Nektar Therapeutics

STATISTICAL ANALYSIS PLAN

A PHASE 3 OPEN-LABEL, RANDOMIZED, MULTICENTER STUDY OF NKTR-102 VERSUS TREATMENT OF PHYSICIAN'S CHOICE (TPC) IN PATIENTS WITH METASTATIC BREAST CANCER WHO HAVE STABLE BRAIN METASTASES AND HAVE BEEN PREVIOUSLY TREATED WITH AN ANTHRACYCLINE, A TAXANE, AND CAPECITABINE

Protocol Number: 15-102-14, Protocol Amendment 4.0

Version: 1.0

Date: 16 April 2018

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ATTACHMENTS

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	anthracycline, taxane, capecitabine
ATCL	Anatomical, Therapeutic Chemical Level classification
BCBM	breast cancer brain metastases
BFI	Brief Fatigue Inventory
BMI	body mass index
BN-20	quality of life assessment specific to brain neoplasms
BSA	body surface area
CBR	clinical benefit rate
CMH	Cochran-Mantel-Haenszel
CNS	central nervous system
CR	complete response
DoR	duration of response
DMC	data monitoring committee
eCRF	electronic case report form
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D-5L™	EuroQol 5D
ESMO-MCBS	European Society for Medical Oncology magnitude of clinical benefit scale
GCP	Good Clinical Practice
GPA	Graded Prognostic Assessment
HER	human epidermal growth factor receptor
HR	hazard ratio
HRQoL	health-related quality of life
IPFI	initial progression-free interval
IRT	interactive response technology
ITT	intent-to-treat
IV	intravenous
KM	Kaplan-Meier

Abbreviation	Definition
KPS	Karnofsky Performance Status
LDH	lactate dehydrogenase
LR	locally recurrent
MBC	metastatic breast cancer
MedDRA	Medical Dictionary for Regulatory Activities
MID	minimal important difference
MMRM	Repeated Measures Linear Mixed Effects Model
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	not evaluable
ORR	objective response rate
OS	overall survival
PD	progressive disease
PE	physical examination
PEG	polyethylene glycol
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	partial response
Q21d	once every 21 days
QLQ-C30	Quality of Life Core 30
QoL	quality of life
RANO-BM	Response Assessment in Neuro-Oncology—Brain Metastases
RECIST	Response Evaluation Criteria in Solid Tumors
RS	raw score
SAE	serious adverse event
SD	stable disease
SN38	7-ethyl-10-hydroxy-camptothecin; the active metabolite of irinotecan
SOC	system organ class
SRS	stereotactic radiosurgery
TEAE	treatment emergent adverse event
TNBC	triple-negative breast cancer
TPC	Treatment of Physician’s Choice
UGT1A1	uridine diphosphate-glucoronosyl transferase 1A1
WHO-DDE	World Health Organization Drug Dictionary Enhanced

1.0 ADMINISTRATIVE STRUCTURE

This study will be managed via partnership between Nektar Therapeutics and a contract research organization. Central clinical laboratories will be used for processing of safety specimens, biomarkers, and pharmacokinetic (PK) samples. An interactive response technology (IRT) service provider will manage the randomization system, study drug, and comparator distribution and inventory management. A data monitoring committee (DMC) will be established to review the interim efficacy and safety data.

2.0 INTRODUCTION

This document describes of the planned statistical analyses of the data captured according to Nektar Therapeutics Protocol 15-102-14 “A Phase 3 Open-Label, Randomized, Multicenter Study of NKTR-102 versus Treatment of Physician’s Choice (TPC) in Patients with Metastatic Breast Cancer Who Have Stable Brain Metastases and Have Been Previously Treated with an Anthracycline, a Taxane, and Capecitabine” version dated 16 April 2018.

This Phase 3 study is conducted in accordance with the protocol, Good Clinical Practice (GCP), and all applicable regulatory requirements.

Analyses of PK, biomarkers, and pharmacoeconomic data will be addressed in separate analysis plans.

3.0 OBJECTIVES

3.1 Primary Efficacy Objective

To compare overall survival (OS) of patients who receive 145 mg/m^2 NKTR-102 given once every 21 days (q21d) with OS of patients who receive TPC selected from the following list of 7 single-agent intravenous (IV) therapies: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel. TPC will be administered per standard of care.

3.2 Secondary Objectives

- To compare the objective response rates (ORR) from NKTR-102 treatment with that of TPC; assessment of tumor outside the CNS will use the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1; assessment of central nervous system (CNS) metastases will use the Response Assessment in Neuro-Oncology Brain Metastases (RANO-BM)

- To compare progression-free survival (PFS) from NKTR-102 treatment with that of TPC; assessment of tumor outside the CNS will use RECIST version 1.1; assessment of CNS metastases will use RANO-BM

- To compare the clinical benefit rate (CBR) from NKTR-102 treatment with that of TPC (i.e., the proportion of patients having complete response [CR], partial response [PR], or stable disease [SD] for at least 4 months); CBR for peripheral lesions and for CNS lesions will be separately described

- To compare the duration of response (DoR) from NKTR-102 treatment with that of TPC

- To compare the time to CNS disease progression in those patients with CNS lesions present at study entry

- To compare the time to CNS recurrence in those patients without CNS lesions present at study entry

- To evaluate the safety profiles of NKTR-102 and TPC

- To compare health-related quality of life (HRQoL) from NKTR-102 treatment with that of TPC using the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Core 30 (QLQ-C30) with the specific brain neoplasms (BN-20) questionnaire, the EuroQol 5D (EQ-5D-5LTM) questionnaire, and the Brief Fatigue Inventory (BFI)

- To obtain PK data (in patients randomized to NKTR-102 only)

- To correlate presence of reduced function uridine diphosphate-glucuronosyl transferase 1A1 (UGT1A1) variants with NKTR-102 safety (in patients randomized to NKTR-102 only)
- To evaluate the pharmacoeconomic implications of NKTR-102 therapy using selected measures of health care utilization
- To evaluate the magnitude of clinical benefit using European Society for Medical Oncology magnitude of clinical benefit scale (ESMO-MCBS)

3.3 Exploratory Objective

- To identify biomarkers that correlate with response, PFS, and OS

4.0 STUDY ENDPOINTS

4.1 Primary Efficacy Endpoint

Overall survival defined as date of randomization to the date of death from any cause. Patients who are lost to follow-up will be censored at the time they were last known to be alive. Patients who are alive at the time of OS analysis will be censored at the date of data cut-off for OS analysis.

4.2 Secondary Endpoints

- PFS with respect to all lesions outside the CNS per RECIST version 1.1
- PFS in brain metastasis per the RANO-BM
- PFS overall (CNS and peripheral)
- ORR defined as proportion of patients with a confirmed CR or confirmed PR per RECIST version 1.1 for lesions outside the CNS; RANO-BM for CNS lesions based upon the best response as assessed by the Investigator.
- CBR with respect to all lesions outside the CNS per RECIST version 1.1 defined as the proportion of patients with CR, PR, or SD for at least 4 months from the date of randomization (≥120 days). CBR with respect to all CNS lesions per RANO-BM defined as the proportion of patients with SD for at least 4 months (≥120 days)
- DoR with respect to all lesions outside the CNS per RECIST version 1.1, defined as the time from the first documented CR or PR to the time of earliest evidence of disease progression or death from any cause.
- HRQoL using the EORTC QLQ-C30 with BN-20 questionnaire, BFI and EQ-5D-5L
- Evaluation of the pharmacoeconomic implications of NKTR-102 therapy using selected measures of health care utilization
- Magnitude of clinical benefit using ESMO-MCBS
- Incidence of grade 3 and higher toxicities including adverse events (AEs) and clinical significant laboratory tests, serious AEs (SAEs), AEs with outcome of death, AEs leading to dose reduction, dose delay, and dose withdrawal. Duration and intensity of study drug exposure.
- AEs of special interest including neutropenia, diarrhea, neuropathy, and hypersensitivity.

5.0 OVERALL STUDY DESIGN AND PLAN

5.1 Study Design

This open-label, randomized, two-arm, multicenter, international, Phase 3 study of NKTR-102 in patients with breast cancer brain metastases (BCBM) who have stable brain metastases will evaluate single-agent NKTR-102 (145 mg/m^2 q21d) in patients who have previously received an anthracycline, a taxane, and capecitabine (ATC) versus a comparator arm consisting of an active single-agent TPC.

In Group A, NKTR-102 will be administered at a dose level of 145 mg/m^2 on a q21d schedule as a 90-minute IV infusion on Day 1 of each treatment cycle. In Group B, TPC will be administered per standard of care. Patients randomized to TPC will receive single-agent IV chemotherapy, limited to choice of one of the following 7 agents: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel.

Up to 220 patients will be randomized using a 1:1 randomization ratio. The following stratification factors for randomization were selected to balance the treatment groups for known factors that influence prognosis in patients with metastatic breast cancer and brain metastases:

- geographic region (US versus the rest of world)
- hormone and human epidermal growth factor receptor (HER) status (triple-negative breast cancer [TNBC], HER2+/HR any, and HR+/HER2-)
- Eastern Cooperative Oncology Group (ECOG) Performance Status (ECOG 0 and ECOG 1)

Cross-over from Group B to Group A is not permitted. The duration of the study will be 26-47 months to final analysis. An independent DMC will assess interim safety and efficacy data.

All patients will undergo tumor assessments performed at the participating study center or at a radiology facility associated with the site. Tumor measurements will be evaluated locally per RECIST 1.1 and RANO-BM criteria. All tumor imaging (head, chest, abdomen and other as appropriate) and digital photography will be forwarded to a central imaging facility to permit blinded independent review (local assessment will be used for patient management).

5.2 Study Medications

5.2.1 Arm A: NKTR-102

5.2.1.1 Dosage and Administration

Body surface area (BSA) will be determined before the start of each cycle, based on institutional guidelines and will be capped at 2.4 m^2. In the instance where there are no institutional guidelines, baseline height and most recent weight will be used to calculate BSA. Each patient's NKTR-102 dose will be determined by multiplying the most recent BSA (capped at 2.4 m^2) by the starting dose of 145 mg/m^2. NKTR-102 for Injection will be administered as an IV infusion over 90 minutes (± 15 minutes). Premedications are not required to be administered prior to the initial infusion, but may be used for an individual patient, as needed.

5.2.2 Arm B: Treatment of Physician's Choice

Patients randomized to TPC will receive single-agent chemotherapy, limited to one of the following 7 agents: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel. TPC must consist of single-agent IV therapy, not combination therapy. Choice of Group B agent (TPC) for an individual patient will depend on the TPC drug products available at each medical center. For TPC products without generic versions (e.g., eribulin, ixabepilone, and nab-paclitaxel) the branded product must be commercially available at the medical center. Selection of a TPC drug product should be based on what would have been offered to the patient within that medical center if the patient was not participating in a clinical study.

6.0 STATISTICAL CONSIDERATIONS

6.1 General Considerations

Although this is an open label study, analyses for treatment comparisons prior to final database lock will not be performed except for the pre-planned interim safety and efficacy analyses as described in Section 9.0 of this document and in the DMC charter.

Blinding for aggregated treatment information during the trial and data unblinding for interim analysis and after final database lock will be provided in separate documents.

Summary statistics for continuous variables will include the mean, standard deviation, median, minimum, maximum, and 25% and 75% quartiles. The mean will be presented to one decimal place beyond which the data were captured. The standard deviation will be presented to two decimal places beyond which the data were captured. The minimum and maximum will be presented to the precision with which the data were captured.

Categorical variables will be presented as frequency counts and percentages. A row or column denoted 'Missing' will be included in count tabulations where necessary to account for dropouts and missing values. Percentages will be rounded to 1 decimal place and the percent will be suppressed when the count is zero. The denominator will be the number of patients in that dose schedule within the population of interest unless otherwise noted.

Time-to-event variables will be analyzed using the Kaplan-Meier (KM) method. The KM estimates for quartiles and the 95% confidence interval for the median will be presented. All time to event variables will be plotted using the KM method.

Data listings will be created to support each table and to present all data. Data listings will be presented by treatment group and patient number.

6.2 Determination of Sample Size

The study is powered for detecting superiority of NKTR-102 compared with TPC in OS and up to 220 patients will be enrolled. The number of OS events needed to provide 80% conditional power for the final analysis will be determined at an interim analysis when approximately 82 OS events are available using the promising zone adaptive method (Mehta & Pocock, 2011). The minimum and maximum number of events for the final analysis will be 106 and 191, respectively. Two-sided alpha of 0.001 will be used to test efficacy at the interim analysis (efficacy zone-as part of promising zone design). One hundred ninety-one (191) events will be able to demonstrate statistical significance for any observed hazard ratio of 0.75 or better, or a corresponding median difference of approximately 2 months or better if the OS median for the control arm is 6 months and the proportional hazard assumption is approximately met. The

detailed event size adaptation rules based on conditional power will be provided in the DMC Charter Appendix – Event Size Adaption Rule for Clinical Study Protocol 15-102-14.

6.3 Analysis Populations

Intent-to-Treat (ITT) Population: The ITT population includes all patients who are randomized in the study. The primary endpoint OS and secondary efficacy analyses (except ORR and DoR) will utilize the ITT population.

Response Evaluable Population: The Response Evaluable population includes all patients who are randomized in the study with measurable disease in the periphery by RECIST 1.1 at baseline as determined by the Investigator. The secondary endpoint analyses of ORR and DoR will utilize the Response Evaluable Population.

Safety Population: The Safety Population consists of all patients who are randomized and receive at least one dose (or partial dose) of study drug (NKTR-102 or TPC). Safety analyses will be conducted using this population and the treatment arm (A or B) received, based on the first dose of study drug.

6.4 Handling of Missing Data

- Missing data will be handled as follows:
- For time from initial diagnosis to informed consent and time from diagnosis of metastatic disease to informed consent, partial dates for initial diagnosis and diagnosis of metastatic disease, will be imputed as follows:
  - No imputation will be done if the year is missing.
  - If the year is before informed consent date then missing days will be imputed as the first day of the month and missing months will be imputed as July.
  - If the year is the current year of informed consent date then missing days will be imputed as the first day of the month and missing months will be imputed as January.
- For duration of AEs, partial dates for start of AE, will be imputed as follows:
  - Missing day, month, and year should be queried. In case of non-resolution of missing month and/or year, no imputation will be performed.
  - Start day of AE is missing
    - If the reported month of occurrence of AE is after the month of Cycle 1 Day 1 dose then day will be imputed as the first day of the month of occurrence of AE.

    - If the reported month of occurrence of AE is the month of Cycle 1 Day 1 dose then the missing day will be imputed as the same day as Cycle 1 Day 1.
- For duration of AEs, partially or completely missing dates for stop of AE, will be imputed as follows:
  - Missing day, month, and year are not allowed and should be queried. In case of non-resolution of missing month and year, no imputation will be performed.
  - If the day and/or month is missing, but year is present, the last day of that month and December, the date of discontinuation from the study, or the date of death, whichever is earlier, will be used as the stop date.
- For determination of prior or concomitant medication, missing or partial dates will be handled as follows:
  - Missing day of month will be imputed as the first day of the month. Missing month will be imputed as July if the year is before Cycle 1 Day 1; missing month as January if the year is the year of Cycle 1 Day 1.
    - If the stop date of a medication is before the month of Cycle 1 Day 1 dose then the medication will be classified as prior medication.
    - If the month of the stop date is the month of Cycle 1 Day 1 and day of the month is missing, then the medication will be classified as both prior medication and concomitant medication if the start date is prior to Cycle1 Day 1; or concomitant medication only if the start date is post Cycle1 Day 1.
    - If the date is completely missing, the medication will be classified as both prior and concomitant medication unless stop date can be used to classify specifically.

Handling of missing data for analysis of PFS and QoL are described in the statistical analysis sections below.

No imputation of other missing data is planned.

6.5 Stratification and Pooling

For stratified analyses, the stratification factors (Region, Receptor Status, and ECOG performance status) entered into the interactive voice response system at the time of randomization will be used. If the proportion of patients assigned to incorrect stratums exceeds 5% of the total patients randomized, a sensitivity analysis using stratums based upon data in the clinical database will be performed. In the BEACON trial, the incidence of patients with HER2+ breast cancer and brain metastases was 13.4% (9 out of 67 patients); the incidence of ECOG=0 patients was 23.9% (16 out of 67 patients). To avoid unstable estimates, in this trial, if any cells

(defined by the three stratification factors) have fewer than 5 patients, patients may be pooled (for example, patients with HER2+ breast cancer may be pooled with patients who have HR+/HER- or TNBC, whichever group has fewer patients, i.e., Receptor Status will be collapsed into two levels from three levels).

6.6 Analytic Definitions

General:

- Baseline: Baseline will be defined as the last assessment result on or before the first study drug administration.
- Study day: There is no study day zero. Day 1 is the date of first study drug administration.
  For post-treatment events, study day is calculated as:
  Date of visit or assessment – date of Day 1 + 1
  For pre-treatment events, study day is calculated as:
  Date of visit or assessment – date Day 1

Baseline Characteristics:

- Body Mass Index (BMI) (kg/m^2) is defined as: (Weight in kg) / $(\text{Height in m})^2$
- BSA will be calculated based on the following formula (Mosteller, 1987) and used for summary purposes:

$$BSA\ (m^2) = ([\text{Height (cm)} \times \text{Weight (kg)}]/\ 3600)^{1/2}$$

Exposure:

- Exposure duration (days): date of last dose – date of first dose (Cycle 1 Day 1) + 1.
- Number of cycles: Total number of complete or partial treatment cycles the patient received.
- Cumulative dose (mg): Total actual dose (mg) the patient received across all cycles, defined as the sum of actual dose (mg) received across all cycles.
- Average dose per infusion: Mean of (the actual dose receive/BSA at that cycle) across all cycles.
- Duration of infusion: Completion time of infusion – start time of infusion.
- Calculated cumulative dose level (mg/m^2): Cumulative dose (mg) divided by the average BSA (m^2) across all cycles.

For Arm A (NKTR-102), the following parameters will be calculated:

- Dose intensity (mg/m^2/week): [Calculated cumulative dose level (mg/m^2) / (exposure duration (days) + 20 days)] × 7
- Expected dose intensity (mg/ m^2/week) = (145 mg/m^2)/(3 week) = 48.3 mg/m^2/week
- Relative dose intensity (%): ((Dose intensity) / (expected dose intensity)) × 100

For Arm B, the following parameters will be calculated:

- Dose intensity (mg/m^2/week): {Calculated cumulative dose level (mg/m^2) /[(exposure duration (days) + ADD ON PERIOD – 1 day)]} x 7 ; where ADD ON PERIOD = cycle length as planned for Cycle 1 – (date of last dose of the last cycle – date of 1st dose in the last cycle)
- Expected dose intensity (mg/m^2/week): {[initial planned Cycle 1 Day 1 dose (mg/m^2) × number of doses given in Cycle 1] / cycle length as planned for Cycle 1} × 7
- Relative dose intensity (%): ((Dose intensity) / (expected dose intensity)) × 100.

Visit Windows:

For statistical summary by visit or cycle, data will be assigned to a derived visit window. If more than one eligible result is reported within the same window, the result collected closest to the target day will be used for the descriptive statistics and the worst value (most conservative) result will be used in the shift table or summary for abnormalities.

The following visit windows will be used to assign analysis visits for by-week summaries including tumor assessments and HRQoL:

Table 1: Visit Windows

Planned Visit	Target Day	Starting Day for Visit Window	Ending Day for Visit Window
Screening	Screening	Day -28	randomization
Week 8	56 days since randomization	1 day since randomization (and post Cycle 1 Day 1)	83 days since randomization
Week 16	112 days since randomization	84 days since randomization	139 days since randomization
Week 8*X	8*X time 7 days since randomization	(8*X-4)*7 days since randomization	(8*X+4)*7 – 1 days since randomization

7.0 STATISTICAL ANALYSIS

7.1 Patient Disposition

A summary of patient disposition will display the number of patients who were randomized and who comprised each analysis population by treatment arm. In addition, the number of patients who discontinued study drug and the number of patients who exited the study, both overall and by reason, will be presented by treatment arm.

The three randomization stratification factors, geographic region , and receptor status based upon local review of disease (HER2+, HR+/HER2-, and TNBC) will be tabulated as documented on the electronic case report form (eCRF) and as recorded in the IRT system by treatment arm as randomized. Discrepancies between those two sources will be identified and summarized.

All disposition data will be presented in a data listing.

7.2 Protocol Deviations

Important protocol deviations will be defined in a separate document. These protocol deviations will be captured by searching relevant data fields reported in the clinical database using computer algorithm and site monitoring by clinical research associates tracked in the clinical tracking system. Final determination of important deviations will be reviewed by the study team at regular intervals throughout the trial (including clinical study manager, the Medical Monitor, and the statistician).

The number and percentage of patients in each important and non-important protocol deviations category and sub-category will be summarized by treatment arm for the ITT population. All protocol deviations will be listed.

7.3 Demographics and Baseline Characteristics

The following baseline data will be summarized and listed for all analysis populations: age (years), sex, race, ethnicity, ECOG performance status, reproductive status, height (cm), weight (kg), BSA (m^2), BMI (kg/m^2), and key laboratory tests including absolute neutrophil count (ANC), hemoglobin, platelets, bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), albumin, creatinine, and creatinine clearance.

7.3.1 Medical and Cancer History

Cancer history will include time since initial breast cancer diagnosis, stage, histology and receptor status at initial breast cancer diagnosis, time since diagnosis of locally advanced or

metastasis breast cancer, initial disease-free interval, time since diagnosis of metastatic disease to the brain, time since most recent prior treatment for BCBM such as whole brain radiation, stereotactic radiation and/or surgical resection, type of prior therapy for BCBM, radiographic evidence of metastatic brain lesions at baseline (Yes/No), number of sites of disease, receptor status at last biopsy, and Graded Prognostic Assessment (GPA) score (Table 6) at study entry.

Medical history collected at screening will be mapped by the Dictionary for Regulatory Activities (MedDRA) and will be summarized by system organ class (SOC) and preferred term for the ITT population. For summary tables, a patient will be counted only once per SOC and preferred term.

7.3.2 Prior Systemic Cancer/Oncology Therapies

Prior systemic cancer/oncology therapies will be tabulated using the World Health Organization Drug Dictionary Enhanced (WHO-DDE) and Anatomical, Therapeutic, or Chemical Level (ATCL)-2 classifications and preferred term. If the ATCL-2 classification is missing, the next non-missing higher level of classification will be used (Level 1). If a patient reports the same medication multiple times, then the frequency of that medication will be incremented by only one in the applicable arm. As with the medication, if a patient reports multiple medications within the same ATCL-2 classification, then the frequency of that ATCL-2 classification will be incremented by only one in the applicable arm. Percentages will be calculated using the number of patients in the ITT population.

The number and percent of patients who received prior anthracycline, taxane, and capecitabine treatment will be summarized. The number and percent of patients who received prior hormonal therapies, HER2-directed therapies, cytotoxic chemotherapy regimens for breast cancer, neoadjuvant regimens, adjuvant regimens, regimens for locally recurrent, metastatic disease, secondary primary, and chemoradiation will be calculate using the ITT population. Duration of prior anthracycline, taxane, capecitabine and eribulin use will be summarized descriptively. If a patient received multiple treatments of a therapy, the sum of duration will be calculated.

All prior systemic cancer/oncology therapies will be listed.

7.3.3 Surgical History and Prior Radiotherapy

The number and percent of patients who had had prior cancer-related CNS and non-CNS surgery will be summarized by type of surgery for each treatment arm using the ITT population.

The number and percent of patients who had prior radiotherapy will be summarized by site and type of the radiotherapy for each treatment arm using the ITT population.

All prior radiotherapy and surgical procedures will be listed.

7.3.4 Concomitant Radiotherapy, Procedures and Surgery for CNS Lesion.

The number and percent of patients who had concomitant radiotherapy will be summarized by site and type of radiotherapy for each treatment arm using the ITT population.

The number and percent of patients who had concomitant procedures will be summarized by each coded procedure preferred term for each treatment arm using the ITT population.

The number and percent of patients who had concomitant surgery for CNS lesion will be summarized by type of surgical procedure for each treatment arm using the ITT population.

All concomitant radiotherapy, procedures, and surgery for CNS lesion will be listed.

7.4 Treatments and Medications

7.4.1 Prior and Concomitant Medications

Prior and concomitant medications will be coded to ATCL and preferred drug name using WHO-DDE.

Concomitant medications are defined as medications taken on or after the date of first dose, including medications initiated prior to the date of first dose and continued during treatment, and medications initiated on or after the date of first dose.

Concomitant medications will be tabulated for the Safety Population by WHO-DDE ATCL-2 classifications and preferred term. If the ATCL -2 classifications is missing, the next non-missing higher level of classification will be used (Level 1). If a patient reports the same medication multiple times, then the frequency reported for that medication will be incremented by only one. As with the medication, if a patient reports multiple medications within the same ATCL-2 classification then the frequency for that ATCL-2 classification will be incremented by only one. Percentages will be calculated using the total number of patients in the Safety population.

Prior medications are defined as medications taken starting prior to the first dose. Prior medications will be summarized for the ITT population using the same analytical procedures as concomitant medications.

Prior and concomitant medications will be presented in a data listing.

7.4.2 Subsequent Anti-Cancer Therapy, Radiotherapy, and Surgery.

Subsequent anti-cancer systemic therapy will be summarized based on the ITT population. Subsequent Anti-cancer therapy will be tabulated using the WHO-DDE and ATCL -2 classifications and preferred term. The number and percent of patients took at least one subsequent anti-cancer therapy will be calculated and presented by treatment group, ATCL-2 classifications, and preferred term.

The number and percent of patients who had post-treatment radiotherapy will be summarized by site and type of radiotherapy for each treatment arm using the ITT population.

The number and percent of patients who had post-treatment surgery for CNS lesion will be summarized by type of surgical procedure for each treatment arm using the ITT population.

7.4.3 Exposure / Study Treatments

Overall exposure to study treatments (Arm A and B) and study treatment administration for each cycle and all cycles combined will be summarized for the Safety population. Overall exposure to study treatment will be summarized in terms of exposure duration, number of cycles, cumulative dose (mg), and average infusion duration (minutes) and will be calculated for both treatment arms. Average dose per infusion (mg/m^2), dose intensity (mg/m^2/week), and relative dose intensity will be calculated. The number of patients with dose reduction, dose interruption, and dose delay will be tabulated along with their reason.

Association between number of treatment cycles received and baseline characteristics including geographic region, receptor status, ECOG, GPA score, baseline ANC, hemoglobin and platelets, baseline AST and creatinine) will be examined.

For patients who are randomized to NKTR-102, the quartiles of dose intensity and duration of exposure will be calculated. Patients will be categorized into ≤25%, <25% to ≤ 50%, < 50% to ≤75%, and >75% group. For each group, the number and percent of patients who have selected toxicity (neutropenia, diarrhea, neuropathy, and hypersensitivity) will be summarized. In addition, the number and percent of patients who have grade 3 or more selected toxicity will be summarized.

All study exposure data will be presented in a data listing.

7.5 Efficacy Analysis

7.5.1 Analysis of Primary Endpoint

The primary efficacy endpoint, which will be calculated on the ITT population, is the overall survival defined as the time from the date of randomization to death from any cause. Patients will be followed until their date of death, loss to follow-up, and withdrawal of consent for further follow-up for survival. Patients who do not have date of death will be censored at last date shown to be alive. Patients who do not have date of death at final database closure will be censored at the date of event cut-off for OS analysis. Patients who do not have any follow up since randomization will be censored at the date of randomization.

At final analysis, the primary analysis of OS for statistical significance claim will be based on the CHW version of the logrank test statistic with weights $\sqrt{(2/3)}$ and $\sqrt{(1/3)}$ (Cui, Hung, & Wang, 1999) at two-sided 0.0499 level. The conventional logrank test with equal weights for every patient will be conducted as a sensitivity analysis.

The median survival times and their 95% confidence intervals as well as survival curves will be estimated using the Kaplan-Meier method and will be summarized by treatment group.

The proportion of patients that are alive at 6 and 12 months and the corresponding 95% confidence intervals will be calculated as well. A single hazard ratio comparing NKTR-102 to TPC and its 95% confidence interval will be calculated using a Cox regression model adjusting for geographic region, receptor status and ECOG performance status.

Above analysis of OS (i.e. log rank test and Cox regression model) will be repeated excluding patients from both arms who selected ixabepilone as their TPC agent prior to randomization.

If more than 10% of study population (more than 22 patients) received concomitant stereotactic radiosurgery (SRS) during the study, the proportion of patient received SRS during the study between treatment groups will be compared using Fisher's exact test. Use of post-study SRS and its impact on OS will also be examined.

The impact of SRS on OS will be evaluated by a Cox regression model comparing patients received SRS to those who did not receive it. The treatment effect will be adjusted for SRS by a Cox regression model with treatment arm and usage of SRS as covariates.

Subgroup analyses of overall survival will be performed to assess whether the treatment effect is concordant among subgroups. The following variables will be used to define subgroups:

- Number of prior regimens (≤3, and ≥4)

- Age (< 65 versus ≥ 65)
- Race (white, black, asian, other)
- Time from initial breast cancer diagnosis to randomization (≤ 4 years versus > 4 years)
- Time since LR/MBC diagnosis (≤ 2 years versus > 2 years)
- Time since BCBM diagnosis (≤ 6 months versus > 6 months)
- Time since local therapy for BCBM (≤ 3 months versus > 3 months)
- Initial progression-free interval (IPFI, < 2 year versus >= 2 years. IPFI will be calculated from the date of first dose to the date of progressive disease (PD) for the first systemic cancer treatment for breast cancer in locally recurrent or metastatic setting
- Patients with evidence of brain metastatic lesions on head imaging at entry (Yes versus No)
- Patients with 1-3 brain lesions versus >3 brain lesions
- Prior whole brain radiation vs. stereotactic radiation (with or without surgical resection)
- GPA score 0-2 versus 2.5-4 at study entry
- Steroid use for control of neurological signs/symptoms at baseline (Yes versus No)
- Initial disease free interval (≤ 2 years versus > 2 years)
- Geographic region (US versus other)
- Geographic region (by Country)
- ECOG performance status (0 versus 1)
- Receptor status based on local review of disease (TNBC versus HER2+ versus Other)
- HER2 status (positive versus negative versus unknown)
- Hormone receptor status (positive versus negative versus unknown)
- Number of sites involved (≤2 versus ≥3); disease site will be mapped based on the Target or Non-target tumor location reported in the database

- Prior eribulin use (Yes versus No)
- Prior hormonal therapy (Yes versus No)
- Metastasis in liver (Yes versus No, Yes is defined as at least one Target or Non-target tumor located in liver at baseline)
- Metastasis in lung (Yes versus No, Yes is defined as at least one Target or Non-target tumor located in lung at baseline)

The benefit of NKTR-102 compared to TPC will be evaluated by a single hazard ratio (NKTR-102/TPC) with its 95% confidence interval based on a Cox regression model for each subgroup. The hazard ratios and their 95% confidence intervals will be displayed in a forest plot. Subgroup analyses above may not be performed if the sample size is too small to provide accurate estimate.

7.5.2 Analysis of Secondary Efficacy Endpoints

Analysis for secondary endpoints will not include any adjustment for multiplicity. Statistical tests will be two-sided with a type I error rate of 0.05.

7.5.2.1 Progression-Free Survival (Outside the CNS)

Progression-free survival analyses will be performed for the ITT population.

The revised RECIST guidelines (Eisenhauer, 2009) will be used to determine disease progression. Date of PD is the earliest date of the imaging method to determine PD, or if the patient is followed solely by physical examination (PE) the date of the PE showing PD. The date of global deterioration" or "symptomatic deterioration" will not be used as the date of PD.

Progression-free survival is defined as the time from the date of randomization to the earliest evidence of documented PD or of death from any cause. For the primary analysis of PFS, the censoring methods described in Table 2 will be used. These rules are considered as the censoring criteria described in (Stone. 2011) and are consistent with censoring rules described in the FDA Guidance for Industry document – Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Appendix 3 Table A (PFS includes documented progression only) (FDA 2007).

Table 2: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-off	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to start date of the anti-cancer therapy	Censored
Prematurely discontinued study drug due to non-progression related reasons	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to study drug discontinuation	Censored

Abbreviations: PD = progressive disease; PE = physical examination

The first sensitivity analyses of PFS will be performed based on the censoring methods provided in Table 3. The difference between the rules described in Table 2 and Table 3 is that patients who prematurely discontinue the study drug due to reasons other than progression will be censored in Table 2 while in Table 3, they will not be censored. These rules are consistent with censoring criteria described in (Stone. 2011). These rules are also consistent with the considerations of informative censoring as outlined in the Guideline on the Evaluation of Anticancer Medicinal Products in Man – Appendix 1 (EMA 2012).

Table 3: Alternative censoring approach: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD.	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to start date of the anti-cancer therapy	Censored

Abbreviations: PD = progressive disease; PE = physical examination

The second sensitivity analysis of PFS will be performed based on the censoring methods described in Table 4. The differences between the rules described in Table 2 and Table 4 are that patients who prematurely discontinue the study drug due to reasons other than progression or patients who start any new anti-cancer therapy will be censored in Table 2 while in Table 4 they will not be censored. These rules are consistent with the censoring criteria described by Stone, 2011. These rules are also consistent with the considerations of informative censoring as outlined in the Guideline on the Evaluation of Anticancer Medicinal Products in Man – Appendix 1 (EMA 2012).

Table 4: The second alternative censoring approach: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-off	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD	Date of earliest evidence of PD	Progressed

Abbreviations: PD = progressive disease; PE = physical examination

Progression-free survival will be compared with a two sided, stratified, log-rank test with the same stratification factors that were used for randomization. Median PFS time, its 95% confidence intervals as well as 25% and 75% quartiles of PFS times will be summarized using KM method for each treatment group.

A single hazard ratio (NKTR-102/TPC) and its 95% confidence interval will be calculated using a stratified Cox regression model adjusting for geographic region, hormone receptor status, and baseline ECOG performance status.

7.5.2.2 Progression-Free Survival in Brain Metastasis

Progression-free survival in brain metastasis for patients with CNS lesions at baseline will be analyzed separately to patients without CNS lesions at baseline. For patients with CNS lesions at baseline, progression-free survival in brain metastasis is defined as the time from the date of randomization to the earliest evidence of documented PD per RANO-BM in brain metastases or death from any cause. For patients without CNS lesion at baseline, progression-free survival in brain metastasis is defined as the time from the date of randomization to the earliest evidence of disease recurrence in the CNS or death from any cause.

The censoring rule specified in Table 2 will be used. The same statistical methods used for the analysis of PFS outside the CNS will be used for the analysis PFS-brain metastasis.

7.5.2.3 Progression-Free Survival (Overall)

Progression-free survival (CNS and peripheral) is defined as the time from the date of randomization to the earliest evidence of documented PD in either CNS per RANO-BM or peripheral per RECIST 1.1 or death from any cause. For patients who are censored per RECIST 1.1 (Table 2) but have date of PD per RANO-BM, the date of disease progression per RANO-BM will be used. For patients who are censored per RANO-BM but have date of PD per RECIST 1.1, the date of disease progression per RECIST 1.1 will be used. For patients who are censored by both RECIST 1.1 and RANO-BM, patients will be censored on the date of the earliest censoring per RECIST 1.1 and RANO-BM.

The same statistical methods used for the analysis of PFS (outside CNS) will be used for the analysis of PFS (Overall).

7.5.2.4 Objective Response Rate

Objective response rate analyses will be performed for the Response Evaluable population and defined as the proportion of patients with a confirmed CR or PR for lesions outside the CNS per RECIST 1.1. The analysis of ORR will be performed using Fisher's exact test between the two treatment arms. Clopper-Pearson exact 2-sided 95% confidence limits will be calculated for the proportion of patients with ORR in each arm.

The primary analysis of ORR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

In addition, the best response using response categories CR, PR, SD, PD, and not evaluable (NE) per RECIST 1.1 outside CNS and per RANO-BM in CNS will be tabulated. The number and percent of patients in each response category will be calculated. Patients who do not have any post-baseline tumor assessment will be counted under the category NE.

7.5.2.5 Clinical Benefit Rate.

Clinical benefit rate outside CNS is defined as the proportion of patients having a CR, PR, or SD for at least 4 months (≥ 120 days) per RECIST 1.1. The SD duration of 4 months is selected to reflect the shorter life expectancy of the study population given that the median OS for the TPC arm in patients with history of brain metastases was 4.8 months from the BEACON trial. CBR analyses will be performed for the ITT Population. CBR will be calculated and compared between the treatment arms using a Cochran Mantel-Haenszel (CMH) test stratified by the randomization factors. Clopper Pearson exact 2-sided 95% confidence limits will be calculated for CBR of each treatment arm.

In addition, the CBR in CNS will be defined as the proportion of patients having a SD for at least 4 months (≥ 120 days) per RANO-BM. Clinical benefit rate in the brain will be compared between the treatment groups using a CMH test stratified by the randomization factors. Clopper Pearson exact two-sided 95% confidence limits will be calculated in determining the CBR of each group.

The primary analysis of CBR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

7.5.2.6 Duration of Response

Duration of response outside CNS is defined as the time from first documented CR or PR until the earliest evidence of disease progression per RECIST 1.1 or death from any cause. The algorithm for calculation of date of progression and censoring rule (Table 2) will be the same as defined in the analysis for PFS. Median duration, its 95% confidence intervals as well as 25% and 75% quartiles of DoR will be summarized using Kaplan-Meier method for each treatment arm.

The primary analysis of DoR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

7.5.3 HRQoL

The EORTC QLQ-C30 with BN-20, BFI and EQ-5D-5L will be used to measure the health outcome, QoL and assess the symptoms and side effects of treatment and their impact on everyday life.

The QLQ-C30 questionnaire is composed of 5 multi-item scales (physical, role, social, emotional, and cognitive functioning), a global health status/QoL scale, and 9 symptoms (pain, fatigue, financial impact, appetite loss, nausea/vomiting, diarrhea, constipation, sleep disturbance, and dyspnea). Most items are scaled 1 to 4, except items contributing to the global health status/QoL, which are 7-point questions. Raw scores will be transferred using a linear transformation to standardize the results such that scores range from 0 to 100.

The calculation for scoring these scales is the same in all cases:

1. Calculate the average of the items that contribute to the scale; this is the raw score.

2. Use a linear transformation to standardize the raw score, so that scores range from 0 to 100; a higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms.

Calculations for raw score and linear transformation are as follows:

In practical terms, if items $I_1, I_2, \ldots I_n$ are included in a scale, where n is number of items in the scale, the procedure is as follows:

Raw score

$RS\ (Raw\ Score) = (I_1 + I_2 + I_3 + \ldots + I_n)/n$

Linear transformation

Apply the linear transformation to 0-100 to obtain the score S,

Functional scales: $S = \{1 - (RS - 1)/range\} \times 100$

Symptom scales / items: $S = \{(RS - 1)/range\} \times 100$

Global health status / QoL: $S = \{(RS - 1)/range\} \times 100$

Range is the difference between the maximum possible value of *RS* and the minimum possible value.

Raw score and linear transformed score will be calculated when at least 50% of the items from the scale have been answered. Otherwise, the score will be set as missing.

The structure of this questionnaire and scoring are presented in Table 5.

Table 5: Scoring for QLQ-C30 version 3.0

Scale	Scale abbreviation	Number of items	Individual score range	Item Number
Global health status / QoL				
Global health	QL	2	1-7	29, 30
Functional scales				
Physical functioning	PF	5	1-4	1 to 5
Role functioning	RF	2	1-4	6, 7
Emotional functioning	EF	4	1-4	21 to 24
Cognitive functioning	CF	2	1-4	20, 25
Social functioning	SF	2	1-4	26, 27
Symptom scales / items				
Fatigue	FA	3	1-4	10, 12, 18
Nausea and vomiting	NV	2	1-4	14, 15
Pain	PA	2	1-4	9, 19
Dyspnea	DY	1	1-4	8
Insomnia	SL	1	1-4	11
Appetite loss	AP	1	1-4	13
Constipation	CO	1	1-4	16
Diarrhea	DI	1	1-4	17
Financial difficulties	FI	1	1-4	28

At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group for each subscale. Changes from baseline will be compared between treatment groups using Repeated Measures Linear Mixed Effects Model (MMRM) with treatment, time, treatment and time interaction, baseline score, and stratification variables as covariates. Unstructured covariance matrix will be used unless the model fails to converge. Treatment effect on the change from baseline will be compared by the difference in the least-square means and corresponding 95% confidence interval.

Additional analysis, including generalized linear mixed effect model for sub-scales with small number of categories, maybe explored,

Patients will be classified as improved, stable, or worsened. Minimal important difference (MID) thresholds, a threshold of 5 score points, will be used to categorize patients as improved (≥ 5 decrease), stable (0 to <5 decrease), or worsened (increase) on the EORTC QLQ-C30 scores. The proportion of patients within each status group using the best change from baseline

will be calculated for each subscale. The proportion of patients improved, stable, or worsened will be calculated and compared between treatment groups using the CMH test stratified by the factor used for randomization.

Selected symptoms (e.g., neuropathy or diarrhea) will be identified by medical review. Time to progression of selected symptoms will be calculated and compared between treatment groups. Progression is defined as an increase in at least 1 point from baseline. Time to progression of a selected symptom will be calculated between the date of randomization and date of first observation of the progression. Patients who do not experience progression of a symptom will be censored at the date of their last EORTC QLQ assessment or at their date of randomization (if no post-baseline EORTC QLQ assessment is completed). A KM method will be used to estimate the median times and their 95% confidence interval for each treatment group. A two-sided log-rank test will be performed to compare treatment groups for time to progression in each selected symptom. If the proportion of patients with progression of selected symptoms is below 30%, descriptive statistics will be used to provide summary time to progression.

The brain cancer module (QLQ-BN20) is intended for patients undergoing chemotherapy or radiotherapy. It includes 20 items (individual score range 1-4) that aggregates into 4 multi-item scales of future uncertainty (4 items), visual disorder (3 items), motor dysfunction (3 items), communication deficits (3 items); and 7 single-item scales of headaches, seizures, drowsiness, itchy skin, hair loss, weakness of legs, and bladder control. All scale scores and items are linearly transformed to a 0–100 scale with higher scores indicating more severe symptoms.

At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group for each subscale. In addition, change from baseline score in QLQ-BN20 will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

The BFI was developed for the assessment of the severity of fatigue and the impact of fatigue on daily functioning. It consists of 9 numeric rating scales of 0 to 10: 3 items that measure severity and 6 items that measure interference with daily activities. A scale score will be calculated if a minimum of 5 items have been answered. A global fatigue score will be obtained by averaging the answered items on the BFI. At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group. In addition, change from baseline score in BFI will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

The EQ-5D-5L consists of descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D-5L descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems,

slight problems, moderate problems, severe problems, and extreme/unable to perform activity. The responses from each of 5 dimensions will be combined in a 5-digit number describing the respondent's health state. It should be noted that the numerals 1-5 have no arithmetic properties and should not be used as a cardinal score. At each assessment point, the EQ-5D-5L levels will be divided into 'no/slight problem', 'moderate problem' or 'severe/extreme problems'. The proportion of patients within each status will be calculated and compared between treatment groups using the CMH test stratified by the factors used in randomization.

The EQ VAS records the patient's self-rated health on a vertical, visual analogue scale where the endpoints are labeled 'Best imaginable health state' and 'Worst imaginable health state'. This information can be used as a quantitative measure of health outcome as judged by the individual patients. At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group. In addition, change from baseline score will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

7.5.4 Graded Prognostic Assessment Index

The GPA tool, which assigns scores for prognostic indices using three prognostic factors for survival, Karnofsky performance status (KPS), tumor subtype, and age has been developed to predict estimated survival in patients with newly diagnosed brain metastases.

The GPA will be calculated as the sum of scores from the three factors according to Table 6.

Table 6: GPA Calculation

Factor\ Score	0	0.5	1.0	1.5	2.0
KPS	50 or less	60	70-80	90-100	n/a
Tumor Subtype	Basal	n/a	Luminal A	Her2+	Luminal B
Age	60 or greater	Less than 60	n/a	n/a	n/a

Abbreviations: KPS = Karnofsky Performance Status; n/a = not applicable

The following tables outline the mapping from ECOG performance status score to KPS score and from receptor status to tumor subtype

Table 7: ECOG Performance Status versus Karnofsky Performance Status

ECOG PERFORMANCE STATUS	KARNOFSKY PERFORMANCE STATUS
0—Fully active, able to carry on all pre-disease performance without restriction	100—Normal, no complaints; no evidence of disease 90—Able to carry on normal activity; minor signs or symptoms of disease
1—Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work	80—Normal activity with effort, some signs or symptoms of disease 70—Cares for self but unable to carry on normal activity or to do active work
2—Ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours	60—Requires occasional assistance but is able to care for most of personal needs 50—Requires considerable assistance and frequent medical care
3—Capable of only limited self-care; confined to bed or chair more than 50% of waking hours	40—Disabled; requires special care and assistance 30—Severely disabled; hospitalization is indicated although death not imminent
4—Completely disabled; cannot carry on any self-care; totally confined to bed or chair	20—Very ill; hospitalization and active supportive care necessary 10—Moribund
5—Dead	0—Dead
Tumor Subtype	
Her2+	ER-, PR- and Her2 +
Basal	ER-, PR- and Her2- (TNBC)
Luminal A	ER+ and/or PR+ and HER2-
Luminal B	ER+ and/or PR+ and HER2+

The GPA categories range from 0-4 with a lower score predictive of a worse prognosis for survival following a diagnosis of brain metastasis. The association between GPA Index and survival will be evaluated using a Cox regression model. GPA-adjusted hazard ratio will be presented if the association is significant.

7.5.5 Magnitude of Clinical Benefit

The magnitude of clinical benefit of NKTR-102 will be assessed by the ESMO magnitude of clinical benefit scale (ESMO-MCBS v1.0). The ESMO-MCBS can be applied to comparative outcome studies evaluating the relative benefit of treatments using outcomes of survival, QoL, surrogate outcomes for survival or treatment toxicity in solid cancers. For cancer therapies, the ESMO-MCBS scale provides a clear, well-structured and validated mechanism to indicate the

magnitude of benefit in addition to the level of evidence that can inform both national and international (e.g. ESMO) guidelines. Table 8 (downloaded from http://www.esmo.org/Policy/Magnitude-of-Clinical-Benefit-Scale/Scale-Evaluation-Forms) will be used to calculate the magnitude of clinical benefits.

Table 8: ESMO Magnitude of Clinical Benefit Scale – Form 2a: for Therapies That are not Likely to be Curative with Primary Endpoint of OS

Grade 4	**Mark with X if relevant**
HR ≤ 0.65 AND Gain ≥ 3 months	
Increase in 2 year survival alone ≥ 10%	
Grade 3	
HR ≤ 0.65 AND Gain 2.5-2.9 months	
Increase in 2 year survival alone 5 - <10%	
Grade 2	
HR > 0.65-0.70 OR Gain 1.5-2.4 months	
Increase in 2 year survival alone 3 - <5%	
Grade 1	
HR > 0.70 OR Gain <1.5 months	
Increase in 2 year survival alone <3%	

Abbreviations: HR = hazard ratio; OS = overall survival

Preliminary magnitude of clinical benefit grade (highest grade scored)

4	**3**	**2**	**1**

Quality of Life assessment /grade 3-4 toxicities assessment*

Does secondary endpoint quality of life show improvement?

Are there statistically significantly less grade 3-4 toxicities impacting on daily well-being*	

*This does not include alopecia, myelosuppression, but rather chronic nausea, diarrhea, fatigue, etc.

Adjustments

Upgrade 1 level if improved quality of life and/or less grade 3-4 toxicities impacting daily well-being are shown

Final adjusted magnitude of clinical benefit grade

5	4	3	2	1

7.5.6 Health Economics

A separate analysis plan for health economics will be provided.

7.6 Pharmacokinetics and Biomarkers

A separate analysis plan for PK and biomarker will be provided.

8.0 SAFETY ANALYSIS

The safety data will include AEs, SAEs, ECOG performance status, and clinical laboratory tests. Summaries will use the Safety population and will be presented separately for NKTR-102 and TPC treated patients. All safety data will be presented in data listings.

8.1 Adverse Events and Death

Adverse events will be coded by SOC and preferred term using MedDRA. Adverse event severity will be based on NCI CTCAE Grade (version 4.03).

A treatment emergent adverse event (TEAE) is defined as an AE that was not present prior to treatment with study drug but appeared following treatment, or was present prior to treatment initiation but worsened during treatment. An AE that was present at treatment initiation but resolved and then reappeared while the patient was on treatment is a TEAE (regardless of the intensity of the AE when the treatment was initiated). The treatment-emergent period will be defined as the period of time from the date of the first dose of study drug up to 30 days after the date of the last dose of study drug or the day prior to the initiation of subsequent anticancer treatment, whichever occurs first.

Adverse events leading to permanent discontinuation of study drug and/or leading to death will be listed and tabulated. The frequency of TEAEs and SAEs will be tabulated by preferred term and system organ class. The maximum NCI CTCAE grade and frequency of AEs will be summarized. For patient level of summaries (except for time to onset and time to resolution of diarrhea), patients with multiple occurrences of events of the same preferred terms and SOC will be counted once at the highest Grade and the strongest relationship to study drug for each preferred term, system organ class. Adverse events that are reported as possibly, probably, or definitely related to study drug will be counted as related to study drug; AEs with a missing relationship will be considered “Related” for this summary.

The following adverse events summaries will be provided:

- TEAEs by SOC and Preferred Term
- TEAEs by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs by SOC and Preferred Term and CTCAE Grade
- Grade 3 or Higher TEAEs by Preferred Term and by Descending Incidence of Overall Patients

- Serious TEAEs by Preferred Term and by Descending Incidence of Overall Patients
- Serious TEAEs by SOC and Preferred Term and CTCAE Grade
- TEAEs Related to Study Drug by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs Related to Study Drug by SOC and Preferred Term and CTCAE Grade
- Grade 3 or Higher TEAEs Related to Study Drug by SOC and Preferred Term and CTCAE Grade
- Serious TEAEs Related to Study Drug by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs Leading to Study Drug Discontinuation by SOC and Preferred Term and CTCAE Grade
- TEAEs Leading to Dose delay, Reduction, and Interruption by SOC and Preferred Term and CTCAE Grade
- TEAEs with Fatal Outcome by SOC and Preferred Term
- TEAEs Related to Study Drug with Fatal Outcome by SOC and Preferred Term

The following data listings will be produced:

- All AEs
- SAEs
- TEAEs Leading to Study Drug Discontinuation
- TEAE with Fatal Outcome

8.2 Clinical Laboratory Evaluations

Clinical laboratory tests will be conducted according to the Schedule of Assessments. Clinical laboratory tests will be performed by a designated central laboratory. Central laboratory results must be used to determine patient eligibility. For clinical management, local laboratories may be used for each site as the treating physician deems necessary by the treating physician. If a local laboratory result is considered a clinically significant AE, it should be reported as an AE.

For hematology and chemistry, the absolute values and change from baseline at each post-baseline study assessment will be summarized by descriptive statistics. In addition, the minimum and maximum absolute values and corresponding change from baseline across all post-baseline study assessments will be provided.

For selected laboratory tests with CTCAE grade, shift tables comparing post-baseline to baseline toxicity grade will be provided at each post-baseline study assessment. In addition, the worst shift from baseline will be summarized.

All central laboratory test results will be presented in data listings.

8.3 Additional Safety Analyses

Additional safety analysis will be performed to understand selected toxicities including diarrhea, neutropenia, neuropathy, and hypersensitivity reactions.

The total number of diarrhea events will be summarized. The number and percent of patients with at least one diarrhea event will be summarized by toxicity grade. A patient is only counted once using the highest grade. The number and percent of patients who had serious diarrhea events will be summarized.

Time to onset of first diarrhea event will be defined as the time between date of first dose and date of first diarrhea event and summarized using descriptive statistics. Time to onset of first grade 2 or higher, first grade 3 or higher diarrhea events will be defined and summarized similarly. The relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher diarrhea events will be plotted. Median duration of diarrhea events of any severity grade will be calculated for each patient. The median duration of diarrhea events for each treatment arm will be summarized by descriptive statistics using the median duration of each patient in that arm. Duration of grade 2 or higher diarrhea event, grade 3 or higher diarrhea event will be calculated and summarized similarly.

The number and percent of patients who discontinued study drug due to diarrhea with and without resolution will be summarized. For these patients, time to resolution will be calculated as the stop date of the event minus the date of last dose plus one and summarized using descriptive statistics.

In addition, selected analysis for diarrhea maybe repeated for patients who had prophylactic loperamide versus loperamide treatment only.

The total number of neutropenia events will be summarized. The number and percent of patients with at least one neutropenia event will be summarized by toxicity grade. A patient is only

counted once using the highest grade. The number and percent of patients who had serious neutropenia events will be summarized.

Time to onset of first neutropenia event will be defined as the time between date of first dose and date of first neutropenia event and summarized using descriptive statistics. Time to onset of first grade 2 or higher, first grade 3 or higher neutropenia events will be defined and summarized similarly. The relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher neutropenia events will be plotted. Median duration of neutropenia events of any severity grade will be calculated for each patient. The median duration of neutropenia events for each treatment arm will be summarized by descriptive statistics using the median duration of each patient in that arm. Duration of grade 2 or higher neutropenia event, grade 3 or higher neutropenia event will be calculated and summarized similarly.

Neuropathy-related events will be identified using preferred term by medical review. The total number of neuropathy-related events will be summarized. The number and percent of patients with at least one neuropathy events will be summarized by toxicity grade. A patient is only counted once using the highest grade. The number and percent of patients for each neuropathy-related event will be calculated under each preferred term by toxicity grade. Time to onset of neuropathy-related events, relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher neuropathy-related events, and duration of neuropathy-related events will be summarized similar to diarrhea-related events. The number and percent of patients with dose reduction and delay due to neuropathy-related events will be summarized. In addition, the number and percent of patient who discontinued study treatment due to neuropathy-related events with and without resolution will be summarized.

Neutropenia-related events will be identified, summarized and analyzed similarly to the neuropathy-related events.

Hypersensitivity events will be identified under selected SOC and preferred term by medical review. The total number of hypersensitivity events and the number and percent of patients who had at least one hypersensitivity events will be summarized . In addition, the number and percent of patients who had hypersensitivity events will also be summarized for each SOC, preferred term and severity. For patients with multiple events under a SOC or preferred term, patients will be counted only once using the highest severity. The duration of hypersensitivity events will be summarized by descriptive statistics. For patients who had more than one hypersensitivity events, the median duration of all events within the patient will be used for calculating the descriptive statistics. An analysis of hypersensitivity events and on-study development of polyethylene glycol (PEG) antibodies will be undertaken. This analysis will occur separately from the main analysis of the trial. The proportion of patients with PEG antibodies will be calculated. The timing of appearance of PEG antibodies will be examined.

Pulmonary toxicity events (potentially indicative of pneumonitis or interstitial lung disease) will be identified by medical review. The number and percent of patients who have pulmonary toxicity will be summarized for each treatment arm.

UGT1A1 polymorphisms may provide a useful diagnostic tool to predict *in vivo* glucuronidation of 7-ethyl-10-hydroxy-camptothecin (SN38), a molecule that is a major determinant of irinotecan and NKTR-102 metabolism and toxicity. All patients in this study treated with NKTR-102 will be tested for UGT1A1 alleles. The incidence and severity of diarrhea and neutropenia will be summarized for patients who randomized to NKTR-102 and with different UGT1A1 alleles (i.e., Poor, Intermediate, and Normal). These data will also be part of a larger population analysis across NKTR-102 clinical trials.

9.0 INTERIM ANALYSIS

A DMC Charter will be approved and finalized by the independent DMC members prior to the initiation of any interim analysis; the DMC Charter and meeting minutes will be submitted as part of the final Clinical Study Report.

Two-sided alpha of 0.001 will be used to test efficacy in OS at this interim analysis as the efficacy zone of the adaptive design. If OS does not reach statistical significance, the number of events for final OS analysis is determined per Mehta & Pocock (2011) to maintain approximately 80% conditional power for the CHW test. In order to prevent back calculation of the interim treatment effect as pointed out in (Liu and Hu 2016), the actual adaptation rule will take the step-function form and maintain approximate 80% conditional power. The adaptation rule – Event Size Adaption Rule for Clinical Study Protocol 15-102-14 will be included in the appendix of DMC charter and can only be accessed by DMC members, the Sponsor Executive Committee members and the sponsor design statistician who are not involved in the study conduct.

10.0 DATA MONITORING COMMITTEE

An independent DMC will be established to review interim efficacy and safety analyses results when approximately 82 events have occurred. The DMC consists of two external physicians and one external biostatistician.

The role of the DMC is to review the efficacy and safety interim analysis following the rules provided in the DMC charter and charter appendix.

A DMC charter will be approved and finalized by the DMC members prior to the initiation of any interim analysis. The DMC charter describes details including the primary responsibilities of the DMC, its relationship with other trial components, its membership, and the purpose and timing of its meetings, statistical monitoring guidelines to be implemented and statistical analysis for the open and closed sessions. The DMC Charter and meeting minutes will be submitted as part of the final Clinical Study Report.

11.0 REFERENCES

1. Anders CK, Deal AM, Miller R, et al. The Prognostic Contribution of Clinical Breast Cancer Subtype, Age, and Race Among Patients With Breast Cancer Brain Metastases. Cancer. 2010;117(8):1602-1611.

2. Brown PD, Asher AL, Ballman KV, et al. NCCTG N0574 (Alliance): A phase III randomized trial of whole brain radiation therapy (WBRT) in addition to radiosurgery (SRS) in patients with 1 to 3 brain metastases. Neuro-Oncology. 2015;17(suppl5):v45.5-v46

3. Crozier JA, Cornell LF, Rawal B, et al. Breast cancer brain metastases: Comprehensive review of tumor extent, histology and treatment on survival. Journal of Clinical Oncology. 2015;33 (supple; abstract 1059).

4. Cui L, Hung JHM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics 1999; 55: 853-857.

5. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247.

6. Halasz LM, Uno H, Hughes M, et al. Comparative Effectiveness of Stereotactic Radiosurgery Versus Whole-Brain Radiation Therapy for Patients With Brain Metastases From Breast or Non–Small Cell Lung Cancer. Cancer. 2016.

7. Khuntia D. Contemporary review of the management of brain metastasis with radiation. Hindawi Publishing Corporation. Advances in Neuroscience. 2015;2015:1-3.

8. Liu Y, and Hu M. Testing multiple primary endpoints in clinical trials with sample size adaptation. Pharmaceut. Statist. 2016, 15: 37–45. doi: 10.1002/pst.1724.

9. Mehta C, Pocock SJ. Adaptive increase in sample size when interim results are promising: A practical guide with examples. Statistics in Medicine 2011; 30: 3267-3284.

10. Mosteller RD. "Simplified calculation of body-surface area". N Engl J Med 1987; 317:1098. PMID 3657876.

11. Niwinska A, Murawska M, Pogoda K. Breast cancer brain metastases: differences in survival depending on biological subtype, RPA RTOG prognostic class and systemic treatment after whole-brain radiotherapy (WBRT). Annals of oncology. 2010;21(5):942-948.

12. Paul Brown et al. A phase III randomized trial of whole brain radiation therapy (WBRT) in addition to radiosurgery (SRS) in patients with 1 to 3 brain metastases . 2015 ASCO Annual Meeting.

13. Stone A.M. et al. Research outcomes and recommendations for the assessments of progression in cancer clinical trials from a PhRMA working group. 2011; 47: 1763-1771.

14. Subbiah IM, Lei X, Weinberg JS, et al. Validation and development of a modified breast graded prognostic assessment as a tool for survival in patients with breast cancer and brain metastases. Journal of Clinical Oncology. 2015;33(20):2239-2245.

ATTACHMENT 1: TABLE AND LISTING SHELLS FOR THE PRIMARY ANALYSIS

Table and listings shells for the planned statistical analysis for the study will be stored in a separate file.

ATTACHMENT 2: TABLE AND LISTING SHELLS FOR THE DMC ANALYSIS

Tables and listings for the interim analysis are outlined in the DMC report tables and listing shells.

Nektar Therapeutics

STATISTICAL ANALYSIS PLAN

A PHASE 3 OPEN-LABEL, RANDOMIZED, MULTICENTER STUDY OF NKTR-102 VERSUS TREATMENT OF PHYSICIAN'S CHOICE (TPC) IN PATIENTS WITH METASTATIC BREAST CANCER WHO HAVE STABLE BRAIN METASTASES AND HAVE BEEN PREVIOUSLY TREATED WITH AN ANTHRACYCLINE, A TAXANE, AND CAPECITABINE

Protocol Number: 15-102-14, Protocol Amendment 4.0

Date of Original SAP Version (v1.0): 16 April 2018

Date of SAP Amendment #1 (v2.0): 17 Dec 2018

TABLE OF CONTENTS

[redacted]

[redacted]

LIST OF TABLES

[redacted]

LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	anthracycline, taxane, capecitabine
ATCL	Anatomical, Therapeutic Chemical Level classification
BCBM	breast cancer brain metastases
BFI	Brief Fatigue Inventory
BMI	body mass index
BN-20	quality of life assessment specific to brain neoplasms
BSA	body surface area
CBR	clinical benefit rate
CMH	Cochran-Mantel-Haenszel
CNS	central nervous system
CR	complete response
DoR	duration of response
DMC	data monitoring committee
eCRF	electronic case report form
ECOG	Eastern Cooperative Oncology Group
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D-5L™	EuroQol 5D
ESMO-MCBS	European Society for Medical Oncology magnitude of clinical benefit scale
GCP	Good Clinical Practice
GPA	Graded Prognostic Assessment
HER	human epidermal growth factor receptor
HR	hazard ratio
HRQoL	health-related quality of life
IPFI	initial progression-free interval
IRT	interactive response technology
ITT	intent-to-treat
IV	intravenous
KM	Kaplan-Meier

[redacted]

Abbreviation	Definition
KPS	Karnofsky Performance Status
LDH	lactate dehydrogenase
LR	locally recurrent
MBC	metastatic breast cancer
MedDRA	Medical Dictionary for Regulatory Activities
MID	minimal important difference
MMRM	Repeated Measures Linear Mixed Effects Model
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	not evaluable
ORR	objective response rate
OS	overall survival
PD	progressive disease
PE	physical examination
PEG	polyethylene glycol
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	partial response
Q21d	once every 21 days
QLQ-C30	Quality of Life Core 30
QoL	quality of life
RANO-BM	Response Assessment in Neuro-Oncology—Brain Metastases
RECIST	Response Evaluation Criteria in Solid Tumors
RS	raw score
SAE	serious adverse event
SD	stable disease
SN38	7-ethyl-10-hydroxy-camptothecin; the active metabolite of irinotecan
SOC	system organ class
SRS	stereotactic radiosurgery
TEAE	treatment emergent adverse event
TNBC	triple-negative breast cancer
TPC	Treatment of Physician's Choice
UGT1A1	uridine diphosphate-glucoronosyl transferase 1A1
WHO-DDE	World Health Organization Drug Dictionary Enhanced

[redacted]

1.0 ADMINISTRATIVE STRUCTURE

This study will be managed via partnership between Nektar Therapeutics and a contract research organization. Central clinical laboratories will be used for processing of safety specimens, biomarkers, and pharmacokinetic (PK) samples. An interactive response technology (IRT) service provider will manage the randomization system, study drug, and comparator distribution and inventory management. A data monitoring committee (DMC) will be established to review the interim efficacy and safety data.

[redacted]

2.0 INTRODUCTION

This document describes the planned statistical analyses of the data captured according to Nektar Therapeutics Protocol 15-102-14 "A Phase 3 Open-Label, Randomized, Multicenter Study of NKTR-102 versus Treatment of Physician's Choice (TPC) in Patients with Metastatic Breast Cancer Who Have Stable Brain Metastases and Have Been Previously Treated with an Anthracycline, a Taxane, and Capecitabine" version dated 16 April 2018.

This Phase 3 study is conducted in accordance with the protocol, Good Clinical Practice (GCP), and all applicable regulatory requirements.

Analyses of PK, biomarkers, and pharmacoeconomic data will be addressed in separate analysis plans.

[redacted]

3.0 OBJECTIVES

3.1 Primary Efficacy Objective

To compare overall survival (OS) of patients who receive 145 mg/m^2 NKTR-102 given once every 21 days (q21d) with OS of patients who receive Treatment of Physician's Choice (TPC) selected from the following list of 7 single-agent intravenous (IV) therapies: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel. TPC will be administered per standard of care.

3.2 Secondary Objectives

- To compare the objective response rates (ORR) from NKTR-102 treatment with that of TPC; assessment of tumor outside the CNS will use the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1; assessment of central nervous system (CNS) metastases will use the Response Assessment in Neuro-Oncology Brain Metastases (RANO-BM)
- To compare progression-free survival (PFS) from NKTR-102 treatment with that of TPC; assessment of tumor outside the CNS will use RECIST version 1.1; assessment of CNS metastases will use RANO-BM
- To compare the clinical benefit rate (CBR) from NKTR-102 treatment with that of TPC (i.e., the proportion of patients having complete response [CR], partial response [PR], or stable disease [SD] for at least 4 months); CBR for peripheral lesions and for CNS lesions will be separately described
- To compare the duration of response (DoR) from NKTR-102 treatment with that of TPC
- To compare the time to CNS disease progression in those patients with CNS lesions present at study entry
- To compare the time to CNS recurrence in those patients without CNS lesions present at study entry
- To evaluate the safety profiles of NKTR-102 and TPC
- To compare health-related quality of life (HRQoL) from NKTR-102 treatment with that of TPC using the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Core 30 (QLQ-C30) with the specific brain neoplasms (BN-20) questionnaire, the EuroQol 5D (EQ-5D-5LTM) questionnaire, and the Brief Fatigue Inventory (BFI)
- To obtain PK data (in patients randomized to NKTR-102 only)

[redacted]

- To correlate presence of reduced function uridine diphosphate-glucuronosyl transferase 1A1 (UGT1A1) variants with NKTR-102 safety (in patients randomized to NKTR-102 only)
- To evaluate the pharmacoeconomic implications of NKTR-102 therapy using selected measures of health care utilization
- To evaluate the magnitude of clinical benefit using European Society for Medical Oncology magnitude of clinical benefit scale (ESMO-MCBS)

3.3 Exploratory Objective

- To identify biomarkers that correlate with response, PFS, and OS

[redacted]

4.0 STUDY ENDPOINTS

4.1 Primary Efficacy Endpoint

Overall survival, defined as date of randomization to the date of death from any cause. Patients who are lost to follow-up will be censored at the time they were last known to be alive. Patients who are alive at the time of OS analysis will be censored at the last date known to be alive for OS analysis.

4.2 Secondary Endpoints

- PFS with respect to all lesions outside the CNS per RECIST version 1.1
- PFS in brain metastasis per the RANO-BM
- PFS overall (CNS and peripheral)
- ORR defined as proportion of patients with a confirmed CR or confirmed PR per RECIST version 1.1 for lesions outside the CNS; RANO-BM for CNS lesions based upon the best response as assessed by the Investigator.
- CBR with respect to all lesions outside the CNS per RECIST version 1.1 defined as the proportion of patients with CR, PR, or SD and for SD it has to be at least 4 months from the date of randomization (≥120 days). CBR with respect to all CNS lesions per RANO-BM defined as the proportion of patients with SD or better and for SD it has to be at least 4 months (≥120 days)
- DoR with respect to all lesions outside the CNS defined for patients who have a confirmed CR or PR as the date from first documented CR or PR per RECIST 1.1 to the date of the documentation of disease progression or death due to any cause, whichever is earlier
- HRQoL using the EORTC QLQ-C30 with BN-20 questionnaire, BFI and EQ-5D-5L
- Evaluation of the pharmacoeconomic implications of NKTR-102 therapy using selected measures of health care utilization
- Magnitude of clinical benefit using ESMO-MCBS
- Incidence of grade 3 and higher toxicities including adverse events (AEs) and clinical significant laboratory tests, serious AEs (SAEs), AEs with outcome of death, AEs leading to dose reduction, dose delay, and dose withdrawal. Duration and intensity of study drug exposure.

[redacted]

- AEs of special interest including neutropenia, diarrhea, neuropathy, and hypersensitivity.

[redacted]

5.0 OVERALL STUDY DESIGN AND PLAN

5.1 Study Design

This open-label, randomized, two-arm, multicenter, international, Phase 3 study of NKTR-102 in patients with breast cancer brain metastases (BCBM) who have stable brain metastases will evaluate single-agent NKTR-102 (145 mg/m^2 q21d) in patients who have previously received an anthracycline, a taxane, and capecitabine (ATC) versus a comparator arm consisting of an active single-agent TPC.

In Group A, NKTR-102 will be administered at a dose level of 145 mg/m^2 on a q21d schedule as a 90-minute IV infusion on Day 1 of each treatment cycle. In Group B, TPC will be administered per standard of care. Patients randomized to TPC will receive single-agent IV chemotherapy, limited to choice of one of the following 7 agents: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel.

Up to 220 patients will be randomized using a 1:1 randomization ratio. The following stratification factors for randomization were selected to balance the treatment groups for known factors that influence prognosis in patients with metastatic breast cancer and brain metastases:

- geographic region (US versus the rest of world)
- hormone and human epidermal growth factor receptor (HER) status (triple-negative breast cancer [TNBC], HER2+/HR any, and HR+/HER2-)
- Eastern Cooperative Oncology Group (ECOG) Performance Status (ECOG 0 and ECOG 1)

Cross-over from Group B to Group A is not permitted. The duration of the study will be 26-47 months to final analysis. An independent DMC will assess interim safety and efficacy data.

All patients will undergo tumor assessments performed at the participating study center or at a radiology facility associated with the site. Tumor measurements will be evaluated locally per RECIST 1.1 and RANO-BM criteria. All tumor imaging (head, chest, abdomen and other as appropriate) and digital photography will be forwarded to a central imaging facility to permit blinded independent review (local assessment will be used for patient management).

[redacted]

5.2 Study Medications

5.2.1 Arm A: NKTR-102

5.2.1.1 Dosage and Administration

Body surface area (BSA) will be determined before the start of each cycle, based on institutional guidelines and will be capped at 2.4 m^2. In the instance where there are no institutional guidelines, baseline height and most recent weight will be used to calculate BSA. Each patient's NKTR-102 dose will be determined by multiplying the most recent BSA (capped at 2.4 m^2) by the starting dose of 145 mg/m^2. NKTR-102 for Injection will be administered as an IV infusion over 90 minutes (± 15 minutes). Premedications are not required to be administered prior to the initial infusion, but may be used for an individual patient, as needed.

5.2.2 Arm B: Treatment of Physician's Choice

Patients randomized to TPC will receive single-agent chemotherapy, limited to one of the following 7 agents: eribulin, ixabepilone, vinorelbine, gemcitabine, paclitaxel, docetaxel, or nab-paclitaxel. TPC must consist of single-agent IV therapy, not combination therapy. Choice of Group B agent (TPC) for an individual patient will depend on the TPC drug products available at each medical center. For TPC products without generic versions (e.g., eribulin, ixabepilone, and nab-paclitaxel) the branded product must be commercially available at the medical center. Selection of a TPC drug product should be based on what would have been offered to the patient within that medical center if the patient was not participating in a clinical study.

[redacted]

6.0 STATISTICAL CONSIDERATIONS

6.1 General Considerations

Although this is an open label study, analyses for treatment comparisons prior to final database lock will not be performed except for the pre-planned interim safety and efficacy analyses as described in Section 9.0 of this document and in the DMC charter.

Blinding for aggregated treatment information during the trial and data unblinding for interim analysis and after final database lock will be provided in separate documents.

Summary statistics for continuous variables will include the mean, standard deviation, median, minimum, maximum, and 25% and 75% quartiles. The mean will be presented to one decimal place beyond which the data were captured. The standard deviation will be presented to two decimal places beyond which the data were captured. The minimum and maximum will be presented to the precision with which the data were captured.

Categorical variables will be presented as frequency counts and percentages. A row or column denoted 'Missing' will be included in count tabulations where necessary to account for dropouts and missing values. Percentages will be rounded to 1 decimal place and the percent will be suppressed when the count is zero. The denominator will be the number of patients in the population of interest unless otherwise noted.

Time-to-event variables will be analyzed using the Kaplan-Meier (KM) method. The KM estimates for quartiles and the 95% confidence interval for the median will be presented. All time to event variables will be plotted using the KM method.

Data listings will be created to support each table and to present all data. Data listings will be presented by treatment group and patient number.

6.2 Determination of Sample Size

The study is powered for detecting superiority of NKTR-102 compared with TPC in OS and up to 220 patients will be enrolled. The number of OS events needed to provide 80% conditional power for the final analysis will be determined at an interim analysis when approximately 82 OS events are available using the promising zone adaptive method (Mehta & Pocock, 2011). The minimum and maximum number of events for the final analysis will be 106 and 191, respectively. Two-sided alpha of 0.001 will be used to test efficacy at the interim analysis (efficacy zone-as part of promising zone design). One hundred ninety-one (191) events will be able to demonstrate statistical significance for any observed hazard ratio of 0.75 or better, or a corresponding median difference of approximately 2 months or better if the OS median for the control arm is 6 months and the proportional hazard assumption is approximately met. The

[redacted]

detailed event size adaptation rules based on conditional power will be provided in the DMC Charter Appendix – Event Size Adaption Rule for Clinical Study Protocol 15-102-14.

6.3 Analysis Populations

Intent-to-Treat (ITT) Population: The ITT population includes all patients who are randomized in the study. The primary endpoint of OS and secondary efficacy analyses (except ORR and DoR) will utilize the ITT population. Patients will be analyzed by the treatment arm to which they are randomized.

Response Evaluable Population: The Response Evaluable population includes all patients who are randomized in the study with measurable disease in the periphery by RECIST 1.1 at baseline as determined by the Investigator. The secondary endpoint analyses of ORR and DoR will utilize the Response Evaluable Population. For DoR, only patients having confirmed CR/PR will be included in the analysis. Patients will be analyzed by the treatment arm to which they are randomized.

Safety Population: The Safety Population consists of all patients who are randomized and receive at least one dose (or partial dose) of study drug (NKTR-102 or TPC). Safety analyses will be conducted using this population and the treatment arm (A or B) received, based on the first dose of study drug.

6.4 Handling of Missing Data

Missing data will be handled as follows:

- For prior systemic cancer therapies and medical history of cancer, the study day corresponding to the start and stop date of the regimen will be calculated when calculating duration of the therapy/history and relevant time to the initiation of the study, etc. For partially missing start dates, missing day of the month will be imputed as the first of the month and missing month will be imputed as January. For partially missing stop dates, missing day of the month will be imputed as the last day of the month and missing month will be imputed as December and at least 1 day after the start date. No imputation will be done if the year is missing.
- For prior radiotherapy and surgery, the relevant time to the initiation of the study, etc. will be calculated using the calculated study day corresponding to the date of radiotherapy or date of procedure for surgery The imputation rules described for prior systemic cancer therapy will be applied to prior radiotherapy and surgery.
- For time from initial diagnosis to informed consent and time from diagnosis of metastatic disease to informed consent, partial dates for initial diagnosis and diagnosis of metastatic disease will be imputed as follows:

[redacted]

- No imputation will be done if the year is missing.
- If the year is before informed consent date then missing days will be imputed as the first day of the month and missing months will be imputed as July.
- If the year is the current year of informed consent date then missing days will be imputed as the first day of the month and missing months will be imputed as January.

- For duration of AEs, partial dates for **start** of AE will be imputed as follows:
  - Missing day, month, and year should be queried. In case of non-resolution of missing month and/or year, no imputation will be performed.
  - Start day of AE is missing and the year is same as Cycle 1 Day 1 (C1D1)
    - If the reported month of occurrence of AE is after the month of C1D1 dose then missing day will be imputed as the first day of the month of occurrence of AE.
    - If the reported month of occurrence of AE is the month of C1D1 dose then the missing day will be imputed as the same day as C1D1.
  - Start day of AE is missing and the year is after the year of C1D1
    - Missing day will be imputed as the first day of the month of occurrence of AE
- For duration of AEs, partially missing dates for **stop** of AE will be imputed as follows:
  - Missing day, month, and year are not allowed and should be queried. In case of non-resolution of missing month and year, no imputation will be performed.
  - If only the day is missing, the last day of that month, the date of discontinuation from the study, or the date of death, whichever is earlier, will be used as the stop date.
  - If month and day are missing, then December 31st, the date of discontinuation from the study, or the date of death, whichever is earlier, will be used as the stop date.
- For determination of prior medication, any medication with a start date prior to C1D1 will be classified as prior medication regardless of when the stop date is. Missing or partial dates will be handled as follows:
  - If missing day and/or month of the start date, the medication will be classified as prior unless the month and/or year of the start date is after C1D1
  - A medication with completely missing start date will be classified as prior, unless the stop date is on or after C1D1
- For determination of concommitant medication, the following will be classified as concomitant medication:
  - Any medication with a start date prior to or on C1D1 and continued to take after C1D1

[redacted]

  - Any medication with a start date after C1D1, but prior to or on the last dose date
- Missing or partial dates for concomitant medication will be handled as follows:
  - If missing day and/or month of the start date, the medication will be excluded from concomitant if the month and/or year of the start date is after the last dose date
  - If missing day and/or month of the stop date, the medication will be excluded from concomitant if the month and/or year of the stop date is prior to C1D1
  - A medication with completely missing start and stop dates will be classified as concomitant

Handling of missing data for analysis of PFS and QoL are described in the statistical analysis sections below.

No imputation of other missing data is planned.

6.5 Stratification and Pooling

For stratified analyses, the stratification factors (Region, Receptor Status, and ECOG performance status) entered into the interactive voice response system at the time of randomization will be used. If the proportion of patients assigned to incorrect stratums exceeds 5% of the total patients randomized, a sensitivity analysis using stratums based upon data in the clinical database will be performed. In the BEACON trial, the incidence of patients with HER2+ breast cancer and brain metastases was 13.4% (9 out of 67 patients); the incidence of ECOG=0 patients was 23.9% (16 out of 67 patients). To avoid unstable estimates, in this trial, if any cells (defined by the three stratification factors) have fewer than 5 patients, patients may be pooled (for example, patients with HER2+ breast cancer may be pooled with patients who have HR+/HER- or TNBC, whichever group has fewer patients, i.e., Receptor Status will be collapsed into two levels from three levels).

6.6 Analytic Definitions

General:

- Baseline: Baseline will be defined as the last assessment result on or before the first study drug administration.

- Study day: There is no study day zero. Day 1 is the date of first study drug administration. For post-treatment events, study day is calculated as:

  Date of visit or assessment – date of Day 1 + 1

  For pre-treatment events, study day is calculated as:

[redacted]

Date of visit or assessment – date Day 1

Baseline Characteristics:

- Body Mass Index (BMI) (kg/m^2) is defined as: (Weight in kg) / (Height in m)^2
- BSA will be calculated based on the following formula (Mosteller, 1987) and used for summary purposes:

$$BSA\ (m^2) = ([Height\ (cm) \times Weight\ (kg)]/\ 3600)^{1/2}$$

Exposure:

- Exposure duration (days): date of last dose – date of first dose (Cycle 1 Day 1) + 1.
- Number of cycles: Total number of complete or partial treatment cycles the patient received.
- Cumulative dose (mg): Total actual dose (mg) the patient received across all cycles, defined as the sum of actual dose (mg) received across all cycles.
- Average dose per infusion: Mean of (the actual dose receive/BSA at that cycle) across all cycles.
- Duration of infusion: Completion time of infusion – start time of infusion.
- Calculated cumulative dose level (mg/m^2): Cumulative dose (mg) divided by the average BSA (m^2) across all cycles.

<u>For Arm A (NKTR-102), the following parameters will be calculated:</u>

- Calculated dose (mg) per cycle: planned dose level (145 mg/m^2) x BSA (m^2)
- Dose intensity (mg/m^2/week): [Calculated cumulative dose level (mg/m^2) / (exposure duration (days) + 20 days)] × 7
- Expected dose intensity (mg/ m^2/week) = (145 mg/m^2)/(3 week) = 48.3 mg/m^2/week
- Relative dose intensity (%): ((Dose intensity) / (expected dose intensity)) × 100

<u>For Arm B, the following parameters will be calculated:</u>

- Dose intensity (mg/m^2/week): {Calculated cumulative dose level (mg/m^2) /[(exposure duration (days) + ADD ON PERIOD – 1 day)]} x 7 ; where ADD ON PERIOD = cycle length as planned for Cycle 1 – (date of last dose of the last cycle – date of 1st dose in the last cycle)

[redacted]

- Expected dose intensity (mg/m^2/week): {[initial planned Cycle 1 Day 1 dose (mg/m^2) × number of doses given in Cycle 1] / cycle length as planned for Cycle 1} × 7
- Relative dose intensity (%): ((Dose intensity) / (expected dose intensity)) × 100.

Visit Windows:

For statistical summary by visit or cycle, data will be assigned to a derived visit window. If more than one eligible result is reported within the same window, the result collected closest to the target day will be used for the descriptive statistics and the worst value (most conservative) result will be used in the shift table or summary for abnormalities.

The following visit windows will be used to assign analysis visits for by-cycle summaries for assessments/tests conducted at each treatment cycle:

Planned Visit	Target Day	Starting Hour/Day for Visit Window	Ending Hour/Day for Visit Window
Screening	Screening	Day -28	Day -1
Cycle 1 Day 1	Cycle 1 Day 1	< 0 hr/same day of Cycle 1 Day 1	< 0 hr of Cycle 1 Day 1
Cycle 2 Day 1	Cycle 2 Day 1	> 0 hr of Cycle 1 Day 1	<0 hr of Cycle 2 Day 1
Rest of Cycles Day 1	Cycle x Day 1	> 0 hr of Cycle (x-1) Day 1	< 0 of Cycle x Day 1

Assessments/tests performed after the last treatment cycle will be summarized under the patient's last treatment cycle + 1.

The following visit windows will be used to assign analysis visits for by-week summaries for tumor assessments:

Planned Visit	Target Day	Starting Day for Visit Window	Ending Day for Visit Window
Screening	Screening	Day -28	randomization
Week 8	56 days since randomization	1 day since randomization (and post Cycle 1 Day 1)	83 days since randomization
Week 16	112 days since randomization	84 days since randomization	139 days since randomization
Week 24	168 days since randomization	140 days since randomization	209 days since randomization
Week 12*X (X ≥ 3)	12*X*7 days since randomization	(12*X-6)*7 days since randomization	(12*X+6)*7 – 1 days since randomization

[redacted]

Depending on the analysis, single values may be required for each analysis window. For example change from baseline by visit usually requires a single value, whereas a time-to-event analysis does not. If there are multiple valid, non-missing records in an analysis window, the following selection rules will be used to select a single value as needed:

- For baseline value,
  - In general, the baseline value will be the last, non-missing value on or prior to the first dose date of study drug, unless otherwise specified. If multiple measurements occur on the same day, the last non-missing value prior to the first dose date of study drug will be considered the baseline value. If multiple measurements occur at the same time or the time is not available, the average of the measurments (for continuous data) will be considered the baseline value.
- For post-baseline value,
  - The record closest to the target day for that visit will be chosen
  - If there are 2 records equidistance from the target visit day, the later record will be chosen
  - If there is more than 1 record on the selected day, the average will be taken, unless otherwise specified.

[redacted]

7.0 STATISTICAL ANALYSIS

7.1 Patient Disposition

A summary of patient disposition will display the number of patients who were randomized and who comprised each analysis population by treatment arm. In addition, the number of patients who discontinued study drug and the number of patients who exited the study, both overall and by reason, will be presented by treatment arm.

The three randomization stratification factors, geographic region , and receptor status based upon local review of disease (HER2+, HR+/HER2-, and TNBC) will be tabulated as documented on the electronic case report form (eCRF) and as recorded in the IRT system by treatment arm as randomized. Discrepancies between those two sources will be identified and summarized.

All disposition data will be presented in a data listing.

7.2 Protocol Deviations

Important protocol deviations will be defined in a separate document. These protocol deviations will be captured by searching relevant data fields reported in the clinical database using computer algorithm and site monitoring by clinical research associates tracked in the clinical tracking system. Final determination of important deviations will be reviewed by the study team at regular intervals throughout the trial (including clinical study manager, the Medical Monitor, and the statistician).

The number and percentage of patients in each important and non-important protocol deviations category and sub-category will be summarized by treatment arm for the ITT population. All protocol deviations will be listed.

7.3 Demographics and Baseline Characteristics

The following baseline data will be summarized and listed for all analysis populations: age (years), sex, race, ethnicity, ECOG performance status, reproductive status, height (cm), weight (kg), BSA (m^2), BMI (kg/m^2), and key laboratory tests including absolute neutrophil count (ANC), hemoglobin, platelets, bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), albumin, creatinine, and creatinine clearance.

7.3.1 Medical and Cancer History

Cancer history will include time since initial breast cancer diagnosis, stage, histology and receptor status at initial breast cancer diagnosis, time since diagnosis of locally advanced or

[redacted]

metastasis breast cancer, initial disease-free interval, time since diagnosis of metastatic disease to the brain, time since most recent prior treatment for BCBM such as whole brain radiation, stereotactic radiation and/or surgical resection, type of prior therapy for BCBM, radiographic evidence of metastatic brain lesions at baseline (Yes/No), number of sites of disease, receptor status at last biopsy, and Graded Prognostic Assessment (GPA) score (Table 5) at study entry.

Medical history collected at screening will be mapped by the Dictionary for Regulatory Activities (MedDRA) and will be summarized by system organ class (SOC) and preferred term for the ITT population. For summary tables, a patient will be counted only once per SOC and preferred term.

7.3.2 Prior Systemic Cancer/Oncology Therapies

Prior systemic cancer/oncology therapies will be tabulated using the World Health Organization Drug Dictionary Enhanced (WHO-DDE) and Anatomical, Therapeutic, or Chemical Level (ATCL)-2 classifications and preferred term. If the ATCL-2 classification is missing, the next non-missing higher level of classification will be used (Level 1). If a patient reports the same medication multiple times, then the frequency of that medication will be incremented by only one in the applicable arm. As with the medication, if a patient reports multiple medications within the same ATCL-2 classification, then the frequency of that ATCL-2 classification will be incremented by only one in the applicable arm. Percentages will be calculated using the number of patients in the ITT population.

The number and percent of patients who received prior anthracycline, taxane, and capecitabine treatment will be summarized. The number and percent of patients who received prior hormonal therapies, HER2-directed therapies, cytotoxic chemotherapy regimens for breast cancer, neoadjuvant regimens, adjuvant regimens, regimens for locally recurrent, metastatic disease, secondary primary, and chemoradiation will be calculate using the ITT population. Duration of prior anthracycline, taxane, capecitabine and eribulin use will be summarized descriptively. If a patient received multiple treatments of a therapy, the sum of duration will be calculated.

All prior systemic cancer/oncology therapies will be listed.

7.3.3 Surgical History and Prior Radiotherapy

The number and percent of patients who had had prior cancer-related CNS and non-CNS surgery will be summarized by type of surgery for each treatment arm using the ITT population.

The number and percent of patients who had prior radiotherapy will be summarized by site and type of the radiotherapy for each treatment arm using the ITT population.

[redacted]

All prior radiotherapy and surgical procedures will be listed.

7.3.4 Concomitant Radiotherapy, Procedures and Surgery for CNS Lesion.

The number and percent of patients who had concomitant radiotherapy will be summarized by site and type of radiotherapy for each treatment arm using the ITT population.

The number and percent of patients who had concomitant procedures will be summarized by each coded procedure preferred term for each treatment arm using the ITT population.

The number and percent of patients who had concomitant surgery for CNS lesion will be summarized by type of surgical procedure for each treatment arm using the ITT population.

All concomitant radiotherapy, procedures, and surgery for CNS lesion will be listed.

7.4 Treatments and Medications

7.4.1 Prior and Concomitant Medications

Prior and concomitant medications will be coded to ATCL and preferred drug name using WHO-DDE.

Concomitant medications are defined as medications taken on or after the date of first dose, including medications initiated prior to the date of first dose and continued during treatment, and medications initiated on or after the date of first dose.

Concomitant medications will be tabulated for the Safety Population by WHO-DDE ATCL-2 classifications and preferred term. If the ATCL -2 classifications is missing, the next non-missing higher level of classification will be used (Level 1). If a patient reports the same medication multiple times, then the frequency reported for that medication will be incremented by only one. As with the medication, if a patient reports multiple medications within the same ATCL-2 classification then the frequency for that ATCL-2 classification will be incremented by only one. Percentages will be calculated using the total number of patients in the Safety population.

Prior medications are defined as medications taken starting prior to the first dose. Prior medications will be summarized for the ITT population using the same analytical procedures as concomitant medications.

Prior and concomitant medications will be presented in a data listing.

[redacted]

7.4.2 Subsequent Anti-Cancer Therapy, Radiotherapy, and Surgery.

Subsequent anti-cancer systemic therapy will be summarized based on the ITT population. Subsequent anti-cancer therapy will be tabulated using the WHO-DDE and ATCL -2 classifications and preferred term. The number and percent of patients who took at least one subsequent anti-cancer therapy will be calculated and presented by treatment group, ATCL-2 classifications, and preferred term.

The number and percent of patients who had post-treatment radiotherapy will be summarized by site and type of radiotherapy for each treatment arm using the ITT population.

The number and percent of patients who had post-treatment surgery for CNS lesion will be summarized by type of surgical procedure for each treatment arm using the ITT population.

7.4.3 Exposure / Study Treatments

Overall exposure to study treatments (Arms A and B) and study treatment administration for each cycle and all cycles combined will be summarized for the Safety population. Overall exposure to study treatments will be summarized in terms of exposure duration, number of cycles, cumulative dose (mg), and average infusion duration (minutes) and will be calculated for both treatment arms. Average dose per infusion (mg/m^2), dose intensity (mg/m^2/week), and relative dose intensity will be calculated. The number of patients with dose reduction, dose interruption, and dose delay will be tabulated along with their reason.

Association between number of treatment cycles received and baseline characteristics including geographic region, receptor status, ECOG, GPA score, baseline ANC, hemoglobin and platelets, baseline AST and creatinine) will be examined.

For patients who are randomized to NKTR-102, the quartiles of dose intensity and duration of exposure will be calculated. Patients will be categorized into ≤25%, >25% to ≤ 50%, > 50% to ≤75%, and >75% groups. For each group, the number and percent of patients who have selected toxicity (neutropenia, diarrhea, neuropathy, and hypersensitivity) will be summarized. In addition, the number and percent of patients who have Grade 3 or higher selected toxicity will be summarized.

All study exposure data will be presented in a data listing.

[redacted]

7.5 Efficacy Analysis

7.5.1 Analysis of Primary Endpoint

The primary efficacy endpoint, which will be calculated on the ITT population, is the overall survival defined as the time from the date of randomization to death from any cause. Patients will be followed until their date of death, loss to follow-up, and withdrawal of consent for further follow-up for survival. Patients who are lost to follow up will be censored at last date shown to be alive. Patients who do not have date of death at final database closure will be censored at the last date shown to be alive for OS analysis. Patients who do not have any follow up since randomization will be censored at the date of randomization.

Two-sided alpha of 0.001 and 0.049 are allocated to the interim analysis and the final analysis respectively to control the overall Type I error rate at 0.05 level for primary OS analysis. At final analysis, the primary analysis of OS for statistical significance claim will be based on the CHW version of the logrank test statistic with weights $\sqrt{(2/3)}$ and $\sqrt{(1/3)}$ (Cui, Hung, & Wang, 1999). Based on the correlation between interim and final CHW test statistic, statistical significance can be claimed if the two-sided p-value for CHW test statistic is no greater than 0.0499 (see detailed calculation in the Appendix). The conventional logrank test with equal weights for every patient will be conducted as a sensitivity analysis.

The median survival times and their 95% confidence intervals as well as survival curves will be estimated using the Kaplan-Meier method and will be summarized by treatment group.

The proportion of patients that are alive at 6 and 12 months and the corresponding 95% confidence intervals will be calculated as well. A hazard ratio comparing NKTR-102 to TPC and its 95% confidence interval will be calculated using a Cox regression model adjusting for geographic region, receptor status and ECOG performance status.

Above analysis of OS (i.e. log rank test and Cox regression model) will be repeated excluding patients from both arms who selected ixabepilone as their TPC agent prior to randomization.

If more than 10% of study population (more than 22 patients) received concomitant stereotactic radiosurgery (SRS) during the study, the proportion of patient received SRS during the study between treatment groups will be compared using Fisher's exact test. Use of post-study SRS and its impact on OS will also be examined.

The impact of SRS on OS will be evaluated by a Cox regression model comparing patients who received SRS to those who did not. The treatment effect will be adjusted for SRS by a Cox regression model with treatment arm and usage of SRS as covariates.

[redacted]

Subgroup analyses of overall survival will be performed to assess whether the treatment effect is concordant among subgroups. The following variables will be used to define subgroups:

- Number of prior regimens (≤3, and ≥4)
- Age (< 65 versus ≥ 65)
- Race (white, black, asian, other)
- Time from initial breast cancer diagnosis to randomization (≤ 4 years versus > 4 years)
- Time since LR/MBC diagnosis (≤ 2 years versus > 2 years)
- Time since BCBM diagnosis (≤ 6 months versus > 6 months)
- Time since local therapy for BCBM (≤ 3 months versus > 3 months)
- Initial progression-free interval (IPFI, < 2 year versus >= 2 years; IPFI will be calculated from the date of first dose to the date of progressive disease (PD) for the first systemic cancer treatment for breast cancer in locally recurrent or metastatic setting. If progression date is not reported, the earliest start date of the next regimen or the date of last scan prior to randomization is used.
- Patients with evidence of brain metastatic lesions on head imaging at entry (Yes versus No)
- Patients with 1-3 brain lesions versus >3 brain lesions
- Prior whole brain radiation vs. stereotactic radiation (with or without surgical resection)
- GPA score 0-2 versus 2.5-4 at study entry
- Steroid use for control of neurological signs/symptoms at baseline (Yes versus No)
- Initial disease free interval (IDFI, ≤ 2 years versus > 2 years; IDFI will be calculated as the duration between the initial diagnosis of breast cancer and the start date of first treatment in a locally recurrent or metastatic setting)
- Geographic region (US versus other)
- Geographic region (by Country)
- ECOG performance status (0 versus 1)

[redacted]

- Receptor status based on local review of disease (TNBC versus HER2+ versus Other)
- HER2 status (positive versus negative versus unknown)
- Hormone receptor status (positive versus negative versus unknown)
- Number of sites involved (≤2 versus ≥3); disease site will be mapped based on the Target or Non-target tumor location reported in the database
- Prior eribulin use (Yes versus No)
- Prior hormonal therapy (Yes versus No)
- Metastasis in liver (Yes versus No, Yes is defined as at least one Target or Non-target tumor located in liver at baseline)
- Metastasis in lung (Yes versus No, Yes is defined as at least one Target or Non-target tumor located in lung at baseline)

The benefit of NKTR-102 compared to TPC will be evaluated by a hazard ratio (NKTR-102/TPC) with its 95% confidence interval based on a Cox regression model for each subgroup. The hazard ratios and their 95% confidence intervals will be displayed in a forest plot. Subgroup analyses above may not be performed if the sample size is too small to provide accurate estimate.

7.5.2 Analysis of Secondary Efficacy Endpoints

Analysis for secondary endpoints will not include any adjustment for multiplicity. Statistical tests will be two-sided with a type I error rate of 0.05.

7.5.2.1 Progression-Free Survival (Outside the CNS)

Progression-free survival analyses will be performed for the ITT population.

The revised RECIST 1.1 guidelines (Eisenhauer, 2009) will be used to determine disease progression. Date of PD is the earliest date of the imaging method to determine PD, or if the patient is followed solely by physical examination (PE) the date of the PE showing PD. The date of global deterioration" or "symptomatic deterioration" will not be used as the date of PD.

Progression-free survival is defined as the time from the date of randomization to the earliest evidence of documented PD or of death from any cause. For the primary analysis of PFS, the censoring methods described in Table 1 will be used. These rules are considered as the

censoring criteria described in (Stone. 2011) and are consistent with censoring rules described in the FDA Guidance for Industry document – Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Appendix 3 Table A (PFS includes documented progression only) (FDA 2007).

Table 1: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-off	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD or death	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to start date of the anti-cancer therapy	Censored
Prematurely discontinued study drug due to non-progression related reasons	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to study drug discontinuation	Censored

Abbreviations: PD = progressive disease; PE = physical examination

The first sensitivity analyses of PFS will be performed based on the censoring methods provided in Table 2. The difference between the rules described in Table 1 and Table 2 is that patients who prematurely discontinue the study drug due to reasons other than progression will be censored in Table 1 while in Table 2, they will not be censored. These rules are consistent with censoring criteria described in (Stone. 2011). These rules are also consistent with the considerations of informative censoring as outlined in the Guideline on the Evaluation of Anticancer Medicinal Products in Man – Appendix 1 (EMA 2012).

[redacted]

Table 2: Alternative censoring approach: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD or death	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to start date of the anti-cancer therapy	Censored
Prematurely discontinued study drug due to non-progression related reasons prior to observing PD or death	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to study drug discontinuation	Progressed

Abbreviations: PD = progressive disease; PE = physical examination

The second sensitivity analysis of PFS will be performed based on the censoring methods described in Table 3. The differences between the rules described in Table 1 and Table 3 are that patients who prematurely discontinue the study drug due to reasons other than progression or patients who start any new anti-cancer therapy will be censored in Table 1 while in Table 3 they will not be censored. These rules are consistent with the censoring criteria described by Stone, 2011. These rules are also consistent with the considerations of informative censoring as outlined in the Guideline on the Evaluation of Anticancer Medicinal Products in Man – Appendix 1 (EMA 2012).

[redacted]

Table 3: The second alternative censoring approach: Date of Progression or Censoring for Progression-free Survival

Situation	Date of Disease Progression or Censoring	Outcome
No baseline assessments for tumor response	Date of randomization	Censored
Disease progression or death after two or more consecutive missed tumor response assessments	Date of last scan/PE for tumor assessment showing no evidence of disease progression that is before the first missed visit	Censored
Not known to have progressed or died according to data in the database as of data-cut-off	Date of last scan/PE for tumor assessment showing no evidence of disease progression	Censored
Disease progression reported on multiple response assessments or multiple scans were performed	Date of earliest evidence of PD	Progressed
Disease progression or death after one missed tumor response assessments	Date of earliest evidence of PD	Progressed
Death without PD or new anti-cancer therapy	Date of death	Progressed
Starting anti-cancer therapy prior to observing PD or death	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to start date of the anti-cancer therapy	Progressed
Prematurely discontinued study drug due to non-progression related reasons prior to observing PD or death	Date of last scan/PE for tumor assessment showing no evidence of disease progression prior to study drug discontinuation	Progressed

Abbreviations: PD = progressive disease; PE = physical examination

Progression-free survival will be compared with a two sided, stratified, log-rank test with the same stratification factors that were used for randomization. Median PFS time, its 95% confidence intervals as well as 25% and 75% quartiles of PFS times will be summarized using KM method for each treatment group.

A single hazard ratio (NKTR-102/TPC) and its 95% confidence interval will be calculated using a stratified Cox regression model adjusting for geographic region, hormone receptor status, and baseline ECOG performance status.

7.5.2.2 Progression-Free Survival in Brain Metastasis

Progression-free survival in brain metastasis for patients with CNS lesions at baseline will be analyzed separately from patients without CNS lesions at baseline. For patients with CNS lesions at baseline, progression-free survival in brain metastasis is defined as the time from the date of randomization to the earliest evidence of documented PD per RANO-BM in brain metastases or death from any cause. For patients without CNS lesion at baseline, progression-free survival in

[redacted]

brain metastasis is defined as the time from the date of randomization to the earliest evidence of disease recurrence in the CNS or death from any cause.

The censoring rule specified in Table 1 will be used, and no sensitivity analyses are performed. The same statistical methods used for the analysis of PFS outside the CNS will be used for the analysis PFS-brain metastasis.

7.5.2.3 Progression-Free Survival (Overall)

Progression-free survival (CNS and peripheral) is defined as the time from the date of randomization to the earliest evidence of documented PD in either CNS per RANO-BM or peripheral per RECIST 1.1, CNS recurrence for the patients without CNS at baseline, or death from any cause. For patients who are censored per RECIST 1.1 (Table 1) but have date of PD per RANO-BM, the date of disease progression per RANO-BM will be used. For patients who are censored per RANO-BM but have date of PD per RECIST 1.1, the date of disease progression per RECIST 1.1 will be used. For patients who are censored by both RECIST 1.1 and RANO-BM, patients will be censored on the date of the earliest censoring per RECIST 1.1 and RANO-BM. No sensitivity analyses are performed.

The same statistical methods used for the analysis of PFS (outside CNS) will be used for the analysis of PFS (Overall).

7.5.2.4 Objective Response Rate

Objective response rate analyses will be performed for the Response Evaluable population and defined as the proportion of patients with a confirmed CR or PR for lesions outside the CNS per RECIST 1.1. The analysis of ORR will be performed using Fisher's exact test between the two treatment arms. Clopper-Pearson exact 2-sided 95% confidence limits will be calculated for the proportion of patients with ORR in each arm.

The primary analysis of ORR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

In addition, the best response using response categories CR, PR, SD, PD, and not evaluable (NE) per RECIST 1.1 outside CNS and per RANO-BM in CNS will be tabulated. The number and percent of patients in each response category will be calculated. Patients who do not have any post-baseline tumor assessment will be counted under the category NE.

[redacted]

7.5.2.5 Clinical Benefit Rate

Clinical benefit rate outside CNS is defined as the proportion of patients having a confirmed CR, PR, or SD and for SD it has to be at least 4 months (≥ 120 days) per RECIST 1.1. The SD duration of 4 months is selected to reflect the shorter life expectancy of the study population given that the median OS for the TPC arm in patients with history of brain metastases was 4.8 months from the BEACON trial. CBR analyses will be performed for the ITT Population. CBR will be calculated and compared between the treatment arms using a Cochran Mantel-Haenszel (CMH) test stratified by the randomization factors. Clopper Pearson exact 2-sided 95% confidence limits will be calculated for CBR of each treatment arm.

In addition, the CBR in CNS will be defined as the proportion of patients having a SD or better and for SD it has to be at least 4 months (≥ 120 days) per RANO-BM. Clinical benefit rate in the brain will be compared between the treatment groups using a CMH test stratified by the randomization factors. Clopper Pearson exact two-sided 95% confidence limits will be calculated in determining the CBR of each group.

The primary analysis of CBR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

7.5.2.6 Duration of Response

Duration of response outside the CNS is defined for patients who have a confirmed CR or PR as the date from first documented CR or PR per RECIST 1.1 to the date of the documentation of disease progression or death due to any cause, whichever is earlier. The algorithm for calculation of date of progression and censoring rule (Table 1) will be the same as defined in the analysis for PFS. Median duration, its 95% confidence intervals as well as 25% and 75% quartiles of DoR will be summarized using Kaplan-Meier method for each treatment arm.

The primary analysis of DoR will be based on central imaging facility assessment of tumor response. As a secondary analysis, assessment of tumor response by investigator will be used.

7.5.3 HRQoL

The EORTC QLQ-C30 with BN-20, BFI and EQ-5D-5L will be used to measure the health outcome, QoL and assess the symptoms and side effects of treatment and their impact on everyday life.

The QLQ-C30 questionnaire is composed of 5 multi-item scales (physical, role, social, emotional, and cognitive functioning), a global health status/QoL scale, and 9 symptoms (pain, fatigue, financial impact, appetite loss, nausea/vomiting, diarrhea, constipation, sleep

disturbance, and dyspnea). Most items are scaled 1 to 4, except items contributing to the global health status/QoL, which are 7-point questions. Raw scores will be transformed using a linear transformation to standardize the results such that scores range from 0 to 100.

The calculation for scoring these scales is the same in all cases:

1. Calculate the average of the items that contribute to the scale; this is the raw score.
2. Use a linear transformation to standardize the raw score, so that scores range from 0 to 100; a higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms.

Calculations for raw score and linear transformation are as follows:

In practical terms, if items $I_1, I_2, \ldots I_n$ are included in a scale, where n is number of items in the scale, the procedure is as follows:

Raw score

$RS\ (Raw\ Score) = (I_1 + I_2 + I_3 + \ldots + I_n)/n$

Linear transformation

Apply the linear transformation to 0-100 to obtain the score S,

Functional scales: $S = \{1 - (RS - 1)/range\} \times 100$

Symptom scales / items: $S = \{(RS - 1)/range\} \times 100$

Global health status / QoL: $S = \{(RS - 1)/range\} \times 100$

Range is the difference between the maximum possible value of *RS* and the minimum possible value.

Raw score and linear transformed score will be calculated when at least 50% of the items from the scale have been answered. Otherwise, the score will be set as missing.

The structure of this questionnaire and scoring are presented in Table 4.

[redacted]

Table 4: Scoring for QLQ-C30 version 3.0

Scale	Scale abbreviation	Number of items	Individual score range	Item Number
Global health status / QoL				
Global health	QL	2	1-7	29, 30
Functional scales				
Physical functioning	PF	5	1-4	1 to 5
Role functioning	RF	2	1-4	6, 7
Emotional functioning	EF	4	1-4	21 to 24
Cognitive functioning	CF	2	1-4	20, 25
Social functioning	SF	2	1-4	26, 27
Symptom scales / items				
Fatigue	FA	3	1-4	10, 12, 18
Nausea and vomiting	NV	2	1-4	14, 15
Pain	PA	2	1-4	9, 19
Dyspnea	DY	1	1-4	8
Insomnia	SL	1	1-4	11
Appetite loss	AP	1	1-4	13
Constipation	CO	1	1-4	16
Diarrhea	DI	1	1-4	17
Financial difficulties	FI	1	1-4	28

At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group for each subscale. Changes from baseline will be compared between treatment groups using Repeated Measures Linear Mixed Effects Model (MMRM) with treatment, time, treatment and time interaction, baseline score, and stratification variables as covariates. Unstructured covariance matrix will be used unless the model fails to converge. Treatment effect on the change from baseline will be compared by the difference in the least-square means and corresponding 95% confidence interval.

Additional analysis, including generalized linear mixed effect model for sub-scales with small number of categories, may be explored.

Patients will be classified as improved, stable, or worsened. Minimal important difference (MID) thresholds, a threshold of 5 score points, will be used to categorize patients as improved ($\geq$ 5 decrease), stable (0 to <5 decrease), or worsened (increase) on the EORTC QLQ-C30 scores. The proportion of patients within each status group using the best change from baseline

[redacted]

will be calculated for each subscale. The proportion of patients improved, stable, or worsened will be calculated and compared between treatment groups using the CMH test stratified by the factor used for randomization.

Selected symptoms (e.g., neuropathy or diarrhea) will be identified by medical review. Time to progression of selected symptoms will be calculated and compared between treatment groups. Progression is defined as an increase in at least 1 point from baseline. Time to progression of a selected symptom will be calculated between the date of randomization and date of first observation of the progression. Patients who do not experience progression of a symptom will be censored at the date of their last EORTC QLQ assessment or at their date of randomization (if no post-baseline EORTC QLQ assessment is completed). A KM method will be used to estimate the median times and their 95% confidence interval for each treatment group. A two-sided log-rank test will be performed to compare treatment groups for time to progression in each selected symptom. If the proportion of patients with progression of selected symptoms is below 30%, descriptive statistics will be used to provide summary time to progression.

The brain cancer module (QLQ-BN20) is intended for patients undergoing chemotherapy or radiotherapy. It includes 20 items (individual score range 1-4) that aggregates into 4 multi-item scales of future uncertainty (4 items), visual disorder (3 items), motor dysfunction (3 items), communication deficits (3 items); and 7 single-item scales of headaches, seizures, drowsiness, itchy skin, hair loss, weakness of legs, and bladder control. All scale scores and items are linearly transformed to a 0–100 scale with higher scores indicating more severe symptoms.

At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group for each subscale. In addition, change from baseline score in QLQ-BN20 will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

The BFI was developed for the assessment of the severity of fatigue and the impact of fatigue on daily functioning. It consists of 9 numeric rating scales of 0 to 10: 3 items that measure severity and 6 items that measure interference with daily activities. A scale score will be calculated if a minimum of 5 items have been answered. A global fatigue score will be obtained by averaging the answered items on the BFI. At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group. In addition, change from baseline score in BFI will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

The EQ-5D-5L consists of descriptive system and the EQ visual analogue scale (EQ VAS). The EQ-5D-5L descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 5 levels: no problems,

[redacted]

slight problems, moderate problems, severe problems, and extreme/unable to perform activity. The responses from each of 5 dimensions will be combined in a 5-digit number describing the respondent's health state. It should be noted that the numerals 1-5 have no arithmetic properties and should not be used as a cardinal score. At each assessment point, the EQ-5D-5L levels will be divided into 'no/slight problem', 'moderate problem' or 'severe/extreme problems'. The proportion of patients within each status will be calculated and compared between treatment groups using the CMH test stratified by the factors used in randomization.

The EQ VAS records the patient's self-rated health on a vertical, visual analogue scale where the endpoints are labeled 'Best imaginable health state' and 'Worst imaginable health state'. This information can be used as a quantitative measure of health outcome as judged by the individual patients. At each assessment point, summary statistics of absolute scores and changes from baseline will be calculated by treatment group. In addition, change from baseline score will be compared using the same MMRM model similarly to the analysis of QLQ-C30.

7.5.4 Graded Prognostic Assessment Index

The GPA tool, which assigns scores for prognostic indices using three prognostic factors for survival (Karnofsky performance status (KPS), tumor subtype, and age) has been developed to predict estimated survival in patients with newly diagnosed brain metastases.

The GPA Index will be calculated as the sum of the subscores from the three factors according to Table 5.

Table 5: GPA Calculation

Factor\ Score	0	0.5	1.0	1.5	2.0
KPS	Less than 50	50-60	70-80	90-100	n/a
Tumor Subtype	Basal	n/a	Luminal A	Her2+	Luminal B
Age	60 or greater	Less than 60	n/a	n/a	n/a

Abbreviations: KPS = Karnofsky Performance Status; n/a = not applicable

The following tables outline the mapping from ECOG performance status score to KPS score and from receptor status to tumor subtype:

[redacted]

Table 6: ECOG Performance Status versus Karnofsky Performance Status

ECOG Performance Status	Karnofsky Performance Status
0—Fully active, able to carry on all pre-disease performance without restriction	100—Normal, no complaints; no evidence of disease 90—Able to carry on normal activity; minor signs or symptoms of disease
1—Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work	80—Normal activity with effort, some signs or symptoms of disease 70—Cares for self but unable to carry on normal activity or to do active work
2—Ambulatory and capable of all self-care but unable to carry out any work activities; up and about more than 50% of waking hours	60—Requires occasional assistance but is able to care for most of personal needs 50—Requires considerable assistance and frequent medical care
3—Capable of only limited self-care; confined to bed or chair more than 50% of waking hours	40—Disabled; requires special care and assistance 30—Severely disabled; hospitalization is indicated although death not imminent
4—Completely disabled; cannot carry on any self-care; totally confined to bed or chair	20—Very ill; hospitalization and active supportive care necessary 10—Moribund
5—Dead	0—Dead
Tumor Subtype	**Receptor Status**
Her2+	ER-, PR- and Her2 +
Basal	ER-, PR- and Her2- (TNBC)
Luminal A	ER+ and/or PR+ and HER2-
Luminal B	ER+ and/or PR+ and HER2+

The GPA Index range from 0-4 with a lower score predictive of a worse prognosis for survival following a diagnosis of brain metastasis. The association between GPA Index and survival will be evaluated using a Cox regression model. GPA-adjusted hazard ratio will be presented if the association is significant.

7.5.5 Magnitude of Clinical Benefit

The magnitude of clinical benefit of NKTR-102 will be assessed by the ESMO magnitude of clinical benefit scale (ESMO-MCBS v1.0). The ESMO-MCBS can be applied to comparative outcome studies evaluating the relative benefit of treatments using outcomes of survival, QoL, surrogate outcomes for survival or treatment toxicity in solid cancers. For cancer therapies, the ESMO-MCBS scale provides a clear, well-structured and validated mechanism to indicate the

[redacted]

magnitude of benefit in addition to the level of evidence that can inform both national and international (e.g. ESMO) guidelines. Table 7 (downloaded from http://www.esmo.org/Policy/Magnitude-of-Clinical-Benefit-Scale/Scale-Evaluation-Forms) will be used to calculate the magnitude of clinical benefits.

Table 7: ESMO Magnitude of Clinical Benefit Scale – Form 2a: for Therapies That are not Likely to be Curative with Primary Endpoint of OS

	Mark with X if relevant
Grade 4	
HR ≤ 0.65 AND Gain ≥ 3 months	
Increase in 2 year survival alone ≥ 10%	
Grade 3	
HR ≤ 0.65 AND Gain 2.5-2.9 months	
Increase in 2 year survival alone 5 - <10%	
Grade 2	
HR > 0.65-0.70 OR Gain 1.5-2.4 months	
Increase in 2 year survival alone 3 - <5%	
Grade 1	
HR > 0.70 OR Gain <1.5 months	
Increase in 2 year survival alone <3%	

Abbreviations: HR = hazard ratio; OS = overall survival

Preliminary magnitude of clinical benefit grade (highest grade scored)

4	3	2	1

Quality of Life assessment /grade 3-4 toxicities assessment*

Does secondary endpoint quality of life show improvement?

Are there statistically significantly less grade 3-4 toxicities impacting on daily well-being*	

*This does not include alopecia, myelosuppression, but rather chronic nausea, diarrhea, fatigue, etc.

[redacted]

Adjustments

Upgrade 1 level if improved quality of life and/or less grade 3-4 toxicities impacting daily well-being are shown

Final adjusted magnitude of clinical benefit grade

5	4	3	2	1

7.5.6 Health Economics

A separate analysis plan for health economics will be provided.

7.6 Pharmacokinetics and Biomarkers

A separate analysis plan for PK and biomarker will be provided.

[redacted]

8.0 SAFETY ANALYSIS

The safety data will include AEs, SAEs, ECOG performance status, and clinical laboratory tests. Summaries will use the Safety population and will be presented separately for NKTR-102 and TPC treated patients. All safety data will be presented in data listings.

8.1 Adverse Events and Death

Adverse events will be coded by SOC and preferred term using MedDRA. Adverse event severity will be based on NCI CTCAE Grade (version 4.03).

A treatment emergent adverse event (TEAE) is defined as an AE that was not present prior to treatment with study drug but appeared following treatment, or was present prior to treatment initiation but worsened during treatment. An AE that was present at treatment initiation but resolved and then reappeared while the patient was on treatment is a TEAE (regardless of the intensity of the AE when the treatment was initiated). The treatment-emergent period will be defined as the period of time from the date of the first dose of study drug up to 30 days after the date of the last dose of study drug or the day prior to the initiation of subsequent anticancer treatment, whichever occurs first.

Adverse events leading to permanent discontinuation of study drug and/or leading to death will be listed and tabulated. The frequency of TEAEs and SAEs will be tabulated by preferred term and system organ class. The maximum NCI CTCAE grade and frequency of AEs will be summarized. For patient level summaries (except for time to onset and time to resolution of diarrhea), patients with multiple occurrences of events of the same preferred terms and SOC will be counted once at the highest Grade and the strongest relationship to study drug for each preferred term, system organ class. Adverse events that are reported as possibly, probably, or definitely related to study drug will be counted as related to study drug; AEs with a missing relationship will be considered "Related" for this summary.

The following adverse events summaries will be provided:

- TEAEs by SOC and Preferred Term
- TEAEs by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs by SOC and Preferred Term and CTCAE Grade
- Grade 3 or Higher TEAEs by Preferred Term and by Descending Incidence of Overall Patients

[redacted]

- Serious TEAEs by Preferred Term and by Descending Incidence of Overall Patients
- Serious TEAEs by SOC and Preferred Term and CTCAE Grade
- TEAEs Related to Study Drug by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs Related to Study Drug by SOC and Preferred Term and CTCAE Grade
- Grade 3 or Higher TEAEs Related to Study Drug by SOC and Preferred Term and CTCAE Grade
- Serious TEAEs Related to Study Drug by Preferred Term and by Descending Incidence of Overall Patients
- TEAEs Leading to Study Drug Discontinuation by SOC and Preferred Term and CTCAE Grade
- TEAEs Leading to Dose delay, Reduction, and Interruption by SOC and Preferred Term and CTCAE Grade
- TEAEs with Fatal Outcome by SOC and Preferred Term
- TEAEs Related to Study Drug with Fatal Outcome by SOC and Preferred Term

The following data listings will be produced:

- All AEs
- SAEs
- TEAEs Leading to Study Drug Discontinuation
- TEAE with Fatal Outcome

8.2 Clinical Laboratory Evaluations

Clinical laboratory tests will be conducted according to the Schedule of Assessments. Clinical laboratory tests will be performed by a designated central laboratory. Central laboratory results must be used to determine patient eligibility. For clinical management, local laboratories may be used for each site as deemed necessary by the treating physician. If a local laboratory result is considered a clinically significant AE, it should be reported as an AE.

For hematology and chemistry, the absolute values and change from baseline at each post-baseline study assessment will be summarized by descriptive statistics. In addition, the minimum and maximum absolute values and corresponding change from baseline across all post-baseline study assessments will be provided.

For selected laboratory tests with CTCAE grade, shift tables comparing post-baseline to baseline toxicity grade will be provided at each post-baseline study assessment. In addition, the worst shift from baseline will be summarized.

All central laboratory test results will be presented in data listings.

8.3 Additional Safety Analyses

Additional safety analysis will be performed to understand selected toxicities including diarrhea, neutropenia, neuropathy, and hypersensitivity reactions.

The total number of diarrhea events will be summarized. The number and percent of patients with at least one diarrhea event will be summarized by toxicity grade. A patient is only counted once using the highest grade. The number and percent of patients who had serious diarrhea events will be summarized.

Time to onset of first diarrhea event will be defined as the time between date of first dose and date of first diarrhea event and summarized using descriptive statistics. Time to onset of first grade 2 or higher, first grade 3 or higher diarrhea events will be defined and summarized similarly. The relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher diarrhea events will be plotted. Median duration of diarrhea events of any severity grade will be calculated for each patient. The median duration of diarrhea events for each treatment arm will be summarized by descriptive statistics using the median duration of each patient in that arm. Duration of grade 2 or higher diarrhea event, grade 3 or higher diarrhea event will be calculated and summarized similarly.

The number and percent of patients who discontinued study drug due to diarrhea with and without resolution will be summarized. For these patients, time to resolution will be calculated as the stop date of the event minus the date of last dose plus one and summarized using descriptive statistics.

In addition, selected analysis for diarrhea may be repeated for patients who had prophylactic loperamide versus loperamide treatment only.

The total number of neutropenia events will be summarized. The number and percent of patients with at least one neutropenia event will be summarized by toxicity grade. A patient is only

[redacted]

counted once using the highest grade. The number and percent of patients who had serious neutropenia events will be summarized.

Time to onset of first neutropenia event will be defined as the time between date of first dose and date of first neutropenia event and summarized using descriptive statistics. Time to onset of first grade 2 or higher, first grade 3 or higher neutropenia events will be defined and summarized similarly. The relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher neutropenia events will be plotted. Median duration of neutropenia events of any severity grade will be calculated for each patient. The median duration of neutropenia events for each treatment arm will be summarized by descriptive statistics using the median duration of each patient in that arm. Duration of grade 2 or higher neutropenia event, grade 3 or higher neutropenia event will be calculated and summarized similarly.

Neuropathy-related events will be identified using preferred term by medical review. The total number of neuropathy-related events will be summarized. The number and percent of patients with at least one neuropathy events will be summarized by toxicity grade. A patient is only counted once using the highest grade. The number and percent of patients for each neuropathy-related event will be calculated under each preferred term by toxicity grade. Time to onset of neuropathy-related events, relationship between cumulative dose of NKTR-102 to time to onset of grade 3 or higher neuropathy-related events, and duration of neuropathy-related events will be summarized similar to diarrhea-related events. The number and percent of patients with dose reduction and delay due to neuropathy-related events will be summarized. In addition, the number and percent of patient who discontinued study treatment due to neuropathy-related events with and without resolution will be summarized.

Neutropenia-related events will be identified, summarized and analyzed similarly to the neuropathy-related events.

Hypersensitivity events will be identified under selected SOC and preferred term by medical review. The total number of hypersensitivity events and the number and percent of patients who had at least one hypersensitivity events will be summarized . In addition, the number and percent of patients who had hypersensitivity events will also be summarized for each SOC, preferred term and severity. For patients with multiple events under a SOC or preferred term, patients will be counted only once using the highest severity. The duration of hypersensitivity events will be summarized by descriptive statistics. For patients who had more than one hypersensitivity events, the median duration of all events within the patient will be used for calculating the descriptive statistics. An analysis of hypersensitivity events and on-study development of polyethylene glycol (PEG) antibodies will be undertaken. This analysis will occur separately from the main analysis of the trial. The proportion of patients with PEG antibodies will be calculated. The timing of appearance of PEG antibodies will be examined.

[redacted]

Pulmonary toxicity events (potentially indicative of pneumonitis or interstitial lung disease) will be identified by medical review. The number and percent of patients who have pulmonary toxicity will be summarized for each treatment arm.

UGT1A1 polymorphisms may provide a useful diagnostic tool to predict *in vivo* glucuronidation of 7-ethyl-10-hydroxy-camptothecin (SN38), a molecule that is a major determinant of irinotecan and NKTR-102 metabolism and toxicity. All patients in this study treated with NKTR-102 will be tested for UGT1A1 alleles. The incidence and severity of diarrhea and neutropenia will be summarized for patients randomized to NKTR-102 and with different UGT1A1 alleles (i.e., Poor, Intermediate, and Normal). These data will also be part of a larger population analysis across NKTR-102 clinical trials.

[redacted]

9.0 INTERIM ANALYSIS

A DMC Charter will be approved and finalized by the independent DMC members prior to the initiation of any interim analysis; the DMC Charter and meeting minutes will be submitted as part of the final Clinical Study Report.

Two-sided alpha of 0.001 will be used to test efficacy in OS at this interim analysis as the efficacy zone of the adaptive design. If OS does not reach statistical significance, the number of events for final OS analysis is determined per Mehta & Pocock (2011) to maintain approximately 80% conditional power for the CHW test. In order to prevent back calculation of the interim treatment effect as pointed out in (Liu and Hu 2016), the actual adaptation rule will take the step-function form and maintain approximate 80% conditional power. The adaptation rule – Event Size Adaption Rule for Clinical Study Protocol 15-102-14 will be included in the appendix of DMC charter and can only be accessed by DMC members, the Sponsor Executive Committee members and the sponsor design statistician who are not involved in the study conduct.

[redacted]

10.0 DATA MONITORING COMMITTEE

An independent DMC will be established to review interim efficacy and safety analyses results when approximately 82 OS events have occurred. The DMC consists of two external physicians and one external biostatistician.

The role of the DMC is to review the efficacy and safety interim analysis following the rules provided in the DMC charter and charter appendix.

A DMC charter will be approved and finalized by the DMC members prior to the initiation of any interim analysis. The DMC charter describes details including the primary responsibilities of the DMC, its relationship with other trial components, its membership, and the purpose and timing of its meetings, statistical monitoring guidelines to be implemented and statistical analysis for the open and closed sessions. The DMC Charter and meeting minutes will be submitted as part of the final Clinical Study Report.

11.0 REFERENCES

1. Anders CK, Deal AM, Miller R, et al. The Prognostic Contribution of Clinical Breast Cancer Subtype, Age, and Race Among Patients With Breast Cancer Brain Metastases. Cancer. 2010;117(8):1602-1611.

2. Brown PD, Asher AL, Ballman KV, et al. NCCTG N0574 (Alliance): A phase III randomized trial of whole brain radiation therapy (WBRT) in addition to radiosurgery (SRS) in patients with 1 to 3 brain metastases. Neuro-Oncology. 2015;17(suppl5):v45.5-v46

3. Crozier JA, Cornell LF, Rawal B, et al. Breast cancer brain metastases: Comprehensive review of tumor extent, histology and treatment on survival. Journal of Clinical Oncology. 2015;33 (supple; abstract 1059).

4. Cui L, Hung JHM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics 1999; 55: 853-857.

5. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-247.

6. Halasz LM, Uno H, Hughes M, et al. Comparative Effectiveness of Stereotactic Radiosurgery Versus Whole-Brain Radiation Therapy for Patients With Brain Metastases From Breast or Non–Small Cell Lung Cancer. Cancer. 2016.

7. Khuntia D. Contemporary review of the management of brain metastasis with radiation. Hindawi Publishing Corporation. Advances in Neuroscience. 2015;2015:1-3.

8. Liu Y, and Hu M. Testing multiple primary endpoints in clinical trials with sample size adaptation. Pharmaceut. Statist. 2016, 15: 37–45. doi: 10.1002/pst.1724.

9. Mehta C, Pocock SJ. Adaptive increase in sample size when interim results are promising: A practical guide with examples. Statistics in Medicine 2011; 30: 3267-3284.

10. Mosteller RD. "Simplified calculation of body-surface area". N Engl J Med 1987; 317:1098. PMID 3657876.

11. Niwinska A, Murawska M, Pogoda K. Breast cancer brain metastases: differences in survival depending on biological subtype, RPA RTOG prognostic class and systemic treatment after whole-brain radiotherapy (WBRT). Annals of oncology. 2010;21(5):942-948.

[redacted]

12. Paul Brown et al. A phase III randomized trial of whole brain radiation therapy (WBRT) in addition to radiosurgery (SRS) in patients with 1 to 3 brain metastases . 2015 ASCO Annual Meeting.

13. Stone A.M. et al. Research outcomes and recommendations for the assessments of progression in cancer clinical trials from a PhRMA working group. 2011; 47: 1763-1771.

14. Subbiah IM, Lei X, Weinberg JS, et al. Validation and development of a modified breast graded prognostic assessment as a tool for survival in patients with breast cancer and brain metastases. Journal of Clinical Oncology. 2015;33(20):2239-2245.

15. Wassmer G. Planning and analyzing adaptive group sequential survival trials. Biometrical Journal. 2006; 48: 714-729.

[redacted]

12.0 APPENDIX

Let T_1 denote the interim logrank test statistic with d_1 number of events and T_2 denote the logrank test statistic based on the cumulative data at final analysis with adapted d_2^* number of events. Then we calculate the incremental test statistic as following:

$$\tilde{T}_2^* = \frac{\sqrt{d_2^*}T_2 - \sqrt{d_1}T_1}{\sqrt{d_2^* - d_1}}.$$

The final CHW test statistic with weight $\sqrt{\frac{2}{3}}$ and $\sqrt{\frac{1}{3}}$ can be written as $T_{2,CHW}^* = \sqrt{\frac{2}{3}}T_1 + \sqrt{\frac{1}{3}}\tilde{T}_2^*$.

The overall Type I error rate is the probability of rejecting OS at either interim or final analysis under the null that there is no difference in survival between the treatment and control group (denoted as H_0). The critical value for the interim analysis is denoted as $z_{\frac{0.001}{2}} = \Phi^{-1}(1 - \frac{0.001}{2})$ (where Φ denotes the CDF for standard normal distribution) as two-sided 0.001 level of alpha is spent at this time. The critical value c for the final analysis is calculated so that the overall type I error rate is maintained at 0.05 (two-sided) or 0.025 (one-sided). The probability can be calculated as follows:

$$0.025 = P_{H_0}(T_1 > z_{0.001/2} \text{ or } T_{2,CHW}^* > c) = P_{H_0}(T_1 > z_{0.001/2}) + P_{H_0}(T_1 \le z_{0.001/2} \text{ and } T_{2,CHW}^* > c)$$

Therefore, we need to find c such that $P_{H_0}\left(T_1 \le z_{\frac{0.001}{2}} \text{ and } T_{2,CHW}^* > c\right) = 0.025$. Because T_1 and $T_{2,CHW}^*$ asymptotically follows standard normal distribution with correlation $\sqrt{\frac{2}{3}}$ (**Wassmer 2006**), the critical value c can be found to be $z_{0.0249698}$. This implies that the significant p-value will be 0.0249698 (one-sided) or approximately 0.0499 (two-sided).

[redacted]

Population Pharmacokinetics and Exposure-Response Analysis Plan

TITLE:	A Phase 3 Open-Label, Randomized, Multicenter Study of NKTR-102 versus Treatment of Physician's Choice (TPC) in Patients with Metastatic Breast Cancer Who Have Stable Brain Metastases and Have Been Previously Treated with an Anthracycline, a Taxane, and Capecitabine (Protocol 15-102-14)
REPORT NUMBER:	LS-2019-509
VERSION:	Version 1
STUDY SPONSOR:	NEKTAR Therapeutics Clinical Pharmacology 455 Mission Bay Boulevard South San Francisco, CA 94158
DATE APPROVED:	27 January 2020

Report No. LS-2019-509 (ATTAIN PKAP) **NKTR-102**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Abbreviation or Term	Definition/Explanation
AUC	Area under the Concentration-time Curve
AUCss	Area under the Concentration-time Curve at Steady State
APC	7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino] carbonyloxycamptothecin
BSA	Body Surface Area
C_{max}	Maximum Observed Plasma Concentration
CSR	Clinical Study Report
eGFR	Estimated Glomerular Filtration Rate
EMA	European Medicines Agency
E-R	Exposure-Response
FDA	Food and Drug Administration
IV	Intravenous
ORR	Objective Response Rate
OS	Overall Survival
PD	Pharmacodynamic, Pharmacodynamics
PFS	Progression Free Survival
PK	Pharmacokinetic, Pharmacokinetics
PKAP	Population Pharmacokinetic Analysis Plan
PPK	Population Pharmacokinetic
$T_{1/2}$	Half-life
T_{max}	Time to Maximum Observed Plasma Concentration
SN38	7-ethyl-10-hydroxy-camptothecin; the active metabolite of irinotecan
SN38-G	SN38-glucuronide
UGT1A1	Uridine Diphosphate-glucoronosynl Transferase 1A1
UGT1A1*28	A Genetic Variant of the Drug Metabolizing Enzyme UDP-glucuronosyltransferase 1A1
V_d	Volume of Distribution
V_{ss}	Estimated Volume of Distribution at Steady State

1.0 INTRODUCTION

This document describes the planned population pharmacokinetics (PPK) and exposure-response (E-R) analyses of data captured according to Nektar Therapeutics protocol 15-102-14 (ATTAIN), "A phase 3 open-label, randomized, multicenter study of NKTR-102 vs. treatment of physician's choice (TPC) in patients with metastatic breast cancer who have stable brain metastases and have been previously treated with an anthracycline, a taxane, and capecitabine" (Protocol 12-102-14, Protocol Amendment 4.0). The pharmacokinetic analysis plan (PKAP) was developed based on protocol amendment 4.0, dated 16 April 2018. It extends the descriptions of pharmacokinetic (PK) and
E-R analyses provided in the protocol.

The PPK analysis will characterize the PK of NKTR-102 and several metabolites in patients with metastatic breast cancer who have stable brain metastases (BCBM patients) and received NKTR-102 in ATTAIN. NKTR-102 consists of a 4-arm polyethylene glycol (PEG) polymer with a nominal molecular weight of 20 kDa, a hydrolysable ester-based linker, and one irinotecan molecule at the end of each arm. Upon administration, the linker slowly hydrolyzes, resulting in sustained exposure to irinotecan that is subsequently metabolized to multiple products, including the inactive metabolite 7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino] carbonyloxycamptothecin (APC) and the active metabolite 7-ethyl-10-hydroxy-camptothecin (SN38). SN38 is subject to glucuronidation by UGT enzymes, and published data have shown elevated exposure to SN38 in subjects who are homozygous for the UGT1A1*28 allele (Innocenti, 2014).

The PK and E-R analyses are intended to support global regulatory filings and were developed with consideration of the FDA Guidance for Industry entitled "Population Pharmacokinetics" (July 19, 2019) and "Exposure-Response Relationships - Study Design, Data Analysis, and Regulatory Applications" (Feb 2003) and the EMA "Guideline on Reporting the Results of Population Pharmacokinetic Analyses" (Doc. Ref. CHMP/EWP/185990/06).

2.0 OBJECTIVES OF THE POPULATION PHARMACOKINETICS AND EXPOSURE-RESPONSE ANALYSES

The objectives of the PPK analysis are:

- To use a PPK model to predict PK parameter values for patients in ATTAIN, including individual subject SN38 exposure values (AUC_{ss}) for use in UGT1A1 genotype evaluations and an E-R analysis
- To compare SN38 AUC_{ss} in subjects who are homozygous for UGT1A1*28 versus subjects with all other UGT1A1 genotypes

The E-R analyses will use SN38 AUC_{ss} as the exposure variable (divided into exposure quartiles) for the following objectives:

- To evaluate potential relationships between SN38 exposure and the efficacy endpoint of overall survival (OS)
- If the median of the 4th quartile of SN38 AUCss is >5-fold higher the median of the 1st quartile, then to evaluate potential relationships between SN38 exposure and safety parameters related to diarrhea (i.e., incidence of diarrhea of all grades, ≥ Grade 2, and ≥ Grade 3)

3.0 DATA

3.1 Population Pharmacokinetic Analysis Population and Data Set

The PPK analysis population is defined as all patients who received at least one dose of NKTR-102 in ATTAIN and have at least one measurable post-dose plasma concentration of NKTR-102, irinotecan, or SN38.

The PPK analysis dataset will consist of plasma concentration-time data from all subjects in the analysis population. The data files will be assembled with SAS® v.9.4 (SAS Institute, Cary, NC) and prior to use in the PPK analysis will undergo formal QC review by a programmer or analyst other than the data programmer.

Study design features are summarized in Table 1.

Table 1: Summary of Design Features of NKTR-102 Study 15-102-14 (ATTAIN)

Phase	3
Indication	Metastatic breast cancer with stable brain metastases
Treatment arms:	NKTR-102, Treatment of physician's choice (TPC)
Number of patients	220; 1:1 randomization NKTR-102 vs TPC
NKTR-102 Administration	145 mg/m^2, 90 min IV infusion administered every 21 days
PK collection times	A single blood sample prior to each drug administration, starting from the first dose

3.2 Missing Data and Imputations

Actual dates and times of PK samples and doses will be used whenever possible. If the start or the end of infusion time is missing, then the start or end time will be imputed based on either adding or subtracting the 90-minute infusion duration. If actual dates and times for PK samples are missing, data from the associated PK samples will be excluded from the analysis. If the actual dosing date and time are missing for a NKTR-102 dose, then the dosing date and time will be set to those recorded for the pre-dose PK sample collected just before the start of the cycle.

The concentrations below the limit of quantification (BLQ) will be excluded from the analysis. If the number of BLQ samples is greater than 70% for a given analyte, then the analyte will be excluded from the analysis.

3.3 Identification of Outliers

No attempts will be made to identify and exclude outlier data points.

4.0 METHODOLOGY

4.1 Prior Knowledge and Modeling Approach

A 5-analyte PPK model for NKTR-102 and its metabolites (irinotecan, SN38, SN38G, and APC) has been developed (LS-2015-504) [5]. PPK analysis was carried out in two stages. In Stage 1, model parameters, including covariate effects for NKTR-102, irinotecan, and SN38 were estimated in a 3-analyte model (Figure 1). This model was then expanded in Stage 2 to include concentration-time data for SN38G and APC, using fixed Stage 1 model parameter values (Figure 2).

Figure 1: Parameterization of structural PK model for the 3-Analyte model, consisting of NKTR-102, Irinotecan and SN38 (LS-215-504)

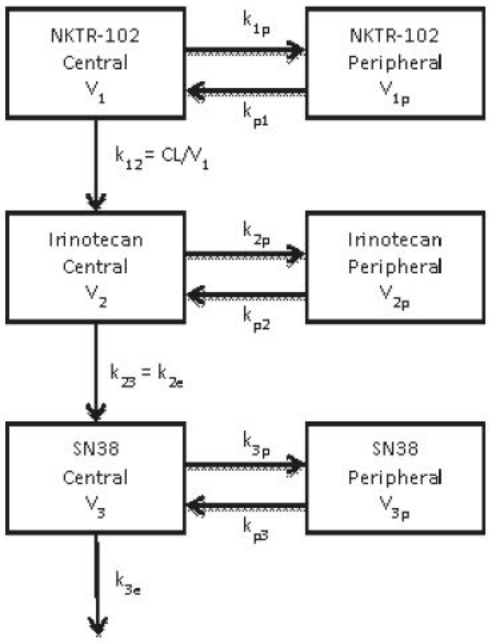
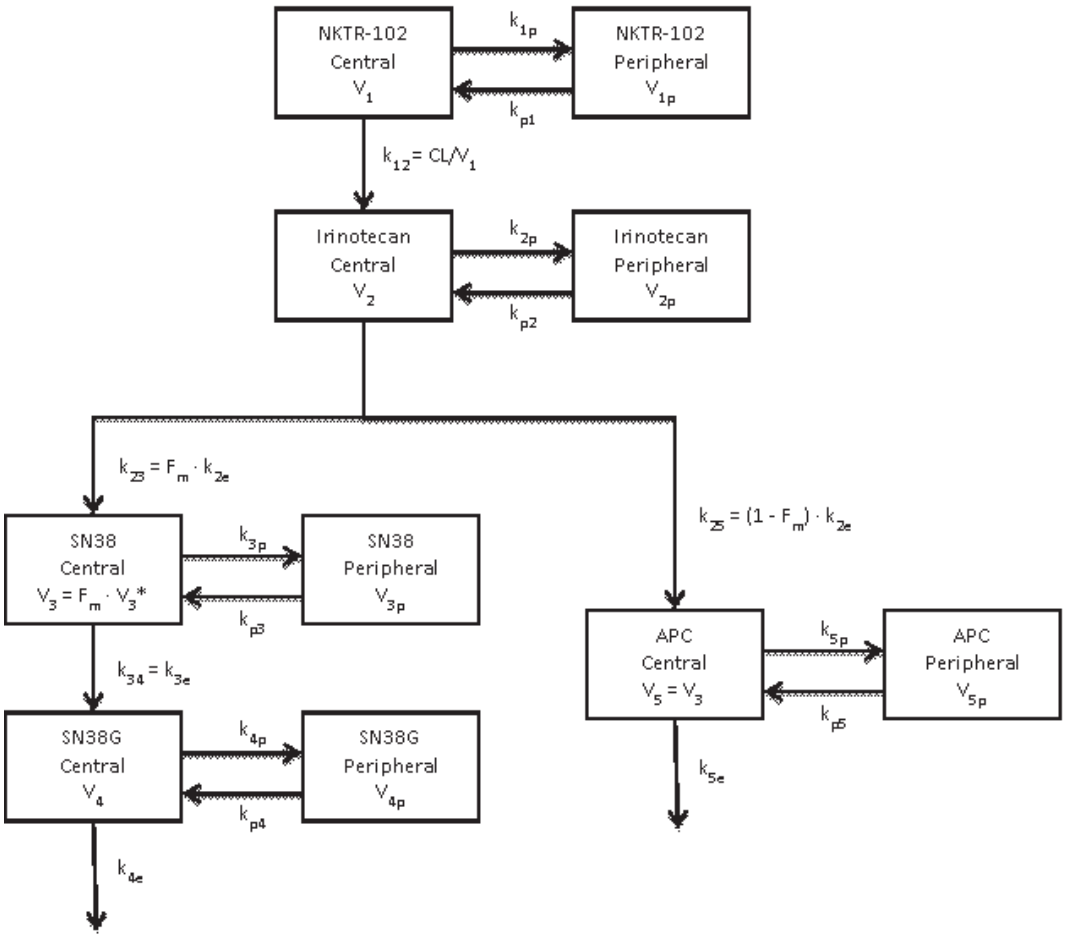


Figure 2: **Parameterization of structural PK model for the 5-Analyte model consisting of NKTR-102, Irinotecan, SN38, SN38G and APC (LS-2015-504)**



The model includes body surface area (BSA) and estimated glomerular filtration rate (eGFR) as covariates of NKTR-102 clearance, BSA and gender as covariates of NKTR-102 central volume of distribution, and UGT1A1 status (homozygous for UGT1A1*28 vs all other genotypes) as a covariate of SN38G central volume of distribution. A summary of the final PPK model parameters is provided in Appendix 2.

4.2 Covariate Analysis

The previously developed covariate model (LS-2015-504) will be used. No new covariate analysis is planned.

4.3 Exploratory Graphical Analysis

An exploratory graphical analysis will be used to assess whether the existing PPK model (LS-2015-504) adequately describes the PK data from ATTAIN. The existing PPK model will be used to generate 90% prediction intervals for population predicted concentrations of NKTR-102, irinotecan, SN38, SN38G, and APC. The observed concentrations for each of these analytes will then be overlaid onto the 90% concentration prediction intervals.

If most of the observed data for NKTR-102, irinotecan, SN38, SN38G, or APC fall within the 90% concentration prediction intervals, then the existing PPK model will be considered fit for purpose for the analyses described in the PKAP for that analyte. If this condition is not met for an individual analyte, observed trough concentration levels by cycle will be summarized using descriptive statistics.

4.4 Estimation of Pharmacokinetic Parameters for ATTAIN Patients

ATTAIN protocol Section 10.8 states the following: “Concentration-time data from this study will be pooled with data from other clinical studies prior to analysis with nonlinear-mixed effects modeling. Pharmacokinetic parameters such as maximum concentration (Cmax), time to maximum concentration (Tmax), area under the concentration-time curve (AUC), clearance (CL), volume of distribution (V), and elimination half-life (T½) will be tabulated and summarized with descriptive statistics.” If there is acceptable agreement between the predicted and observed concentrations values in ATTAIN (Section 5.3), then the existing PPK model will be used to generate posthoc estimates of these PK parameters for subjects in ATTAIN.

To estimate pharmacokinetic parameters, concentration-time data from ATTAIN (mostly consisting of trough concentrations) will be fit with the existing PPK model, allowing adjustment of PPK model parameters within the previously observed variability. A 2-step approach will be followed: NKTR-102, irinotecan, and SN38 concentration will be fit first, followed by addition of SN38G and APC. The individual PPK model parameters will be used to estimate individual patient and summary statistics for Tmax, Cmax, and AUCss for NKTR-102, irinotecan, SN38, SN38G, and APC, and also V CL, and T½ for NKTR-102. T½ will be calculated as previously described using below formula:

$$t_{\frac{1}{2}} = \frac{\ln(2)}{\min(\alpha, \beta)},$$

where $\beta = 1/2(k_{12} + k_{21} + k_{10} + \sqrt{(k_{12} + k_{21} + k_{10})^2 - 4k_{21}k_{10}})$, $\alpha = (k_{21}k_{10})/\beta$ and $k_{10} = CL/V_1$.

4.5 Model Qualification

Formal qualification of the PPK model is not planned. Rather, the exploratory analysis described in Section 4.3 will be performed to evaluate whether an existing PPK (LS-2015-504) is fit for purpose.

4.6 SN38 Exposure Comparison Based on UGT1A1*28 Genotype

Boxplots of SN38 AUC_{ss} values will be used to compare exposure in subjects who were homozygous for the UGT1A1*28 versus subjects with all other UGT1A1 genotypes.

4.7 Exposure-Response Analysis

The previous E-R models based on BEACON data (LS-2015-508) will be used as a framework to investigate E-R relationships in ATTAIN patients. SN38 AUC_{ss} will be the exposure variable in all E-R analyses. SN38 AUC_{ss} values for subjects in the PPK analysis population will be classified into quartiles that divide the data into four approximately equal groups after sorting by rank order. For the E-R efficacy analysis, subjects randomized to the comparator treatment (TPC) will be assigned to a fifth category.

4.7.1 Exposure-Efficacy Analysis

ORR, PFS, and OS will be evaluated using the methods described in the Statistical Analysis Plan (SAP) for ATTAIN. A Kaplan-Meier plot of ORR, PFS, and OS stratified by the exposure quartiles described in Section 4.7 will be generated to assess the relationship between SN38 exposure and efficacy.

4.7.2 Exposure-Safety Analysis

The relationship between SN38 exposure and incidence of diarrhea will be investigated only if the median of the 4th quartile of SN38 AUC_{ss} is >5-fold higher the median of the 1st quartile. The incidence of all grades, ≥Grade 2 and ≥Grade 3 diarrhea will be estimated as described in the Statistical Analysis Plan (SAP) for ATTAIN and will be computed for each SN38 exposure quartile. If the incidence of all grade diarrhea is at least 10% and exhibits a trend across SN38 AUCss quartiles, then significance will be assessed using logistic regression. The presence of prophylactic anti-diarrhea medications will be included as a covariate in the logistic regression model.

5.0 REPORTING

A description of this analysis will be provided in a report according to "Guidance for Industry: Population Pharmacokinetics" issued by the FDA in 1999 and "Guidance on Reporting the Results of Population Pharmacokinetic Analyses" issued by the EMEA in 2007.

The major sections of the report will include a description of the studies/subjects employed in the analysis, data assembly, analysis methods, results, discussions, and conclusions.

Key modeling and output files will be included in the report appendices. Other files together with data sets will be included as electronic appendices. The succeeding subsections provide a minimal list of the in-text tables and figures that will be included in the report, as well as the planned appendices. Additional in-text tables, figures, and appendices may be added to the report as deemed appropriate.

Summary statistics of the estimated NCA parameters will be presented. The descriptive statistics will include the number of subjects, mean (±standard deviation), CV%, median, minimum, and maximum values.

In addition to the report, an Analysis Data Module (ADaM) will be produced, along with Study Data Tabulation Module (SDTM) and define files as appropriate and in the spirit of Clinical Data Interchange Standards Consortium (CDISC) standards.

6.0 REFERENCES

1. Protocol 12-102-14, Protocol Amendment 4.0: A phase 3 open-label, randomized, multicenter study of NKTR-102 vs. treatment of physician's choice (TPC) in patients with metastatic breast cancer who have stable brain metastases and have been previously treated with an anthracycline, a taxane, and capecitabine.
2. Innocenti F, Schilsky RL, Ramirez J, Janisch L, Undevia S, House LK, Das S, Wu K, Turcich M, Marsh R, Karrison T, Maitland ML, Salgia R, Ratain MJ. Dose-finding and pharmacokinetic study to optimize the dosing of irinotecan according to the UGT1A1 genotype of patients with cancer. *J Clin Oncol* 2014, 32(22)**:** 2328-2334
3. Population Pharmacokinetics Guidance for Industry, U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default htm
4. Reporting the results of population pharmacokinetic analyses, European Medicines Agency, https://www.ema.europa.eu/en/reporting-results-population-pharmacokinetic-analyses
5. LS-2015-504: Application and Qualification of a Population Pharmacokinetic Model for NKTR-102 (etirinotecan pegol) Monotherapy in Patients with Advanced Solid Tumors.
6. SAP for protocol 15-102-14, Amendment 4. SAP Amendment #2 (version 2), 17Dec18.

APPENDIX 1: PLANNED LISTINGS, TABLES, AND FIGURES

The index of planned listing, tables and figures is presented below. The final list may be modified prior to soft-lock and mock-ups of the listings, figures and tables will be provided to the clinical programming group.

7.0 PLANNED LISTINGS, TABLES, AND FIGURES

7.1 Planned Tables

7.1.1 Summary Statistics of Plasma Analyte Concentrations

Table 14.10.1.1.1: Summary Statistics of Plasma NKTR-102 Concentrations

Table 14.10.1.1.2: Summary Statistics of the Plasma Irinotecan Concentrations

Table 14.10.1.1.3: Summary Statistics of the Plasma SN38 Concentrations

Table 14.10.1.1.4: Summary Statistics of the Plasma SN38G Concentrations

Table 14.10.1.1.5: Summary Statistics of the Plasma APC Concentrations

7.1.2 Model-predicted Individual and Population Mean Plasma Pharmacokinetic Parameters

Table 14.10.1.2.1: Individual and Mean PPK Model Parameters for NKTR-102, irinotecan, and SN38

Table 14.10.1.2.2: Individual and Mean PPK Model Parameters for SN38G and APC

7.1.3 Individual Subject Plasma Pharmacokinetic Parameters

Table 14.10.1.3.1: Individual Subject Plasma Pharmacokinetic Parameters for NKTR-102 (Cmax, Tmax, AUCss, CL, V, T1/2)

Table 14.10.1.3.2: Individual Subject Plasma Pharmacokinetic Parameters for Irinotecan (Cmax, Tmax, AUCss)

Table 14.10.1.3.3: Individual Subject Plasma Pharmacokinetic Parameters for SN38 (Cmax, Tmax, AUCss)

Table 14.10.1.3.4: Individual Subject Plasma Pharmacokinetic Parameters for SN38G (Cmax, Tmax, AUCss)

Table 14.10.1.3.5: Individual Subject Plasma Pharmacokinetic Parameters for APC (Cmax, Tmax, AUCss)

7.1.4 Summary Statistics of Plasma Pharmacokinetic Parameters

Table 14.10.1.4.1: Summary Statistics of Plasma Pharmacokinetic Parameters for NKTR-102 (Cmax, Tmax, AUCss, CL, V, T1/2)

Table 14.10.1.4.2: Summary Statistics of the Plasma Pharmacokinetic Parameters for Irinotecan (Cmax, Tmax, AUCss)

Table 14.10.1.4.3: Summary Statistics of the Plasma Pharmacokinetic Parameters for SN38 (Cmax, Tmax, AUCss)

Table 14.10.1.4.4: Summary Statistics of the Plasma Pharmacokinetic Parameters for SN38G (Cmax, Tmax, AUCss)

Table 14.10.1.4.5: Summary Statistics of the Plasma Pharmacokinetic Parameters for APC (Cmax, Tmax, AUCss)

7.1.5 Summary of Exposure-Response Analyses

Table 14.10.1.5.1: Median ORR by SN38 AUCss quartiles and TPC arm

Table 14.10.1.5.2: Summary of Cox Proportional Hazard analysis for ORR by SN38 AUCss quartiles and TPC arm

Table 14.10.1.5.3: Median PFS by SN38 AUCss quartiles and TPC arm

Table 14.10.1.5.4: Summary of Cox Proportional Hazard analysis for PFS by SN38 AUCss quartiles and TPC arm

Table 14.10.1.5.5: Median OS by SN38 AUCss quartiles and TPC arm

Table 14.10.1.5.6: Summary of Cox Proportional Hazard analysis for OS by SN38 AUCss quartiles and TPC arm

7.2 Planned Listings

Listing 16.9.1.2.1: Individual Observed Plasma NKTR-102 Concentrations

Listing 16.9.1.2.2: Individual Observed Plasma Irinotecan Concentrations

Listing 16.9.1.2.3: Individual Observed Plasma SN38 Concentrations

Listing 16.9.1.2.4: Individual Observed Plasma SN38G Concentrations

Listing 16.9.1.2.5: Individual Observed Plasma APC Concentrations

7.3 Planned Figures

Figure 14.10.2.3.1: Observed NKTR-102 concentrations overlaid on model-predicted NKTR-102 (typical subject and 90% confidence interval) range of concentrations

Figure 14.10.2.3.2: Observed irinotecan concentrations overlaid on model-predicted etirinotecan (typical subject and 90% confidence interval) range of concentrations

Figure 14.10.2.3.3: Observed SN38 concentrations overlaid on model-predicted SN38 (typical subject and 90% confidence interval) range of concentrations

Figure 14.10.2.3.4: Observed SN38G concentrations overlaid on model-predicted SN38G (typical subject and 90% confidence interval) range of concentrations

Figure 14.10.2.3.5: Observed APC concentrations overlaid on model-predicted APC (typical subject and 90% confidence interval) range of concentrations

Figure 14.10.2.3.6: Individual model fits for NKTR-102, irinotecan, and SN38

Figure 14.10.2.3.7: Individual model fits for APC and SN38G

Figure 14.10.2.3.8: Boxplots of AUC_{ss} for SN38 in subjects homozygous for the UGT1A1*28 allele and subjects with other UGT1A1 genotypes

Figure 14.10.2.3.9: Kaplan-Meier plot of ORR stratified by SN38 AUC_{ss} exposure quartiles and TPC treatment arm

Figure 14.10.2.3.10: Kaplan-Meier plot of PFS stratified by SN38 AUC_{ss} exposure quartiles and TPC treatment arm

Figure 14.10.2.3.11: Kaplan-Meier plot of OS stratified by SN38 AUC_{ss} exposure quartiles and TPC treatment arm

APPENDIX 2: SUMMARY OF FINAL PARAMETER ESTIMATES IN THE 3-ANALYTE MODEL (NKTR-102, IRINOTECAN, SN38) IN A TYPICAL FEMALE PATIENT WITH THE MEDIAN BSA OF 1.79 M^2 AND MEDIAN EGFR OF 83.4 ML/MIN

Analyte	Parameter	Population Parameter, Mean ± SE	Interindividual Variability, Variance ± SE (CV%)‡
NKTR-102	*CL (L/hr)*	0.277 ± 0.006	0.0662 ± 0.0082 (26%)
	V_1 *(L)*	5.01 ± 0.25	0.0700 ± 0.0089 (26%)
	k_{1p} *(hr^{-1})*	0.0048 ± 0.00023	0.228 ± 0.030 (48%)
	k_{p1}*(hr^{-1})*	0.00083 ± 0.00003	ND
	$\theta_{CL,BSA}$	0.944 ± 0.148	ND
	$\theta_{CL,eGFR}$	0.334 ± 0.070	ND
	$\theta_{V_1,BSA}$	0.881 ± 0.180	ND
	$\theta_{V_1,female}$	-0.237 ± 0.058	ND
	Multiplicative Error	0.274 ± 0.004	ND
Irinotecan	$V_2^* = V_2/F_{12}$ *(L)*	4.10 ± 0.20	0.264 ± 0.038(51%)
	k_{2e}*(hr^{-1})*	14.70 ± 0.55	0.0644 ± 0.0147(46%)
	k_{2p}*(hr^{-1})*	15.31 ± 0.75	0.210 ± 0.034(46%)
	k_{p2}*(hr^{-1})*	0.0024 ± 0.00007	ND
	Corr[k_{2e},V_2*]	-0.768 ± 0.059	ND
	Multiplicative Error	0.38 ± 0.006	ND
SN38	$V_3^{**} = \frac{V_3}{F_{12}F_{23}}(L)$	2111 ± 89	0.238 ± 0.031(49%)
	k_{3e}*(hr^{-1})*	0.071 ± 0.0031	0.194 ± 0.033(44%)
	k_{3p}*(hr^{-1})*	0.569 ± 0.032	0.385 ± 0.057(62%)
	k_{p3}*(hr^{-1})*	0.013 ± 0.00049	ND
	Multiplicative Error	0.331 ± 0.006	ND
	Additive Error	0.330 ± 0.028	ND

‡CV% computed as the square root of the variance; ND: not determined.
Source: LS-2015-504, Table 5

Summary of Final Pharmacokinetic Parameter Estimates of SN38G and APC from 5-Analyte Model

Analyte	Parameter	Population Parameter, Mean ± SE	Interindividual Variability, Variance ± SE (CV%)‡
SN38G	$V_4^* = \frac{V_4}{F_{12}(F_{23}+F_{25})F_{34}}(L)$	8.15 ± 0.49	0.279 ± 0.041 (53%)
	$\theta_{V4,UGT1A1}$	0.481 ± 0.160	ND
	$k_{4e}(hr^{-1})$	2.13 ± 0.13	0.298 ± 0.041 (55%)
	$k_{4p}(hr^{-1})$	7.52 ± 2.30	0.602 ± 0.300 (78%)
	$k_{p4}(hr^{-1})$	10.7 ± 4.6	ND
	Exponential Error	0.413 ± 0.0062	ND
APC	$F_{irinotecan \rightarrow SN38}$	0.759 ± 0.0092	0.0229 ± 0.0029 (15%)
	$k_{5e}(hr^{-1})$	0.0349 ± 0.0014	0.154 ± 0.025 (39%)
	$k_{5p}(hr^{-1})$	0.0109 ± 0.0018	1.41 ± 0.32 (119%)
	$k_{p5}(hr^{-1})$	0.00306 ± 0.00033	ND
	Exponential Error	0.343 ± 0.0056	ND
Other parameters	Corr[k_{2e},V_2]	-0.868 ± 0.024	ND

‡CV% computed as square root of variance; ND, not determined.
Source: LS-2015-504, Table 6

NEKTAR

Data Monitoring Committee (DMC) Charter

Nektar Protocol 15-102-14	A Phase 3 Open-Label, Randomized, Multicenter Study of NKTR-102 versus Treatment of Physician's Choice (TPC) in Patients with Metastatic Breast Cancer Who Have Stable Brain Metastases and Have Been Previously Treated with an Anthracycline, a Taxane, and Capecitabine
Sponsor:	Nektar Therapeutics 455 Mission Bay Boulevard South San Francisco, CA 94158 USA Tel: +415.482.5300
Version 1.0 Version 2.0 Version 3.0	24 October 2016 11 October 2018 30 April 2019
Sponsor's Responsible Medical Officer:	Alison L. Hannah, MD
Sponsor's Responsible Statistician:	Yining Du, PhD

[redacted]