refer to section 8 for the definitions.

10.3. Adverse Events, Serious Adverse Events, and Deaths

Adverse events and serious adverse events will be coded using MedDRA v. 25.1 and tabulated by system organ class (SOC) and preferred term (PT). All AEs will be graded by the Investigator according to the NCI CTCAE v.4.0. Summaries will in general focus on treatment emergent adverse events (TEAEs). A TEAE is any adverse event that occurs or worsens on or after first dose of investigational product and up to 28 days after the last dose.

Patient incidence of all TEAEs, SAEs, treatment related TEAEs, treatment related SAEs, TEAEs leading to IP changes, grade 3 or 4 TEAEs, and fatal AEs will be tabulated by SOC and PT, for each treatment group.

Patients on neratinib monotherapy may have, upon disease progression, begun taking another drug in combination with neratinib. Similarly, patients randomized to F or F+T may have crossed over to receive N+F+T. To address these situations, adverse events are summarized three ways:

1. A summary of adverse events that occurred only during the time before combination therapy or crossover. The treatment groups in the table will be the treatment group with which the patient started in the first period, which will be noted as PERIOD 1.
2. For those patients who crossed over to a different therapy, a summary of adverse events that started after the crossover. This will consist of the patients on neratinib monotherapy who crossed over to combination therapy, and the F, and F+T patients who crossed over to N+F+T, and will be presented by the treatment group in the second period, which will be noted as PERIOD 2.
3. A summary of adverse events that occurred anytime during the study.
    a. For neratinib monotherapy patients the patients first treatment group received
    b. For F and F+T patients, events before crossover will be attributed to period 1, and events after crossover will be attributed to period 2.

For each treatment group, AEs, Serious Adverse Events (SAEs) and Treatment Emergent Adverse Events (TEAEs) will be listed in by patient listings sorted by patient ID and study day; SAEs will be flagged. All AE listings will include study day, PT, reported term, dose, AE onset date, AE resolution date, outcome, relationship to drug, action taken, and severity.

Patient deaths are recorded on the End of Study CRF page. Cause of death and the time of death (dichotomized as within 28 days of last dose vs. more than 28 days after last dose) will be summarized via frequencies and percentages. Patient death listings will include all death data available including the date of death, cause of death, and any AEs resulting in death.

10.4. Clinical Laboratory Evaluation

Blood samples for clinical chemistry and hematology will be collected during screening, on day 1 of each treatment cycle (with the exception of cycle 1/day 1 if screening occurred within 72 hours of cycle 1/day 1), and at the treatment discontinuation visit. Samples for pregnancy testing and urinalysis will be collected at screening, and directly before study drug administration.

Descriptive statistics will be calculated on both the actual value and the change from baseline value at the minimum, maximum, and last value for each patient and lab test.

The institutional laboratory will analyze all hematology, routine blood chemistry, and urine samples collected.

10.5. Hepatotoxicity

Abnormalities in liver function tests will be summarized. Patient incidence of the following will be provided:

- 3x-, 5x-, 10x-, and 20xULN elevations of AST, ALT, and either ALT or AST.
- Any elevations of bilirubin; elevated TBL to >2xULN.
- Any elevations of ALP >1.5xULN.
- Elevation of AT (>3xULN) accompanied by elevated bilirubin (>2xULN).

In addition, a listing of potential Hy's Law cases will be provided. Potential hepatotoxicity is identified by the Hy's Law (FDA Guidance for Industry Drug Induced Liver Injury: Pre-marketing Clinical Evaluation, July 2009) as: ALT or AST >3.0 ULN; TBL>2.0 ULN, ALP<2.0 ULN; and no other confounding factors including preexisting or acute liver disease.

10.6. Vital Signs, Physical Findings, and Other Observations Related to Safety

10.6.1. Vital Signs

Vital signs, including systolic and diastolic blood pressure, pulse, and body temperature, as well as weight will be collected during screening, on day 1 of each treatment cycle, and at the treatment discontinuation visit.

Descriptive statistics will be calculated on both the actual value and the change from baseline value for the minimum, maximum, and last observation.

10.6.2. Physical Examinations

Physical examination data will be collected during screening, on day 1 of each treatment cycle and at the treatment discontinuation visit.

A full physical exam will be performed at Screening. Detailed/brief (system-guided) physical examinations will be done at subsequent time points to evaluate any clinically significant abnormalities, including worsening of conditions included in the patient's medical history.

Physical examination results will be listed with vital signs. Height will be collected at screening.
Electrocardiograms

Single standard 12 lead digital ECGs will be performed during screening, and every 8 weeks after that, and at treatment discontinuation, if not done within the previous 8 weeks.

The ECG (measured after resting in a supine position for 5 minutes) will include heart rate, PR, QRS, QT and QT_c intervals. The ECG will be read and interpreted at the investigational site for patient safety monitoring, and documentation stored with the source documents.

An additional correction formula included but considered secondary is QTcF. This additional corrected QT interval is defined as:

- QTcF is the length of the QT interval corrected for heart rate by Fridericia's formula:

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

where RR is the interval from the onset of one QRS complex to the onset of the next QRS complex, measured in seconds, often derived from the heart rate (HR) as 60/HR (here QT is measured in milliseconds).

All ECG parameters and their change from baseline will be summarized at the final, minimum, and maximum study time points in period 1(before crossover) for each treatment group using descriptive statistics.

All ECG values will be listed by patient for each cohort.

10.6.3. Left Ventricular Ejection Fraction (LVEF)

MUGA scan or ECHO to determine LVEF will be performed during screening and repeated every 8 weeks after that, and at treatment discontinuation, if not done within the previous 8 weeks.

MUGA scans or ECHO scans to determine LVEF will be performed as part of the screening procedures within 28 days of Cycle 1/Day 1.

The mean LVEF and mean LVEF change from baseline will be summarized at the last and minimum values in period 1 by treatment arm and overall using descriptive statistics.

All MUGA and ECHO values will be listed by patient for each treatment group.

11. OTHER ASSESSMENTS

11.1. ECOG Performance Status

Eastern Cooperative Oncology Group (ECOG) performance status will be assessed at screening, on day 1 of each treatment cycle, and at the treatment discontinuation visit. ECOG categories are also summarized in the study protocol.

Screening ECOG performance status may be accepted as the Baseline status if the assessment was performed within 72 hours of initiation of investigational product and there were no clinically significant findings.

ECOG status will be included in the baseline and demographic variables. The number and percentage of patients in each ECOG category will be presented for each treatment group, by nominal visit, and minimum and maximum values.

11.2. Patient Reported Outcomes: FACT-G

Patient Reported Outcomes will be evaluated using the FACT-G instrument. Data for each of the 4 subscales and the total score will be summarized and plotted over time for the following treatment groups:

a. Hormone receptor positive (HR+), HER2 negative breast cancer patients receiving N+F+T (randomized and open label)

b. Hormone receptor positive (HR+), HER2 negative breast cancer patients receiving F+T (until crossover to N+F+T)

c. Hormone receptor positive (HR+), HER2 negative breast cancer patients receiving F (until crossover to N+F+T)

d. Lung *EGFR* exon 18 mutation patients receiving N.

Summaries will be created through cycle 6 by nominal visit; the baseline score is the latter of the score at Screening of C1D1, and will be represented as C1D1. All patients with FACT-G scores in the respective treatment group will be summarized regardless of any baseline or post-baseline visits.

FACT-G was collected for patients in other cohorts, but the number of patients is small and a summary is not meaningful. Accordingly, these assessments for these patients, and those above, will be provided in the listings.

12. INTERIM ANALYSES AND DATA MONITORING

Interim analyses may be performed at the completion of each stage for cohorts with a Simon's 2-stage design and for other cohorts as necessary. An IDMC will be established to regularly review accumulating safety data and efficacy data throughout the study.

13. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

13.1. Overall Survival

Long-term follow up was discontinued on January 4, 2022. Accordingly, analysis of overall survival will not be conducted.

13.2. Database Lock

The database will be locked in December 2022, with not all patients off treatment. The analyses will be conducted with the data available in the database; analyses of the data post database lock will be described in an amended SAP.

14. REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics 1982, 38: 29-41.

Chapman PB, Hauschild A, Robert C, et al. Improved Survival with Vemurafenib in Melanoma with BRAF V600E Mutation. *New Engl J Med* 2011; 364(26): 2507-16.

Chow L, Gupta S, Hershman DL, et al. Safety and efficacy of neratinib (HKI-272) in combination with paclitaxel in ErbB2+ metastatic breast cancer. *Ann San An Breast Cancer Symp* 2009. Poster 5081.

Eisenhauer, E.A. et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer 2009; 45: 228-247.

FDA Guidance for Industry Drug Induced Liver Injury: Premarketing Clinical Evaluation, July 2009

FDA Guidance for Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, December, 2018

FDA Guidance for Submission of Abbreviated Reports and Synopses in Support of Marketing Applications, 1999.

Functional Assessment of Cancer Therapy – General. https://www.facit.org/measures/FACT-G.

Hyun O, Lodge, Wahl. Practical PERCIST: A Simplified Guide to PET Response Criteria in Solid Tumors 1.0. Radiology (2016); 280(2).

Koyama T, Chen H. Proper inference from Simon's two-stage designs. Statistics in Medicine 2008; 27(16): 3145-54.

Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* 2007;114(2):97-109. Erratum in: *Acta Neuropathol* 2007 Nov;114(5):547.

Simon R. Optimal two-stage designs for phase II clinical trials. *Control Clin Trials* 1989;10:1-10.

Zhou C, Wu Y-L, Chen G, et al. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *Lancet Oncology* 2011; 12(8): 735-42.