Janssen Research & Development

Statistical Analysis Plan

A Phase 2 Study of Erdafitinib in Subjects with Advanced Solid Tumors and FGFR Gene Alterations

Protocol 42756493CAN2002; Phase 2

JNJ-42756493 (erdafitinib)

Status: Final
Date: Aug 5, 2022
Prepared by: Janssen Research & Development, LLC
Document No.: EDMS-RIM-809941

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

VERSION HISTORY

Table 1– SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
1	Feb. 28, 2020	Not Applicable	Initial release
2	Sep. 14, 2020	Added the language to describe the separate Cholangiocarcinoma Expansion Cohort	A DRC meeting was triggered by the cholangiocarcinoma histology approaching the predefined cap of approximately 30 subjects. In July 2020, a DRC meeting was convened to consider expansion of enrollment of subjects with cholangiocarcinoma. Based on review of safety and efficacy data, the DRC recommended to continue the study and to allow expanded enrollment in the cholangiocarcinoma histology. As a result, a Cholangiocarcinoma Expansion Cohort was added to enroll subjects outside the Broad Panel Cohort.
3	Nov. 03, 2020	Integrated the interim analysis plan within the main study analysis plan	Many elements in the two analysis plans are overlapped thus combining to reduce the number of documents. A specific tumor histology list for futility analysis is pre-defined considering the actual trial enrollment and projected enrollment, for example, Esophageal cancer and non-squamous NSCLC are newly considered to list.
4	Aug. 08, 2022	• Added description and analysis specification for the pediatric cohort. • Added secondary endpoint of clinical benefit rate. • Added more specifications for the timing of primary analysis and final analysis of the study in terms of different cohorts. • Some language modified to be aligned with protocol amendment 3.	To be aligned with protocol amendment 3.
		Updated specification for Cholangiocarcinoma Expansion Cohort, including a sensitivity analysis pooling the Cholangiocarcinoma Expansion Cohort with Broad Panel Cohort.	To specify the sensitivity analysis for the Cholangiocarcinoma Expansion Cohort.
		Added more details for biomarker analysis.	To specify biomarker related analysis, including one sensitivity

SAP Version	Approval Date	Change	Rationale
			analysis considering molecular eligibility.
		Added gamma value for Hwang-Shih-Decani error spending function.	To clarify the intended spending function.
		Added more details for PRO analysis.	To specify PRO instruments and related analysis.
		Modified language for PK analysis.	To further define PK related analysis.
		Modified section on prior and concomitant medications	To clarify the analysis for concomitant medications, prior anti-cancer therapy and to add analysis for subsequent therapy.
		Added language for Ophthalmologic Examination.	To specify safety analysis for Ophthalmologic Examination.
		Added sensitivity analysis pooling cholangiocarcinoma expansion cohort and/or pediatric cohort.	To present efficacy results on all treated subjects with the same molecular eligibility for the Broad Panel Cohort.
		Removed the sensitivity analysis for subjects by centrally confirmed FGFR status.	Analysis for central confirmation with the Foundation Medicine (FMI) test in terms of concordance, efficacy, etc will be executed by FMI in a clinical bridging study but not by Janssen.
		Added supplementary analysis of subjects assessed by RECIST v1.1 or RANO criteria.	To present efficacy results separating subjects assessed by RECIST v1.1 criteria and by RANO criteria for tumor response.
		Modified section adverse events of clinical interest and other safety observations.	To clarify the safety analysis for AE of special interest and clinical importance.
		Modified the ECG RR interval upper limit of normal range.	To be consistent with the normal range for heart rate.
		Added separate efficacy analysis for Japan submission in the Appendix	Different endpoints and analysis are specified for Japan submission, based on the discussion with PMDA
		Updates were made to the endpoints of the Exploratory Cohort.	To clarify the endpoints for Exploratory Cohort analysis.

ABBREVIATIONS

AE	adverse event
ALT/SGPT	alanine aminotransferase
AST/SGOT	aspartate aminotransferase
ATC	anatomic and therapeutic Class
BHM	Bayesian hierarchical model
BOR	best objective response
BSA	body surface area
BUN	blood urea nitrogen
CBR	clinical benefit rate
CI	confidence interval(s)
CR	complete response
CrCl	creatinine clearance
CRF	case report form
CTCAE	Common Terminology Criteria for Adverse Events
CV	coefficient of variation
DBP	diastolic pressure
DCR	disease control rate
DOR	duration of response
DRC	data review committee
ECG	electrocardiogram(s)
ECOG	Eastern Cooperative Oncology Group
FGFR	fibroblast growth factor receptor
FPR	false positive rate
FWER	family-wise error rate
HN	Squamous cell head and neck cancers
HSD	Hwang-Shih-DeCani
ICH	International Conference on Harmonization
IGF-1	Insulin-like growth factor 1
IRC	Independent Review Committee
MedDRA	Medical Dictionary for Regulatory Activities
NGS	next-generation sequencing
ORR	overall response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PMDA	Pharmaceuticals and Medical Devices Agency
PTH	parathyroid hormone
PK	pharmacokinetic(s)
PRO	patient-reported outcome(s)
PT	preferred terms
P-gp	P-glycoprotein
QoL	quality of life
RANO	Response Assessment in Neuro-Oncology
RECIST	Response Evaluation Criteria in Solid Tumors
RCC	renal cell cancer
SAP	Statistical Analysis Plan
SBP	systolic blood pressure
SD	standard deviation
SOC	System Organ Class
STS	soft tissue sarcoma
TEAE	treatment-emergent adverse event(s)
TSH	Thyroid stimulation hormone
ULN	upper limit of normal
WBC	white blood cell
WHO-DD	World Health Organization Drug Dictionary

1. INTRODUCTION

Erdafitinib (JNJ-42756493) is a selective and potent pan FGFR 1-4 inhibitor with demonstrated clinical activity in subjects with solid tumors identified to have alterations in the FGFR pathway. In the United States, erdafitinib (BALVERSA™) was approved on 12 April 2019 for the treatment of adult patients with locally advanced or metastatic urothelial carcinoma, that has susceptible FGFR3 or FGFR2 genetic alterations, and has progressed during or following at least 1 line of prior platinum-containing chemotherapy, including within 12 months of neoadjuvant or adjuvant platinum-containing chemotherapy. This Phase 2, open-label study will evaluate the safety and efficacy of erdafitinib in subjects with advanced solid tumors (other than urothelial tumors), and FGFR gene alterations.

The purpose of the statistical analysis plan (SAP) is to lay out key elements including definitions and statistical methods for the planned analyses for the primary, secondary, exploratory, and other endpoints at the interim, primary, and final analysis.

1.1. Trial Objectives

1.1.1 Primary Objectives

Broad Panel Cohort and Core Panel Cohort

To evaluate the efficacy of erdafitinib in terms of overall response rate (ORR) as assessed by the Independent Review Committee (IRC) in subjects with advanced solid tumors with target FGFR mutations and any gene fusions (Broad Panel Cohort), or in a pre-specified subgroup of subjects with a selected panel of FGFR markers (Core Panel Cohort), or in both cohorts

Pediatric Cohort

To evaluate the efficacy of erdafitinib in terms of ORR as assessed by the IRC in pediatric subjects with advanced solid tumors with FGFR mutations, any gene fusions, or FGFR internal tandem duplication (Pediatric Cohort), including adolescent subjects with target FGFR mutations and any gene fusions

1.1.2 Secondary Objectives

The second objectives of the study are:

- To evaluate the efficacy of erdafitinib, in terms of ORR, as assessed by the Investigator
- To evaluate the efficacy of erdafitinib in terms of duration of response (DOR)
- To evaluate other measures of efficacy including disease control rate (DCR), clinical benefit rate (CBR), progression-free survival (PFS), and overall survival (OS)
- To evaluate erdafitinib pharmacokinetics (PK)
- To evaluate safety and tolerability of erdafitinib
- To evaluate health-related quality of life (QoL)

1.1.3 Exploratory Objectives

The exploratory objectives of the study include:

- To evaluate the efficacy of erdafitinib in subjects with advanced solid tumors in the Exploratory Cohort, in terms of ORR by investigator, DOR, DCR, CBR, PFS and OS
- To identify FGFR mutations and tumor histologies sensitive to erdafitinib from the Exploratory Cohort
- To explore PK exposure-response relationships
- To assess tumor-specific biomarkers (DNA, RNA, protein), molecular subtypes and gene expression signatures in archival tumor samples, paired tumor biopsies, and blood that potentially predict tumor response or resistance to erdafitinib
- To evaluate the relationship between CYP2C9 polymorphism and PK of erdafitinib
- To assess the distribution of erdafitinib in the cerebrospinal fluid (CSF) (when available)
- To collect Medical Resource Utilization data that may be used in future economic models

1.2. Trial Design

This is a Phase 2, open-label study of the efficacy and safety of erdafitinib in subjects ≥6 years of age with advanced solid tumors (other than urothelial tumors) and FGFR gene alterations. Subjects ≥12 years of age with target FGFR mutations or any FGFR gene fusions may be enrolled into the Broad Panel Cohort. Target FGFR mutations include select mutations based on predicted likelihood for pathogenicity with preclinical sensitivity to erdafitinib, or those with clinical or correlative evidence supporting inclusion. A subgroup of subjects in the Broad Panel Cohort with a select panel of pre-specified FGFR markers will be identified as the Core Panel Cohort. While the Broad Panel Cohort consists of target FGFR mutations and any fusions, the Core Panel Cohort consists of a select subset of FGFR mutations or fusions with known observed clinical activity in previous studies and/or with a high level of recurrence. Subjects with any other FGFR mutations that are not captured in the Broad Panel Cohort will be included in the study as the Exploratory Cohort. A separate Cholangiocarcinoma Expansion Cohort will enroll additional subjects with target FGFR mutations or any FGFR gene fusion once the Broad Panel Cohort has reached the cap of approximately 30 subjects for cholangiocarcinoma. A Pediatric Cohort will enroll all subjects ⩾6 to <18 years of age with locally advanced or metastatic solid tumors harboring FGFR alterations who have either progressed following prior therapies and who have no acceptable standard therapies, or who have a newly diagnosed solid tumor and who have no acceptable standard therapies. Adolescent subjects enrolled in the Broad Panel Cohort (≥12 to 18 years) are also considered part of the Pediatric Cohort.

The Screening Phase will start with the Molecular Eligibility Screening Period. Patients with study-eligible FGFR alterations may be identified by central next-generation sequencing (NGS) from tissue sample or based on locally performed and commercial testing from tissue or blood (NGS tests, direct digital counting methods, or the Qiagen therascreen® FGFR reverse transcription polymerase chain reaction [RT-PCR] test). The Full-study Screening Period will occur after the completion of prior treatment and documentation of disease progression for subjects who meet the molecular screening criteria. The Treatment Phase will continue until disease

progression, intolerable toxicity, withdrawal of consent, or decision by the investigator to discontinue treatment. The posttreatment Follow-up Phase will extend from the End of Treatment Visit until the subject has died, withdraws consent, is lost to follow-up, or the end of study, whichever comes first.

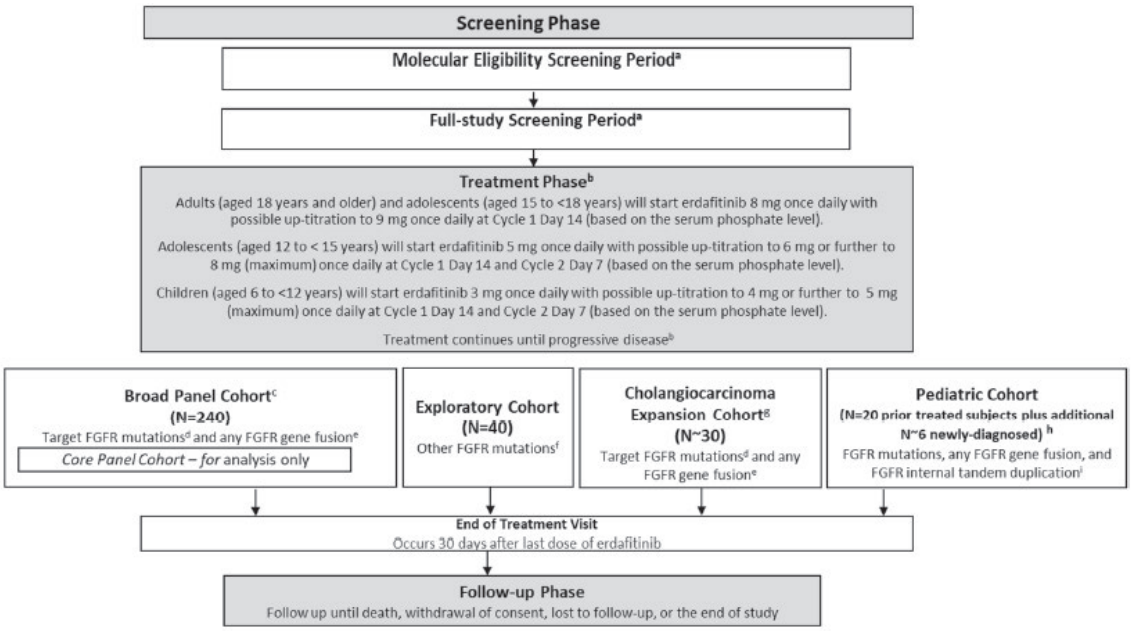
Subjects will take erdafitinib orally once daily for 21 days on a 21-day cycle until disease progression, intolerable toxicity, withdrawal of consent, or decision by the investigator to discontinue treatment. Adults (aged 18 years and older for dosing purposes) and adolescent subjects aged ≥15 to <18 years will start with an erdafitinib dose of 8 mg with possible up-titration to 9 mg based on Cycle 1 Day 14 serum phosphate levels. Adolescent subjects aged ≥12 to <15 years will start with an erdafitinib dose of 5 mg with possible up-titration to 6 mg or further to 8 mg based on Cycle 1 Day 14 and Cycle 2 Day 7 serum phosphate levels. Children aged ≥6 to <12 years will start with an erdafitinib dose of 3 mg with possible up-titration to 4 mg or further to 5 mg based on Cycle 1 Day 14 and Cycle 2 Day 7 serum phosphate levels.

Assessment of response will be performed according to Response Evaluation Criteria in Solid Tumors (RECIST, version 1.1) or Response Assessment in Neuro-Oncology (RANO) by the IRC and investigators. Pharmacokinetic assessments (plasma concentrations of erdafitinib and alpha-1-acid glycoproteins, total protein, and fraction unbound, if required, using venous blood samples), biomarker assessments (molecular screening to determine eligibility for the study; and exploratory DNA, RNA, and protein analyses using archival or fresh biopsy tissue and blood (ctDNA) for exploratory research), patients' health-related quality of life (QoL) assessments, and safety assessments (including adverse event [AE] reports and results of vital sign measurements, electrocardiograms [ECGs], physical examinations, clinical laboratory tests, performance status assessment, and ophthalmologic examinations) will be conducted as described on the Schedule of Activities. Additional safety assessments for children and adolescents include radiographic (growth plate assessment and bone age) imaging, DEXA scan for bone densitometry, and clinical laboratory tests for thyroid stimulation hormone (TSH), total triiodothyronine (T3), and free thyroxine (T4) and insulin-like growth factor 1 (IGF-1) will also be conducted as described on the Schedule of Activities in the protocol.

Approximately 280 subjects ≥12 years of age with FGFR genetic alterations will be enrolled in the Broad Panel Cohort (approximately 240 subjects for 200 response-evaluable subjects) and Exploratory Cohort (approximately 40 subjects). An additional, approximately 30 subjects will be enrolled in the separate Cholangiocarcinoma Expansion Cohort. The Pediatric Cohort is planned to enroll 20 children or adolescent subjects who had progressed following prior therapies and who had no acceptable standard therapies, and approximately 6 additional children or adolescent subjects who had a newly-diagnosed solid tumor and who had no acceptable standard therapies.

Three interim futility analyses for the Broad Panel Cohort are planned when 30%, 50%, and 70% of the subjects in the Broad Panel Cohort (i.e., approximately 60, 100, and 140 subjects) have been treated and are considered response-evaluable by investigator assessment, irrespective of the tumor histologies and the distribution among the tumor histologies. The interim analyses for futility will be based on the ORR by investigator assessment using a Bayesian hierarchical model (BHM), implemented in FACTS v6.2 Enrichment Design – Dichotomous. In addition, an interim efficacy analysis will be conducted for the Broad Panel Cohort at the same time as the second interim futility analysis.

Figure 1: Schematic Overview of the Study



[a] See Protocol Section **Error! Reference source not found.** for information regarding consent during the screening periods. During the Molecular Eligibility Screening Period, either the central or local screening approach may be followed as described in Protocol Section **Error! Reference source not found.**. Starting with protocol amendment 4, only local reports will be used for molecular screening.
[b] See Protocol Section **Error! Reference source not found.** for detailed information regarding titration of erdafitinib based on serum phosphate levels. Treatment is continued until progressive disease, intolerable toxicity, consent withdrawal, or investigator decision to stop treatment. See Protocol Section **Error! Reference source not found.** for continuation of treatment after progressive disease. Note that children and adolescent subjects remain under the titration and dosing schedule under which they were enrolled.
[c] Up to 30 subjects in each tumor histology will be enrolled in the Broad Panel Cohort with the possibility of increasing the number of subjects in a specific tumor histology upon decision by the sponsor's Data Review Committee. The rationale for biomarker selection is described in Protocol Section **Error! Reference source not found.**.
[d] Subjects with target FGFR mutations are eligible for enrollment in the Broad Panel Cohort. The List of Target FGFR Mutations (over 30 validated FGFR mutations) is provided in Protocol Section **Error! Reference source not found.**.
[e] Subjects with any FGFR gene fusion are eligible for enrollment in the Broad Panel Cohort. FGFR gene fusions must have an intact FGFR kinase domain. FGFR gene identifiers for reference are provided in Protocol Section **Error! Reference source not found.**, Criterion 2.
[f] Subjects with FGFR mutations that do not meet the Broad Panel Cohort criteria (ie, not on the List of Target FGFR Mutations) are eligible for enrollment in the Exploratory Cohort.
[g] A separate Cholangiocarcinoma Expansion Cohort will enroll subjects with target FGFR mutations or any FGFR gene fusion once the Broad Panel Cohort has reached the cap of approximately 30 subjects for cholangiocarcinoma.
[h] The Pediatric Cohort will enroll 20 children and adolescent subjects who have progressed following prior therapies and who have no acceptable standard therapies, and approximately 6 additional subjects who have a newly diagnosed solid tumor and who have no acceptable standard therapies. Adolescent subjects enrolled in the Broad Panel Cohort (≥ 12 to 18 years) will be analyzed as part of the Broad Panel Cohort and the Pediatric Cohort; therefore 240 subjects in the Broad Panel Cohort can include subjects from Pediatric Cohort.
[i] Subjects with any FGFR mutation (exclusive of valine gatekeeper and resistance alterations), any FGFR gene fusion, or FGFR internal tandem are eligible for enrollment in the Pediatric Cohort (see protocol Section **Error! Reference source not found.**, Criterion 2).

1.3. Statistical Hypotheses for Trial Objectives

The primary hypothesis of this study is that treatment with erdafitinib will improve ORR with a target rate of 35% over a null rate of 15% in subjects with advanced solid tumors that harbor target FGFR alterations in the Broad Panel Cohort, the Core Panel Cohort, or both cohorts. The corresponding null and alternative hypotheses are as follows:

$$H_0: ORR \leq 15\% \text{ vs } Ha: ORR \geq 35\%.$$

1.4. Sample Size Justification

Approximately 280 subjects ⩾12 years of age with FGFR genetic alterations will be enrolled in the Broad Panel Cohort (approximately 240 subjects for 200 response-evaluable subjects, note that adolescent subjects will be enrolled in the Broad Panel Cohort until the cohort is full or the sample size cap of 30 for the tumor histology is met) and the Exploratory Cohort (approximately 40 subjects). An additional, approximately 30 subjects will be enrolled in the Cholangiocarcinoma Expansion Cohort. A separate Pediatric Cohort (⩾6 to <18 years) will enroll 20 children or adolescent subjects who have progressed following prior therapies and who have no acceptable standard therapies, and approximately 6 additional children or adolescent subjects who have a newly-diagnosed solid tumor and who have no acceptable standard therapies.

1.4.1. Broad Panel Cohort

The Broad Panel Cohort is expected to have 200 response-evaluable subjects. The primary analysis for the Broad Panel Cohort will be conducted 6 months after the first dose of the last Broad Panel Cohort subject. The total number of subjects enrolled to reach 200 response-evaluable subjects may vary depending on the number of non-response-evaluable subjects.

The sample size of 200 response-evaluable subjects for the Broad Panel Cohort is selected based on extensive simulations with various Bayesian Hierarchical Model designs to achieve approximately 80% power to select at least 80% of the tumor histologies (e.g., ≥12 of 15 tumor histologies) at the primary analysis with the true ORR being ≥35% for each histology. For enrollment in each tumor histology, the sample size is capped at approximately 30 subjects. For comparison, a Simon minimax 2-stage design requires 28 subjects for Type I error rate of 0.05 and Type II error rate of 0.20. The cap of 30 applies to a pre-defined list of tumor histologies in Table 3 which also includes a group of "Other" to cover additional tumor histologies not listed and the cap also applies to this group. For the tumor histology that is identified as active and safe, further enrollment beyond the cap of approximately 30 subjects may be allowed at the discretion of the DRC. These subjects will be enrolled in separate cohorts and not included in the primary efficacy analysis of the Broad Panel Cohort.

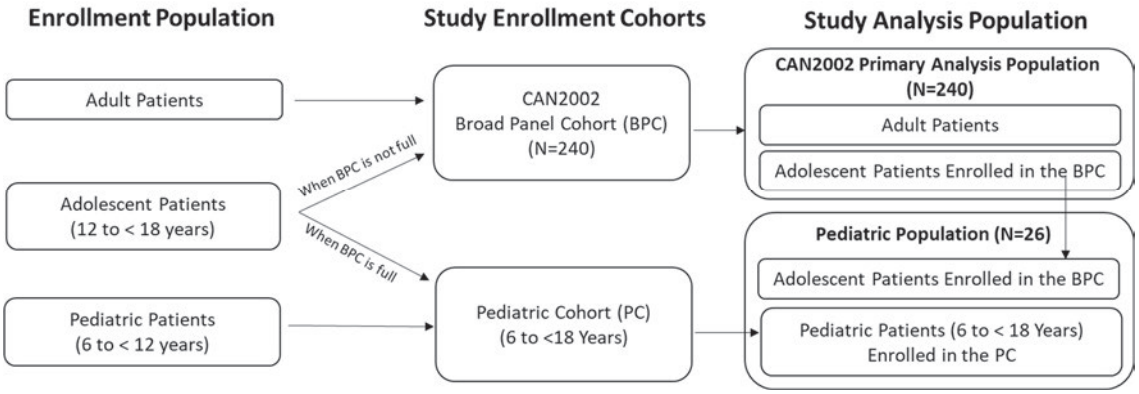
The details about the BHM, selection power, and assumption on tumor histology prevalence for sample size justification will be discussed in the Appendix A.3.

1.4.2. Cholangiocarcinoma Expansion Cohort

A DRC meeting was triggered by the cholangiocarcinoma histology approaching the predefined cap of approximately 30 subjects. In July 2020, a DRC meeting was convened to consider expansion of enrollment of subjects with cholangiocarcinoma. Based on review of safety and efficacy data, the DRC recommended to continue the study and to allow expanded enrollment in the cholangiocarcinoma histology. As a result, a Cholangiocarcinoma Cohort was added to enroll subjects outside the Broad Panel Cohort. Approximately 30 additional subjects are intended to be enrolled in the Cholangiocarcinoma Expansion Cohort to further characterize the clinical activity for this histology.

1.4.3. Pediatric Cohort

A separate Pediatric Cohort will enroll subjects ⩾ 6 to <18 years of age with locally advanced or metastatic solid tumors harboring FGFR alterations. These subjects include 1) 20 planned subjects, including adolescent subjects enrolled in the Broad Panel Cohort, who have either progressed following prior therapies and who have no acceptable standard therapies, and 2) anticipated 6 additional subjects enrolled who have newly diagnosed solid tumors and who have no acceptable standard therapies. Comparable response from erdafitinib treatment is expected in the pediatric and adult study populations. With an observed ORR of 35%, the 20 children and adolescent subjects in 1) above will provide a 95% Clopper-Pearson confidence interval (CI) of 15.4% to 59.2%. Note that the enrollment in the Pediatric Cohort will be unaffected by any futility analysis for the Broad Panel Cohort.



Note: Adolescent subjects (⩾12 to <18 years) will be enrolled in the BPC, if eligible per BPC criteria, until this cohort is filled or the applicable tumor type cap is met. Otherwise, patients will be enrolled in the Pediatric Cohort. Adolescent subjects in the BPC will also be counted in the Pediatric Cohort.

1.5. Randomization and Blinding

As this is an open-label, single arm study, randomization and blinding procedures are not applicable.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Pooling Algorithm for Analysis Centers

The data from all participating centers in the study will be pooled together for analyses.

2.2. Analysis Sets

2.2.1. Efficacy Analysis Set(s)

2.2.1.1. Broad Panel Cohort

The alterations eligible for inclusion in the Broad Panel Cohort consist of target FGFR mutations and any FGFR fusion. The list of Target FGFR mutations eligible for the Broad Panel Cohort is

provided in Protocol Section 10.11. The Broad Panel Cohort will represent the primary cohort of interest for analysis. The efficacy analysis sets of the Broad Panel Cohort are specified below:

- The Treated Population will consist of all subjects (FGFR+) who received at least 1 dose of study drug. The Treated Population is the primary population for analyses of efficacy, baseline demographics and disease characteristics, and patient-reported outcome(s).
- The Response-evaluable Population will include all subjects (FGFR+) who satisfy the following criteria:
  - Met all eligibility criteria for the study;
  - Received at least 1 dose of study drug;
  - Had a baseline and at least 1 adequate post-treatment radiological disease evaluation, or had clinical signs or symptoms of disease progression, or died prior to the first posttreatment disease evaluation (these subjects will be considered non-responders). Adequate disease assessment is defined as having enough evidence to indicate that progression has or has not occurred.

  The Response-evaluable Population will be used for the interim analysis, and supplementary efficacy analysis on key endpoints such as ORR and DOR.
- The Core Panel Cohort, which is a subgroup of the Treated Population in the Broad Panel Cohort with a select panel of pre-specified FGFR markers, will be assessed for efficacy as part of the primary analysis. The subset of FGFR markers was selected based on observed clinical activity in previous studies or a high level of recurrence of these alterations across tumor histologies; the subset of FGFR markers is outlined below:
  - FGFR3 mutations: S249C; Y373C; R248C; G370C
  - FGFR2 mutations: C382R
  - FGFR3 fusions: FGFR3-TACC3
  - FGFR2 fusions: FGFR2-BICC1; FGFR2-TACC2

2.2.1.2. Exploratory Cohort

Subjects enrolled in Exploratory Cohort who receive at least 1 dose of study drug will be evaluated for efficacy and safety as the exploratory analysis.

2.2.2. Safety Analysis Set

The Treated Population in the Broad Panel Cohort (including the Core Panel Cohort), the Cholangiocarcinoma Expansion Cohort, the Exploratory Cohort and the Pediatric Cohort will be used for all safety analyses unless otherwise stated.

2.2.3. Pharmacokinetics (PK) Analysis Set

The PK analysis set will be aligned with efficacy analysis set and safety analysis set as applicable for PK evaluable subjects defined as:

- who have received at least 1 dose of study drug;

- who have at least one evaluable PK sample

A plasma PK sample is considered to be not evaluable if

- there are e.g., missing information of dosing and sampling times
- vomiting occurs within the first 2.5 hours following the last dose intake prior to the predose PK sample draw
- vomiting occurs within first 2.5 hours after Cycle 2 Day 1 dose
- the patient does not take study drug according to the originally assigned dose and assigned dose post up-titration for at least 5 consecutive days without interruption or dose modification prior to the PK sampling day
- For the predose PK sample, the draw occurs outside the 18-30 hours window after the last dose intake
- For the predose PK sample, the draw occurs after the next dose

A cerebrospinal fluid (CSF) PK sample is considered to be not evaluable if there are e.g., missing information of dosing and sampling times. Additionally, a sample can be considered to be not evaluable as per scientific judgement of the clinical pharmacology expert.

2.2.4. Cholangiocarcinoma Expansion Cohort

Subjects enrolled in Cholangiocarcinoma Expansion Cohort with target FGFR mutations or any FGFR fusion who receive at least 1 dose of study drug, will be evaluated for efficacy and safety as a separate analysis. The Cholangiocarcinoma Expansion Cohort will primarily be evaluated separately from the Broad Panel Cohort. Analysis may also be conducted by pooling these subjects with subjects from Broad Panel Cohort.

2.2.5. Pediatric Cohort

The Treated Population in the Pediatric Cohort with locally advanced or metastatic solid tumors harboring FGFR mutations, any gene fusions or FGFR internal tandem duplication who have either progressed following prior therapies and who have no acceptable standard therapies, or who have a newly diagnosed solid tumors and who have no acceptable standard therapies, and receive at least 1 dose of study drug, will be evaluated as a separate statistical analysis for efficacy and safety. The treated subjects (approximately 20) who have either progressed following prior therapies and who have no acceptable standard therapies in the proposed Pediatric Cohort (including the adolescent subjects enrolled in the Broad Panel Cohort) will be the primary analysis population for the Pediatric Cohort. The additional approximately 6 subjects who have a newly diagnosed solid tumors and who have no acceptable standard therapies will be analyzed separately. Pooled analysis may also be conducted by pooling with subjects from other cohorts.

2.3. Definition of Subgroups

Subgroup analysis will be performed for Broad Panel Cohort and Cholangiocarcinoma Expansion Cohort for the selected variables to assess the internal consistency of efficacy and/or safety. The subgroup variables and the cutoff values listed in Table 2 are subject to change if warranted to better represent the data.

Subgroup analysis for countries/regions (e.g. China, East Asia) may also be performed.

Table 2: Subgroup Definition

Subgroup	Definition of Subgroup
Age	≥6 to <12 years (applicable to the Pediatric Cohort only), ≥12 to <18 years, ≥18 to < 65 years, ≥65 years
Sex	Male, Female
ECOG	0, 1
Geography	North America, Europe, Rest of the World
Number of Lines of Prior Therapy	1, 2, ≥3
FGFR alteration type	Mutation, fusion
FGFR gene	FGFR1-4
Up-titration status	Yes, No
Maximum serum phosphate within first 3 months	< 7 mg/dL, >= 7 mg/dL

In addition, tumor histologies (with the pre-defined list in Table 3 or all observed tumor histologies), may also be used to define subgroups for efficacy and/or safety.

Table 3: List of Tumor Histologies for Futility and Cap

Number	Tumor Histologies
1	Cholangiocarcinoma (CCA)
2	High-grade glioma (HGG)
3	Squamous NSCLC (sqNSCLC)
4	Squamous cell head and neck cancers (HNSCC)
5	Gastric Cancer (GSTRC)
6	Breast Cancer (BRST)
7	Endometrial Cancer (EDMTL)
8	Low-grade glioma (LGG)

9	Ovarian Cancer (OVAR)
10	Non-squamous NSCLC (nonsqNSCLC)
11	Colorectal Cancer (CRC)
12	Pancreatic Cancer (PANCR)
13	Cervical Cancer (CRVX)
14	Esophageal (ESOPH)
15	Other

Note: The group Other will enroll on all other tumor histologies not listed. The group "Other" shares the same cap of approximately 30 and will be included in the BHM evaluation for information borrowing only but will not be deemed futile early in the interim analyses and will continue enrollment until cap is reached.

2.4. Study Day and Relative Day

A treated cycle for a specific drug is defined as a cycle in which the subject received any amount of the specific drug. The cycle number will be named according to the sequence of every 21-day cycle for study agent administration.

Assessments will be presented chronologically by study day or cycle day as described below:

Reference date (Day 1) = first dose date of study medication.

Study Day = assessment date – reference date + 1 for assessment performed on or after the reference date; or assessment date – reference date for assessment performed before the reference date.

Cycle Day = assessment date - date of the first dose for the cycle + 1.

There is no "Day 0".

2.5. Baseline and Endpoint

Baseline is defined as the last non-missing observation prior to the start of the first dose of study drug.

Endpoint is defined as the last available postbaseline result within the analysis period. Unscheduled visit results are included in this definition and will be considered as the endpoint value if the unscheduled visit result is the last postbaseline result available within the analysis period.

2.6. Imputation Rules for Missing and Partial Dates

In general, imputation of missing dates will be made for AE onset date, AE resolution date, date of death, start and end dates of concomitant and subsequent therapies, and date of initial diagnosis according to the following rules. Start date will be imputed before end date.

- If date is completely missing, no imputation will be made.
- If year is missing, no imputation will be made.
- If only year is present but month and day are missing, then June 30th will be used.
- If only day is missing but year and month are available, then the 15th of the month will be used.

However, the above imputations will be modified by the following rules:

- If such imputed date for initial diagnosis is on or after the first dose date, then the first dose date - 1 will be used.
- The imputed start date for subsequent therapies will be adjusted sequentially using the following steps:
  - If the imputed start date is before the treatment discontinuation date (or last dose date if no treatment discontinuation date) but in the same year and month, then the treatment discontinuation date (or last dose date if no treatment discontinuation date) will be used.
  - If subsequent therapy end date is not missing and is before the imputed subsequent therapy start date, then the subsequent therapy end date will be used as the start date.
- If the imputed date is for a date of death and is before the last date that the subject is known to be alive, the latter date will be used.
- The imputed AE start date will be adjusted sequentially using the following steps:
  - If the imputed date is in the same year and month as but day before the first dose date, then the first dose date will be used, or if it is in the same year and month as but day after the last dose date + 30 days, then the last dose date + 30 days will be used.
  - If AE end date is not missing and the imputed AE start date is after the AE end date, then the AE end date will be used.
  - If the imputed AE start date and is after date of death, then date of death will be used
  - If the imputed AE start date is in the same month and year but after the 1st subsequent therapy start date, then 1st subsequent therapy start date will be used.
- If the imputed date is for an AE end date and is after the death date, then the death date will be used, or if the imputed AE end date is before the AE start date, then the AE start date will be used.
- The AE imputation rule will be used for concomitant medication.

3. INTERIM ANALYSIS AND DATA REVIEW COMMITTEE

3.1. Data Review Committee (DRC)

A DRC will be established for monitoring of safety data and for review of interim efficacy data, to ensure the safety of the subjects enrolled in this study, and to assess efficacy objectives. This committee will consist of at least one medical expert and at least one statistician; committee membership responsibilities, authorities, and procedures will be documented in the DRC charter.

After each review, the DRC will make recommendations such as:

- The continuation of the study
- The continuation of enrollment in specific tumor histologies
- The allowance of more enrollment above the cap 30 on a tumor histology or multiple tumor histologies
- The early success of the study

Details of the composition of the DRC and their operational procedures will be specified in the DRC Charter.

3.2. Interim Analysis (Broad Panel Cohort)

Three interim futility analyses are planned for the Broad Panel Cohort when approximately 30%, 50%, and 70% of the subjects in the Broad Panel Cohort (ie, approximately 60, 100, and 140 subjects) have been treated and are considered response-evaluable by investigator assessment, irrespective of the tumor histologies and the subject distribution among the tumor histologies (Figure 2, N represents the number of response-evaluable subjects). The tumor histology list for interim futility analysis is pre-defined in Table 3 (except for the group Other including all other tumor histologies not listed). In addition, an interim efficacy analysis will be conducted for the Broad Panel Cohort at the same time as the second interim futility analysis, Figure 2. An early success may be declared if the observed ORR by IRC is 50% or higher with consistent effect across histologies. Regardless of the outcome of the interim efficacy analysis, study enrollment will continue to better characterize the efficacy for the active tumor histologies. The DRC will review the data from the interim analyses (e.g. efficacy, safety, biomarker etc.) and make recommendations to the study team.

The Treated Population in the Broad Panel Cohort will be used to summarize interim population characteristics and safety; the Safety Analysis Set (combining the Broad Panel Cohort, Cholangiocarcinoma Expansion Cohort, Exploratory Cohort and the Pediatric Cohort) will also be used for safety evaluation; the Response-evaluable Population (a subset of the Treated Population) in the Broad Panel Cohort will be used for the interim futility analysis, unless otherwise specified.

For interim analysis, disease progression and response evaluation based on investigator assessment using RECIST v1.1. or RANO will be primarily used to allow quick decisions; however, evaluation based on available IRC assessment may also be provided as supportive data. Note that

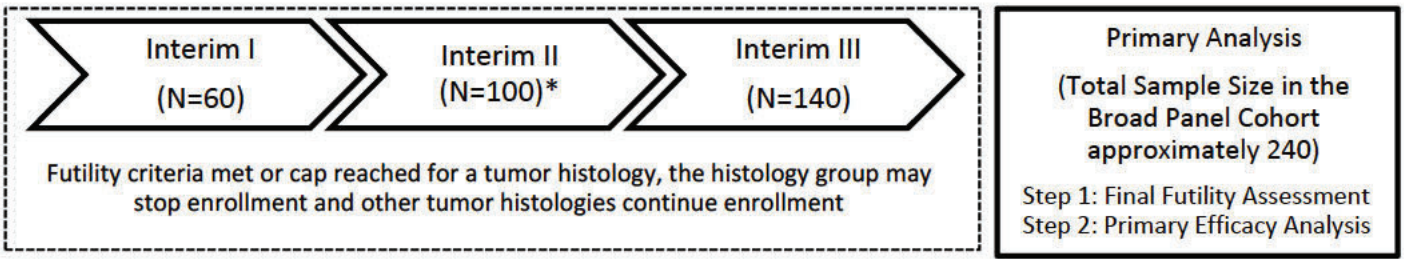
for Interim II, if an early success is to be declared, disease progression and response will be primarily based on IRC assessment.

A Statistical Support Group (SSG), independent of the study team, will support the interim analyses. The SSG will prepare preplanned outputs in Adaptive Clinical Trials Integrated Virtual Environment (ACTIVE) for the futility analysis report, which will be based on overall response rate (ORR) by investigator assessment using a Bayesian hierarchical model, implemented in FACTS v6.2 Enrichment Design-Dichotomous engine (Berry 2013, Chugh 2009). The BHM borrows information across tumor histologies and allows for evaluation of each tumor histology individually given the often-unique features of each group (details in Appendix A.1.). At least 5 response-evaluable subjects in a tumor histology are required for the futility analysis. The futility guideline will be based on the posterior probability of ORR > 25% in the Broad Panel Cohort for a given tumor histology with a threshold of 40%, i.e., a tumor histology meets the model-based futility criteria if the posterior probability Pr(ORR > 25%) < 40%. The operating characteristics of the BHM under different scenarios are detailed in Appendix A.3. For Interim II and III, the futility assessment takes into consideration of more available data from subjects included in previous futility analysis (including the subjects in early futile tumor histologies if applicable) as well as data from newly enrolled subjects after previous interim analysis. In addition, during Interim II and III, histologies that stopped for futility in previous assessment will only be included for BHM information borrowing but will not be assessed for futility again. Note that the group Other will be included in the BHM evaluation for information borrowing only but will not be deemed futile early and will continue enrollment until cap is reached.

Besides the BHM based futility guideline, all available interim data (eg, efficacy, safety, and biomarker data) will be assessed to determine whether enrollment in a given tumor histology should continue, such as the representativeness of biomarker alterations (the expected variant representation by tumor histology based on genomic database assessment). If a tumor histology is deemed futile by the DRC, further enrollment in the histology group will stop; subjects who have been enrolled in the tumor histology group will continue treatment and evaluation as per the protocol. Note that in the primary analysis, a final futility assessment using the same futility guideline as at the interim (i.e, posterior Pr(ORR > 25%) < 40%) but based on ORR by IRC will be performed for all enrolled tumor histologies as part of the primary analysis to identify the tumor histologies for the pooled primary efficacy analysis (Details in section 5.2.3).

Other details of the planned analysis and efficacy/futility assessment are specified in Section 4-6.

Figure 2: Interim Analysis for the Broad Panel Cohort



*At Interim Analysis II, both futility and efficacy will be assessed.

4. SUBJECT INFORMATION

4.1. Demographics and Baseline Characteristics

Subject demographics and baseline disease characteristics will be summarized in Table 4.

Table 4: Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]).
Weight (kg)	
Categorical Variables	
Age (⩾6 to <12 years (applicable to the Pediatric Cohort only), ≥12 and <18 years, ≥18 and <65 years, ≥65 years)	Frequency distribution with the number and percentage of subjects in each category.
Sex (male, female, undifferentiated)	
Race[a] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Other, Multiple)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
Geographical Region (North America, Europe, Rest of the World)	

[a]If multiple race categories are indicated, the Race is recorded as 'Multiple'

Table 5: Baseline Characteristics

Continuous Variables:	**Summary Type**
Time from initial diagnosis to first dose (months)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]).
Hematology	
Hemoglobin	
Chemistry	
Creatinine Clearance	
Categorical Variables	Frequency distribution with the number and percentage of subjects in each category.
Performance Status*	
Metastatic disease sites	
Visceral metastasis	
Number of body sites with metastatic disease	
Prior Radiotherapy	
Prior Systemic Therapy in advanced/metastatic setting	
Prior Cancer-Related Surgery/Procedure	
Number of prior lines of anticancer therapies (1, 2 or ≥ 3)	
FGFR Alteration Type	
FGFR Gene Type	
Tumor Histology Type	

*Baseline performance status (Eastern Cooperative Oncology Group (ECOG) for adults (≥18 years of age), Lansky Score for adolescents (≥6 to <16 years of age), Karnofsky Score for adolescents (≥16 to <18 years of age))

Note that the baseline clinical laboratory tests may not be provided for interim analyses or may be provided by request.

4.2. Disposition Information

The number of subjects in the following disposition categories will be summarized:

- Subjects receiving study agent
- Subjects completed the study

- Subjects who discontinued study agent
- Reasons for discontinuation of study agent
- Subjects who terminated study prematurely
- Reasons for termination of study

A listing of subjects will be provided for the following category:

- Subjects who discontinued study agent

4.3. Treatment Compliance

Relative dose intensity will be summarized descriptively to reflect study agent compliance.

Relative dose intensity will be calculated as follows:

Relative dose intensity (%) = (actual number of dose taken/total number of doses prescribed) x100.

4.4. Extent of Exposure

Descriptive statistics (N, mean, SD, median, and range [minimum, maximum]) for treatment duration, treatment duration excluding dose interruption period, cumulative total dose, and dose intensity will be presented for the safety analysis set and/or the Treated Population in the Broad Panel Cohort. Treatment duration for the study will be calculated as date of last dose of study drug – date of first dose of study drug + 1. Treatment duration excluding dose interruption period are defined as the total number of days that study agent has been administered to the subject (excluding days of study agent interruption).

The number (%) of subjects with a dose adjustment (reduction, interruption, up-titration, re-escalation) will be summarized. Reasons for dose adjustments will also be summarized. Total number of doses skipped and maximum duration of dose skip (days) will also be provided.

4.5. Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact subjects' rights, safety or well-being, or the integrity and/or result of the clinical study. Subjects with major protocol deviations will be identified prior to database lock and the subjects with major protocol deviations will be summarized by category.

- Developed withdrawal criteria but not withdrawn
- Entered study but did not satisfy entry criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Other

A listing of all major protocol deviations will be provided.

4.6. Concomitant Medications

For summarization purposes, concomitant medications will be coded to generic terms using the World Health Organization Drug Dictionary (WHO-DD). According to the protocol, concomitant therapies must be recorded at the time of full-study screening, during the study, and up to 30 days after the last dose of study drug. Concomitant medications used on or after the same day as the first dose of study treatment, including those that started before and continued after the first dose of study treatment will be summarized.

Summaries of concomitant medications will be presented by Anatomical Therapeutic Chemical (ATC) class and drug generic term. The proportion of subjects who receive each concomitant medication will be summarized.

4.7. Prior Anti-cancer Therapy

The number and percentage of subjects who received prior systemic therapy, prior cancer-related surgery/procedure and prior radiotherapy will be summarized. Best response to prior systemic therapy by lines of therapies and last line as well as by therapy types (e.g., chemotherapy, immunotherapy) will be summarized.

4.8. Subsequent Therapies

The number and percentage of subjects who received subsequent therapy (systemic, cancer-related surgery/procedure and radiotherapy) will be summarized.

5. EFFICACY

5.1. Analysis Specifications

The primary analysis for the Broad Panel Cohort will be conducted approximately 6 months after the first dose administration for the last Broad Panel Cohort subject (reaching approximately 200 response-evaluable subjects) and the analysis will be based on Treated Population in the Broad Panel Cohort. At the time of the primary analysis for the Broad Panel Cohort, an analysis of all other cohorts will also be conducted and will include all subjects with the same enrollment cut-off (the date of the last subject enrollment of Broad Panel Cohort) and clinical cut-off as the Broad Panel Cohort. Additional analyses with longer follow-up may be conducted as needed. The final analysis for the Broad Panel Cohort will be performed at the end of the study (per protocol definition).

The Pediatric Cohort will be analyzed separately for efficacy from the Broad Panel Cohort. The final analysis for the Pediatric Cohort will be performed at the end of the study (per protocol definition). The Pediatric Cohort will have the same primary and secondary endpoints as the Broad Panel Cohort. Analysis may also be conducted by pooling subjects from Broad Panel Cohort and Cholangiocarcinoma Expansion Cohort.

The Exploratory Cohort will be analyzed for efficacy based on the Treated Population of the Exploratory Cohort at the time of primary analysis of the Broad Panel Cohort on ORR assessed by investigator.

The Cholangiocarcinoma Expansion Cohort will be evaluated for efficacy at the time of primary analysis and final analysis of the Broad Panel Cohort. Similar endpoints and analyses as Broad Panel Cohort will be conducted for the Cholangiocarcinoma Expansion Cohort.

5.1.1. Level of Significance

In general, statistical inference on the primary endpoint will be tested at 0.025 one-sided level, unless otherwise specified. All interval estimations will be reported using 2-sided 95% CIs.

Multiplicity adjustment for testing the primary endpoint ORR in both Broad Panel Cohort and Core Panel Cohort is based on an error-spending function approach (Spiessens and Debois, 2010). It uses the correlation between the two test statistics in the Core Panel Cohort and the Broad Panel Cohort. The α level for testing the hypothesis in the Core Panel Cohort will use Hwang-Shih-DeCani error spending function with gamma 0 and the information fraction as the ratio of the observed number of subjects in the Core Panel Cohort over the observed number of subjects in the Broad Panel Cohort at the time of primary analysis.

5.1.2. Data Handling Rules

Unless specified otherwise, missing values will not be imputed. Other data handling rules if any will be outlined in the DPS.

5.1.3. General Analysis Considerations

Descriptive statistics and subject listings will be used to summarize the data. For continuous variables, the number of observations, means, standard deviations, medians, and ranges will be used. For discrete variables, frequency and percentage will be presented. For time-to-event variables, Kaplan-Meier estimates will be provided. In general, all interval estimation will be reported using 2-sided 95% CIs.

5.2. Primary Efficacy Endpoint(s)

5.2.1. ORR assessed by IRC Definition

A subject's best objective response (BOR) is defined as the best objective response a subject achieved during the study in the order of Complete Response (CR), Partial Response (PR), Stable Disease (SD), Progressive Disease (PD), where CR and PR are confirmed as per RECIST v1.1 or RANO. Responders are subjects with BOR of CR or PR. Subjects without measurable disease at baseline or without post-baseline tumor assessment are considered as non-responders.

ORR assessed by IRC is defined as the proportion of responders in an analysis population as assessed by IRC.

For the interim analysis (Interim I, II and III), a subject's BOR may also consider the unconfirmed Complete Response(uCR) and unconfirmed Partial Response(uPR). A subject's response will be classified as unconfirmed CR or PR if only one CR or PR has been observed and a subsequent valid disease evaluation has not been performed yet at the time of the interim analysis. However, a subject will be assigned SD/NE if the subject has only one CR or PR and has no subsequent

disease evaluation and is off study or started subsequent anti-cancer therapy. Two sets of analysis based on different definition of responder will be provided for DRC review for decision making:

1) Subjects with unconfirmed response (uCR or uPR) will be considered responders. However, subjects who have initially demonstrated a uCR or uPR but on subsequent evaluation have shown PD will not be considered responders.

2) Subjects with unconfirmed response (uCR or uPR) will not be considered as responders.

5.2.2. Estimand

The primary estimand, the main clinical quantity of interest to be estimated in this study, is defined by the following four components (ICH E9 (R1) Addendum, 2018):

- Population: subjects with advanced solid tumors and target FGFR Gene Alterations.
- Variable: subject responder status (responder – BOR of CR or PR; otherwise – non-responder).
- Intercurrent events: treatment discontinuation and subsequent anti-cancer therapy. The treatment policy strategy will be applied to account for the treatment discontinuation. That is, the value for the variable of interest is used regardless of whether the treatment discontinuation occurs. The while-on-treatment policy will be applied to account for subsequent anti-cancer therapy. That is, only the value for the variable of interest before the subsequent anti-cancer therapy will be used. For instance, if a subject discontinued study treatment but did not start subsequent anti-cancer therapy, the subject's responder status/tumor assessment will be used as if the subject is still on treatment regardless of the treatment discontinuation (i.e. the treatment policy). But if a subject started subsequent anti-cancer therapy, no matter whether the subject has discontinued treatment or not, the subject's responder status/tumor assessment after the subsequent anti-cancer therapy will not be used in the analysis (the while-on-treatment policy).
- Population-level summary: ORR assessed by IRC.

5.2.3. Analysis Methods

5.2.3.1. Primary Analysis

Before the primary efficacy test, a final futility assessment using the same futility guideline at the interim analysis (i.e, Pr(ORR > 25%) < 40%) will be conducted for all enrolled tumor histologies to identify the tumor histologies for the pooled primary efficacy analysis. This futility assessment takes into consideration the possibly enrolled new subjects in tumor histologies after Interim III, as well as more available information of the subjects that continue treatment and become evaluable in the early futile tumor histology groups. All available data (eg, efficacy, safety, and biomarker data) will be assessed by the DRC to determine whether a given tumor histology is futile. The primary efficacy analysis after the final futility assessment will be performed on the pooled data from all tumor histologies that didn't meet futility in this final futility assessment for the Treated Population in the Broad Panel Cohort and the Core Panel Cohort.

The null hypothesis of H_0: ORR ≤ 15% will be tested at a significant level of 0.025 using one-sided exact binomial test. The test will be performed on the pooled data of all tumor histologies that didn't meet the futility criteria in the Broad Panel Cohort. An additional test will be performed on the pooled data from all tumor histologies that didn't meet the futility criteria in the Core Panel Cohort as mentioned in Section 0.

The ORR assessed by IRC and its two-sided 95% Clopper-Pearson CI will also be provided.

5.2.3.2. Supplementary Analysis

Posterior probability of ORR > 25% for Individual Tumor Histologies

As a supplementary analysis, the posterior probability of ORR >25% for each individual tumor histology will also be evaluated. This will provide additional evidence of the contribution of each individual tumor histology that didn't meet the futility criteria in the Broad Panel Cohort to the overall result of the pooled primary efficacy analysis. More details are discussed in the Appendix A.3.

Analysis for Treated Population including Tumor Histologies Deemed Futile

Analysis with all tumor histologies including those deemed futile during the three interim analyses as well as the final futility assessment will also be conducted. This analysis is applicable only when there are tumor histologies deemed futile or excluded from the primary efficacy analysis.

Analysis of Subjects Assessed by RECIST v1.1 or RANO criteria

A supplementary efficacy analysis may be provided for Treated Population in the Broad Panel Cohort separating subjects assessed by RECIST v1.1 criteria and by RANO criteria for tumor response.

5.2.3.3. Sensitivity Analysis

Analysis with Response-evaluable Population

A sensitivity analysis of ORR assessed by IRC may be provided on the pooled data of all tumor histologies that did not meet the futility for the Response-evaluable Population in the Broad Panel Cohort.

Analysis with Subjects Considering Molecular Eligibility

Sensitivity efficacy analysis of ORR assessed by IRC may be performed for Treated Population in the Broad Panel Cohort excluding patients

- without an eligible FGFR alteration (e.g., a patient enrolled with FGFR2-L551F mutation)
- with only known or suspected FGFR germline mutations (FGFR2-A97T, FGFR3-R399C, FGFR3 F384L, FGFR3-A500T, FGFR3-P572A, FGFR3-P572L)

- with only FGFR fusions with a 5'-partner (Partner gene is listed first and FGFR gene is second, eg GENE-FGFR or KLK2-FGFR2), as oncogenic potential of FGFR fusions with a 5'-partner is less well understood as compared to those fusions with 3'partner

- with gatekeeper/resistance FGFR alterations (FGFR1-V561; FGFR2-V564; FGFR3-V555; FGFR4-V550; FGFR1-N546; FGFR2-N549; FGFR3-N540 and FGFR4-N535)

- or with exclusionary co-alterations (For NSCLC subjects only - pathogenic somatic mutations in EGFR or BRAF V600E, KRAS, or any gene fusions in the following genes: ALK, ROS1, or NTRK. For colorectal subjects only – pathogenic somatic mutations in BRAF, KRAS, NRAS and PIK3CA as specified in the protocol)

Analysis Pooling Cholangiocarcinoma Expansion Cohort and/or Pediatric Cohort

A sensitivity efficacy analysis may be provided on the pooled data of Treated Population in the Broad Panel Cohort, Treated Population in the Cholangiocarcinoma Expansion Cohort, and/or Treated Population in the Pediatric Cohort that meet the molecular eligibility for the Broad Panel Cohort.

5.2.3.4. Subgroup Analysis

Subgroup analyses for the ORR by IRC will be conducted for each subgroup (Section 2.3) within the Treated Population, respectively.

5.3. Major Secondary Endpoints

5.3.1. ORR assessed by Investigator

The proportion of subjects in the primary cohort who achieve a CR or PR based on RECIST v1.1 or RANO as assessed by Investigator. The response rate and its two-sided 95% Clopper-Pearson CI will be provided.

Similar analyses as listed the in 5.2.3.1 for the primary endpoint may be conducted as well.

5.3.2. Duration of Response (DOR)

Definition

Duration of response will be analyzed for subjects with the BOR of CR or PR and is defined as the interval between the date of initial documentation of a response and the first documented evidence of PD or death due to any cause. Subjects who are progression-free and alive or have unknown status will be censored at the last tumor assessment. Subjects who had disease progression or death event but started a subsequent anti-cancer therapy before disease progression or death event will be censored at the last disease assessment before the subsequent therapy. Subjects with death or progressive disease after more than one missing DEs will be censored at the date of the last tumor assessment without documented progressive disease.

Analysis method

DOR based on both IRC assessment and investigator assessment will be provided. DOR will be presented with descriptive summary statistics. Kaplan-Meier method will be used for evaluation, if there are sufficient data.

5.3.3. Disease Control Rate (DCR)

Definition

DCR is defined as the proportion of subjects who achieve a best response of CR, PR, or SD based on RECIST v1.1 or RANO (by IRC and Investigator).

For interim analysis, the definition of DCR will also be considered as the proportion of subjects who achieve a best response of CR, PR, uCR, uPR or SD based on RECIST v1.1 or RANO.

Analysis method

The DCR and its two-sided 95% Clopper-Pearson confidence interval will be provided based on both IRC assessment and investigator assessment.

5.3.4. Clinical Benefit Rate (CBR)

Definition

CBR is defined as the proportion of subjects who achieve a best response of CR, PR, or durable SD (defined as a duration of at least 4 months) based on RECIST v1.1 or RANO (by IRC and Investigator).

For interim analysis, the definition of CBR will also be considered as the proportion of subjects who achieve a best response of CR, PR, uCR, uPR or durable SD (defined as a duration of at least 4 months) based on RECIST v1.1 or RANO.

Analysis method

The CBR and its two-sided 95% Clopper-Pearson CI will be provided based on both IRC assessment and investigator assessment.

5.3.5. Progression-free Survival (PFS)

Definition

Progression-free survival is defined as the duration from the date of the first dose of erdafitinib until the date of first documented evidence of progressive disease or death due to any cause, whichever occurs first. Subjects who are progression-free and alive or have unknown status will be censored at the last disease evaluation. Subjects with no baseline or no post-baseline disease evaluation will be censored at the date of first dose. Subjects who had a disease progression or death event after having started a subsequent anti-cancer therapy will be censored at the last disease evaluation before the subsequent therapy. Subjects with death or progressive disease after more than one missing disease evaluations will be censored at the date of the last disease evaluation without documented progressive disease.

Analysis method

PFS based on both IRC assessment and investigator assessment will be presented with descriptive summary statistics; Kaplan-Meier method will be used for evaluation, if there are sufficient data.

5.3.6. Overall Survival (OS)

Definition

OS measured from the date of first dose of study drug to the date of the subject's death from any cause. If the subject is alive or the survival status is unknown, the subject's data will be censored at the date the subject was last known to be alive.

Analysis method

OS will be presented with descriptive summary statistics; Kaplan-Meier method will be used for evaluation, if there are sufficient data.

5.3.7. Patient-reported Outcome (PRO) endpoints

Definition

The patient-reported outcome (PRO) endpoints of interest include domain scales from the EORTC-QLQ-C30 (for subjects ≥18 years of age) or Peds FACT-Br (for subjects <18 years of age), PGIS, PGIC, and EQ-5D-5L (utility value and visual analog scale).

- **EORTC QLQ-C30**

The EORTC QLQ-C30 version 3.0 is a health-related quality of life instrument that has been widely used in cancer patients. It is composed of both multi-item scales and individual items which include: five functional scales, three symptom scales, a global health status/QoL scale, and six individual items. All of the scales and individual items are transformed to a range of 0 to 100 where a higher scale score represents a higher response level. Notably, a high score for a functional scale represents a high level of functioning and a high score for the global health status/QoL represents a high QoL. In contrast, a high score for a symptom scale/item represents a high level of symptomatology/problems.

Scale scores will be computed, per the developers' scoring manual provided responses are available to at least half of the component items. For all scales, the Raw Score (RS) is the mean of the non-missing component items. For the functional scales, Score = 100 x [1 - (RS - 1)/Range]. For the symptom scales/items and the global health status/QoL, Score = 100 x [(RS - 1)/Range]. A summary of the scales is given in Table 1 below.

EORTC QLQ-C30 Scales

Scale	Label	# of Items	Item Range	Item Numbers
Global Health Status/QoL				
Global Health Status/QoL	QL2	2	6	29, 30
Functional Scales				
Physical Functioning	PF2	5	3	1 to 5
Role Functioning	RF2	2	3	6, 7
Emotional Functioning	EF	4	3	21 to 24
Cognitive Functioning	CF	2	3	20, 25
Social Functioning	SF	2	3	26, 27
Symptom Scales				
Fatigue	FA	3	3	10, 12, 18
Nausea and Vomiting	NV	2	3	14, 15
Pain	PA	2	3	9, 19
Dyspnoea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite Loss	AP	1	3	13
Constipation	CO	1	3	16
Diarrhoea	DI	1	3	17
Financial Difficulties	FI	1	3	28

- **EQ-5D-5L**

The EQ-5D-5L is a generic, preference-based instrument for describing and valuing health states in two parts; the EQ-5D descriptive system and the visual analogue scale (VAS). The EQ-5D descriptive system consists of five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is scored using a 5-point Likert scale: (1) no problems, (2) slight problems, (3) moderate problems, (4) severe problems, and (5) extreme problems. While the individual numerals have no arithmetic properties, the digits for the 5 dimensions can be combined into a 5 digit number describing the respondent's health state. Janssen will apply the United Kingdom weights to generate patient utilities from the 5 dimensions of the EQ-5D-5L in this study. The VAS records the respondent's self-rated health state on a vertical thermometer with anchors of 0: Worst imaginable health state, and 100: Best imaginable health state.

- **Patient Global Impression of Severity**

The PGIS is a single-item questionnaire that asks the patient to rate their overall impression of their cancer symptoms. The question text and response options are as follows:

Overall, how would you rate the severity of your cancer symptoms currently?

1. None
2. Mild

3. Moderate
4. Severe
5. Very Severe

- **Patient Global Impression of Change**

The PGIC is a single-item questionnaire that asks the patient to rate their overall impression of the change in their cancer symptoms. The question text and response options are as follows:

Compared to when you received the first treatment in this study, how have your cancer symptoms changed?

1. A lot better now
2. Moderately better now
3. A little better now
4. Neither better, nor worse (no change)
5. A little worse now
6. Moderately worse now
7. A lot worse now

- **FACT-Br**

The Pediatric Functional Assessment of Cancer Therapy – Brain (peds FACT-Br) is available in 2 versions for children and adolescents. It has a recall period of the past 4 weeks. The response scales comprise a 5-point Likert-type scale. Subscale domains address physical well-being, emotional well-being & illness experience, social and family well-being and a brain cancer subscale. Estimated time for completion is approximately 10-15 minutes.

Analysis method

PRO endpoints will be summarized descriptively for the Broad Panel Cohort. Post-hoc or exploratory analyses will be discussed in a separate analysis plan if needed.

Key PRO endpoints:

- EORTC QLQ-C30 global health status/quality of life subscale
- EQ-5D-5L utility score
- EQ-5D-5L VAS

Compliance rates for completion of EORTC QLQ-C30 and EQ-5D-5L at each time point will be generated.

The change from baseline in PRO endpoints at each time point will be summarized with descriptive statistics (N, mean, SD, median, and range [minimum and maximum]). The mean and standard deviation of change from baseline over time will be graphically presented with line plot.

A distribution-based method will be used to define worsening/improvement in EORTC QLQ-C30, i.e., half standard deviation away from the mean score at baseline. Time to worsening and time to improvement will be derived. Time to worsening will be estimated using Kaplan-Meir methods. Death due to disease progression will be considered as worsening. Subjects who have not met the definition of worsening will be censored at the last PRO assessment. Subjects without baseline assessment or post-baseline assessment will be censored at date of treatment start date.

In addition, PGIC and PGIS will be analyzed with descriptive statistics. In addition, change from baseline in PGIS will be summarized with descriptive statistics over time.

Peds FACT-Br (for subjects <18 years of age) will be analyzed for the Pediatric Cohort at the time of final analysis for the Pediatric Cohort.

6. SAFETY

All safety analyses will be based on the safety analysis set, unless otherwise specified.

Safety will be analyzed using the incidence and severity of AEs, physical examinations, vital signs, electrocardiogram (ECG) measurements, clinical safety laboratory assessments, ophthalmologic examinations, performance status and Bone-related Assessments (for adolescents only).

For all continuous safety variables, descriptive statistics will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized using frequency counts and percentages.

6.1.1. Adverse Events

The verbatim terms used in the CRF by investigators to identify AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). These coded AE terms are referred to as preferred terms (PT); classification into System Organ Class (SOC) is a result of the coding process.

Category	**Analysis**	**Sorted By**	**Drug-Related TEAE**
General	Overall summary		✔
	TEAEs	SOC+ PT+ toxicity grade;	✔
	Serious TEAEs	SOC+ PT	✔
	Grade 3 or worse TEAE	SOC+ PT	✔
	TEAEs leading to treatment discontinuation	SOC+ PT	✔
	TEAEs leading to death	SOC+PT	✔

Category	Analysis	Sorted By	Drug-Related TEAE
	TEAEs leading to dose modification or interruption	SOC+ PT	✔
	AEs of clinical importance	SOC+PT + toxicity grade	
	AEs of special interest	SOC+ PT+ toxicity grade	
	Other safety observations (e.g. other malignancies, eye disorder)	PT + toxicity grade	
	Deaths within 30 days of last dose	Reason for death	

6.1.2. All Adverse Events

Treatment-emergent AEs (TEAEs) are defined as 1) those that first occur in TEAE period (defined as the time from first dose date through 30 days after last dose date, or day before subsequent anticancer therapy, whichever occurs first); 2) present before first dose, but worsened in toxicity grade during TEAE period; 3) had missing start date and its end date is during the TEAE period; 4) was a drug-related event. To determine TEAE, partially missing AE start dates will be imputed according to the rules stated in section 2.6.

Treatment-emergent AEs will be summarized by system organ class and preferred terms, by NCI toxicity grade, by relationship to erdafitinib, and by action taken.

The severity of AEs will be graded on a scale of 1 to 5 according to the adult NCI Common Terminology Criteria for Adverse Events (NCI-CTCAE version 5.0 or higher) where higher grades indicate events of higher severity.

For each TEAE, the percentage of subjects who experience at least 1 occurrence of the given event will be summarized. Tables will be sorted by frequency in incidence (the highest to lowest incidence).

For subjects using strong CYP3A inhibitor during treatment-emergent period, all TEAEs will be summarized by SOC, PT, maximum severity, and strong CYP3A inhibitor (Yes vs No). For subjects using strong CYP2C9 inhibitor during treatment-emergent period, all TEAEs will be summarized by SOC, PT, maximum severity, and strong CYP2C9 inhibitor (Yes vs No).

6.1.3. Adverse Events of Special Interest and Clinical Importance

Adverse events of special interest will be summarized based on the following criteria:

- Central serous retinopathy (CSR):
- Growth disorder (including accelerated bone growth and SCFE) and fractures in children and adolescent subjects (≥6 to <18 years of age) is an adverse event of special interest for erdafitinib.

Adverse events of clinical importance on the following criteria will be summarized:

- Eye Toxicity (excluding CSR): Dry eye, Xerophthalmia, Keratitis, Ulcerative keratitis, Conjunctivitis, Blepharitis, Cataract, Cataract subcapsular, Conjunctival haemorrhage, Conjunctival hyperaemia, Conjunctival irritation, Corneal erosion, Corneal infiltrates, Eye inflammation, Eye irritation, Eye pain, Foreign body sensation in eyes, Lacrimation increased, Night blindness, Ocular hyperaemia, Photophobia, Vision blurred, Visual acuity reduced, Visual impairment, Xanthopsia

- Hyperphosphataemia: Hyperphosphataemia

- Nail Toxicity: Nail bed bleeding, Nail discolouration, Nail disorder, Nail dystrophy, Nail ridging, Onychalgia, Onychoclasis, Onycholysis, Paronychia, Nail toxicity, Onychomadesis

- Skin Toxicity: Dry skin, Hyperkeratosis, Skin atrophy, Skin exfoliation, Skin fissures, Skin lesion, Xeroderma, Blister, Erythema, Palmar erythema, Palmar-plantar erythrodysaesthesia syndrome, Plantar erythema, Rash, Rash erythematous, Rash macular, Rash maculo-papular, Skin ulcer, Toxic skin eruption

- Gastrointestinal toxicity: Stomatitis, Dry mouth, Diarrhoea, Aptyalism, Mucosal inflammation

Note that the most up-to-date list of adverse events of special interest and adverse events of clinical importance will be used at the time of analysis.

Incidence of AEs of special interest will be summarized. The TEAE incidence rate by time to 1st onset will be calculated for the time periods 0 to 1 month, >1 month to 2 months, >2 months to 3 months and >3 months. Time to first onset of the TEAE will be summarized based on the observed TEAE values.

$$Incidence\ rate = \frac{Number\ of\ subjects\ with\ 1^{st}\ onset\ of\ TEAE}{Treated\ Subjects}\ \ in\ the\ period\ of\ time$$

Subjects who didn't have the TEAE will be censored, and the date of study completion/discontinuation will serve as the time of censoring.

6.1.4. Deaths

A summary of deaths during the treatment and up to 30 days after last dose will be provided, along with the primary cause of death. In particular, frequencies of deaths due to study treatment-related adverse events will also be reported. A death is study medication-related death if the primary cause is a drug related AE.

6.2. Clinical Laboratory Tests

Laboratory data of hematology and serum chemistry up to 30 days after last dose or the end of treatment visit date, whichever is later, will be reported in SI units.

Applicable laboratory results will be graded according to NCI-CTCAE version 5.0.

The following laboratory tests may be analyzed:

- Hematology: hemoglobin, platelet count, WBC count, ANC
- Chemistry: ALT, chloride, albumin, creatinine (including calculated creatinine clearance), alkaline phosphatase, magnesium, AST, bicarbonate, phosphate, blood glucose, potassium, sodium, total bilirubin, total protein, calcium,
- Parathyroid hormone (PTH)
- TSH (adolescents only): Thyroid stimulating hormone, free thyroxine (T4), and total triiodothyronine (T3)
- IGF-1 (adolescents only)

Descriptive statistics will be calculated for each laboratory analyte at baseline and at each scheduled time point. Changes from baseline results will be presented in pre- versus post-intervention cross-tabulations (with classes for below, within, and above normal ranges). Frequency tabulations of the abnormalities will be made. A listing of subjects with any laboratory results outside the reference ranges will be provided. Parameters with predefined NCI-CTCAE toxicity grades will be summarized. Change from baseline to the worst adverse event grade experienced by the subject during the study will be provided as shift tables.

Liver function abnormality by Hy's Law: For subjects with any elevated liver enzyme (AST or ALT) of >3xULN, ALP <2xULN, and associated with an increase in bilirubin (total) ≥2xULN, a listing for all subjects with all such records will be provided and a summary table of the number of such subjects will be provided.

6.2.1. Creatinine Clearance

Cockcroft-Gault Formula for Estimated Creatinine Clearance for adults shown below will be used to derive creatinine clearance for adults. For adolescents and children, no derivation will be conducted but using the collected creatinine clearance.

$$eCR = \frac{(140\text{-Age}) \times \text{Mass (Kilograms)} \times [0.85 \text{ if female}]}{72 \times \text{Serum Creatinine (in mg/dL}^{1})}$$

for males, the factor is 1 instead of 0.85.

OR

$$eCcr = \frac{(140\text{-Age}) \times \text{Mass (Kilograms)} \times \text{Constant}}{\text{Serum Creatinine (in } \mu\text{mol/L)}}$$

Where Constant = 1.23 for men and 1.04 for women

Reference: http://www.mdcalc.com/creatinine-clearance-cockcroft-gault-equation/

6.3. Ophthalmologic Examination

Summary statistics will be tabulated for the treated subjects whose results of ophthalmologic examinations were obtained from both before and after the study treatment. These summaries will include measures of change from baseline, frequency and percentage shifting from baseline, time to certain eye exam parameters, etc. Listings of treatment-emergent AEs related to the eye toxicity will also be generated.

6.4. Electrocardiogram

Electrocardiogram data will be summarized by ECG parameter. Descriptive statistics will be calculated at baseline and for observed values and changes from baseline at each scheduled time point. Frequency tabulations of the abnormalities will be made.

The ECG variables that will be analyzed are heart rate, PR interval, QRS interval, QT interval, and QT corrected according to Fridericia's formula (QTcF), i.e., QTcF = $QT/\sqrt[3]{RR}$. The Bazett formula will be used for children and adolescent subjects, i.e., QTcB = $QT/\sqrt{RR}$. Descriptive statistics of QTc intervals and changes from baseline will be summarized at each scheduled time point. The percentage of subjects with QTc interval >450 milliseconds, >480 milliseconds, or >500 milliseconds will be summarized, as will the percentage of subjects with QTc interval increases from baseline >30 milliseconds or >60 milliseconds.

Values outside the normal range will be flagged as follows.

- Observed:
    - Heart rate: L <50 bpm; H > 100 bpm
    - RR interval: L < 600 ms; H > 1200 ms
    - QT interval: H > 500 ms
    - QTc interval: H > (450 ms for males, 470 ms for females); increase to >500 ms
- Change from baseline:
    - QTc: 30-60 ms increase; increase >60 ms

All treatment-emergent abnormal findings will be tabulated, displaying the number of subjects with abnormal findings after dosing. An abnormal finding is considered to be treatment-emergent if it occurred during treatment and up to 30 days after the last dose.

6.5. Vital Signs and Physical Examination Findings

All clinically significant abnormalities will be recorded as AEs and summarized for safety.

6.6. Other Safety Measures

Frequencies of performance score will be reported over time. Descriptive statistics of change in performance scores from baseline will also be provided.

Descriptive statistics for bone-related assessments will be provided at baseline, at each scheduled visit and changes from baseline at each scheduled time point.

Pregnancy testing results will be presented in a by-subject listing.

7. PHARMACOKINETICS (PK)

Details of population PK analysis, analysis to evaluate relationships between plasma concentration or metrics of systemic exposure and CYP2C9 polymorphism, markers of pharmacological activities (serum phosphate), efficacy or treatment-emergent adverse events may be explored as data allow using population approaches. If relevant, these analyses will be presented in a separate plan and results will be presented in a separate report. All subjects and samples excluded from the analysis will be clearly documented in the separate study report.

Plasma and CSF pharmacokinetic data will be listed for all subjects with available plasma erdafitinib concentrations. Concentration values below the lower limit of quantitation (LLOQ) will be displayed in listings as zero with a flag and handled as zero. Plasma protein binding data such as alpha-1-acid glycoprotein and fraction unbound (unitless) of erdafitinib based on assessment on C2D1 2-4 hour postdose will be listed for all subjects.

For Pediatric Cohort, an interim Bayesian population PK analysis and, if appropriate, NCA analysis of plasma concentration-time data for erdafitinib and markers of pharmacological activities (e.g., serum phosphate) will be performed after 5 children or adolescents (<18 years of age) are enrolled and have completed Cycle 3 Day 1 visit in the study (see protocol Section 9.4.4). In addition, further interim analyses may be conducted after enrollment of additional subjects. Descriptive statistics, including arithmetic and geometric mean, median, standard deviation, coefficient of variation (CV), geometric CV, minimum and maximum, will be tabulated for erdafitinib PK concentrations and protein binding data by time point and age groups (children [6 to <12], adolescents [12 to <15 years of age, 15 to <18 years of age]). PK and protein binding summary table will be performed based on PK analysis set.

8. BIOMARKERS

The following analyses may be provided with a focus on the treated population in the Broad Panel Cohort.

- Summary and listing of baseline FGFR alterations (alteration types, genes, and variants) will be provided.

- Sensitivity efficacy analysis (see 5.2.3.3 for more details) will be performed excluding patients without an eligible FGFR alteration, with only known or suspected FGFR germline mutations, with only FGFR fusion with a 5-prime partner, with gatekeeper/resistance FGFR alterations, or with exclusionary co-alterations, as specified in the protocol (a corresponding listing of the excluded patients will be provided).

- A co-alteration profile will be provided by tumor histologies, and by FGFR alterations (alteration types, genes, and variants). Additional analyses such as response association may be conducted.

- Biomarker positive frequency based on central screening (with patients who were enrolled by central screening) results will be summarized by tumor histology.

Additional exploratory biomarker analyses may be performed based on the Exploratory Cohort.

9. HEALTH ECONOMICS

Medical resource utilization will be captured, and descriptive summary may be provided.

REFERENCES

1. Berry SM, Kristine RB, Groshen S, Berry DA. Bayesian Hierarchical modeling patient subpopulations: Efficient designs of Phase II oncology clinical trials. Clinical trials. 2013;10:720-734.
2. Chugh R, Wathen JK, Maki RG, et al. Phase II multicenter trial of imatinib in 10 histological subtypes of sarcoma using a Bayesian hierarchical statistical model. J Clin Oncol. 2009;27(19): 3148-3153.
3. ICH E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials, August 2017.
4. Spiessens B, Debois M. Adjusted significance levels for subgroup analysis in clinical trials. Contemporary Clin Trials. 2010;31:647-656.

APPENDIX

The appendix includes details about the BHM, simulation assumptions and design setups, operating characteristics for Broad Panel Cohort only and with multiplicity adjustment for Core Panel Cohort.

A.1. Bayesian Hierarchical Model (BHM)

Denote $\pi_g = \Pr(\text{response}|\text{treatment}, g)$, the BHM models the log-odds of response (i.e. ORR) defined as

$$\theta_g = logit(\pi_g) = \ln(\frac{\pi_g}{1-\pi_g})$$

with weak priors

$$\theta_g \sim N(\mu, \tau^2)$$
$$\mu \sim N(-2.19, 2^2)$$
$\tau^2 \sim$ Inverse Gamma distribution with mean 2 and weight 0.75.

The prior normal mean value of -2.19 for μ corresponds to the logit of 10%. The prior distribution of μ is shown in Figure A1a, and it corresponds to a log odds range of 0% to 99.7%. The parameter τ represents the degree of heterogeneity between the tumor histology groups, thus the degree of borrowing. At one extreme, when τ is 0, there is complete borrowing. At the other extreme, when τ is near infinity, there is no borrowing across the groups. For values between these two extremes, there is an amount of borrowing consistent with the variability across groups. The mean for τ is set to the same value as the prior standard deviation for μ, and the weight is set to be 0.75. The distribution of τ is shown in Figure A1b. It is to ensure certain degree of borrowing.

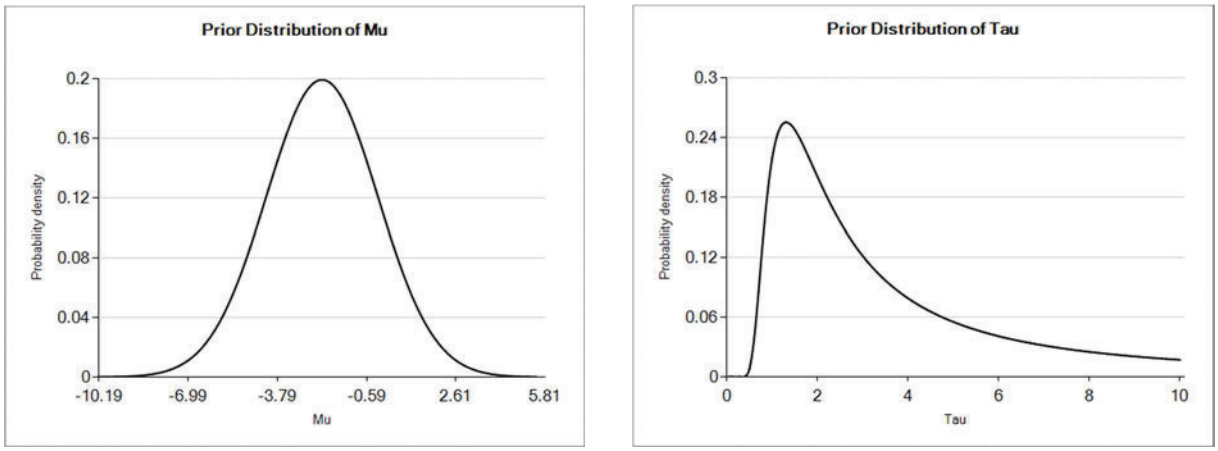


Figure A1: Prior Distribution of Hyperparameters of BHM. a) Prior Distribution of μ. b) Prior Distribution of τ

A.2. Simulation Assumptions and Setups

The simulation study in this section is based on the Broad Panel Cohort which enrolls patients with target FGFR mutations or any gene fusion, considering three protocol specific interim futility analyses and the primary analysis. Note that there will be an efficacy assessment at the second

interim, but it is not considered in the simulation for its ignorable impact with a use of an alpha of < 0.01% in the proposed early success rule. The simulation study has the following assumptions and parameter setups.

Assumptions on Prevalence

The alterations eligible for inclusion in the Broad Panel Cohort consist of target FGFR mutations and any FGFR fusion. The full list of target FGFR mutations eligible for the Broad Panel Cohort will be provided separately on the study portal and in the laboratory manual. Based on the predicted frequency of target FGFR alterations, prevalence of tumor histology, and predictive modeling of distribution; the tumor histologies that were selected for central molecular screening include: glioblastoma; squamous non-small cell lung cancer; squamous cell head and neck cancers; gastric cancer; breast cancer; cholangiocarcinoma; soft tissue sarcoma; cervical, and ovarian cancers; glioma; renal cell cancer; esophageal cancers; hepatocellular carcinoma; and pancreatic cancer (Table A1).

Table A1: Frequency of Preclinically Validated Target FGFR Mutations and Fusions in Advanced Cancer

Advanced Cancer Type	Target FGFR Frequency[a]	TOTAL G7, Drug Treated Patients L2+[b]	Anticipated tumor representation in CAN2002 in L2+ FGFR+ population plus 'Other' Group
Glioblastoma	~21%	16317	~18%
Squamous non-small cell lung cancer	~4%	85177	~18%
Squamous cell head and neck cancers	~9%	25315	~13%
Gastric cancer	~3%	78760	~13%
Breast cancer	~2%	86180	~10%
Cholangiocarcinoma	~11%	8245	~5%
Endometrial	~7%	10482	~4%
Glioma	~5%	13836	~4%
Ovarian cancer	~2%	31363	~3%
Soft tissue sarcoma	~8%	5485	~2%
Renal cell cancer	~1%	43431	~2%
Cervical cancer	~3%	5263	~1%
Esophageal cancer	~1%	22467	~1%
Hepatocellular carcinoma	~2%	10042	~1%
Pancreatic	<1%[c]	-	-
Other	-	-	~5%[d]

Note: a. FGFR mutations and fusions are from The Cancer Genome Atlas (TCGA) and Genomics Evidence Neoplasia Information Exchange (GENIE) databases accessed February 2019, and filtered for FGFR hotspot mutations and likely pathogenic FGFR fusions. Target refers to FGFR mutations and fusions predicted to be likely pathogenic based on genomic features.
b. G7=US, France, Germany, UK. Spain, Italy and Japan, and the treated patients of second and higher line include those are incident or newly recurrent across all these countries.
c. Based on clinical experience.
d. Assumed for simulation purpose

In the simulation studies, all selected central molecular screening tumor histologies are assumed for enrollment except pancreatic cancer due to its neglectable FGFR frequency (~0%) based on The Cancer Genome Atlas (TCGA) and Genomics Evidence Neoplasia Information Exchange (GENIE) databases. In addition, the trial allows enrollment of local molecular screening FGFR+ subjects and the tumor histologies that are not limited to the ones listed in Table A1. Thus a group named 'Other' is included in the analysis to take in all other tumor histologies not listed in Table A1. Considering the uncertainty of the composition in the group Other, it will not stop early for enrollment for interim futility while the same cap of 30 applies. Final futility assessment will be done on the group Other before being pooled for primary efficacy analysis.

The proportion of each tumor histology for enrollment into the trial is listed in the last column in Table A1. In particular, a 5% proportion is assumed for the group Other. Note that other operational factors such as trial countries, open sites, investigators, etc. are not considered into the prevalence assumption.

Assumptions on Accrual Rate

A peak accrual rate of 5 subjects/week with ramp up is assumed. Ramp up defines periods of simple linear increase in the mean recruitment rate from the start of trial accrual. The mean weekly accrual rate is shown in Figure A2. During simulation, a Poisson process was used to simulate the random arrival of subjects with the specified mean accrual rate.

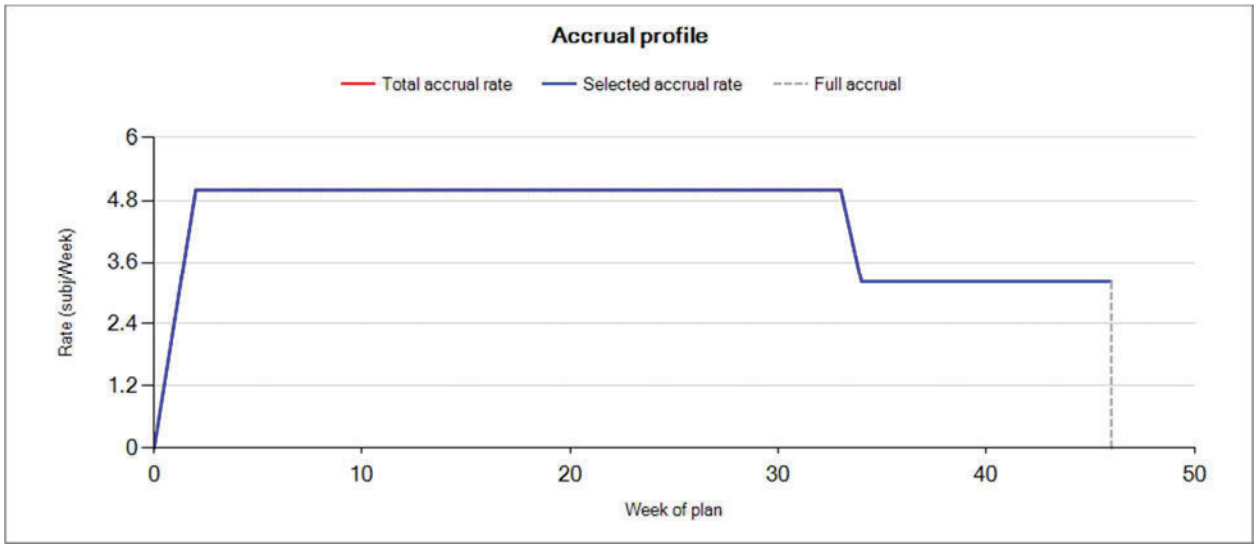


Figure A2: Mean Weekly Accrual Rate Over Time

Parameter Setups

- Sample size: 200 (Response-evaluable Population)
- Cap (Maximum sample size) on each tumor histology: 30
- Three Interim analyses at 60, 100, 140 response-evaluable subjects having completed a specified visit (with 6 weeks to final endpoint)

- Minimum number of subjects required before interim futility stop in a given histology group: 5
- Futility criteria: Less than a 40% chance that ORR > 25% (i.e. posterior Pr(ORR > 25%) < 40%)
- Success criteria at the primary analysis: Greater than an 80% chance that ORR > 25% (i.e. posterior Pr(ORR > 25%) > 80%)
- Follow-up of enrolled subjects is allowed after tumor histology group stopped enrollment due to futility
- No assumption on dropout

Note that the futility threshold 40% and the success threshold 80% were justified through extensive simulations to achieve good balance of design operating characteristics under various scenarios.

A.3. Operating Characteristics

The following simulations results are based on the Broad Panel Cohort only with no multiplicity adjustment. All scenarios were conducted using the FACTS v6.2 Enrichment Design – Dichotomous engine and the results from FACTS are post-processed in R.

A.3.1. Scenarios

As mentioned in section A.2., 14 tumor histologies and one Other group were assumed for enrollment. Below we use the shorthands: GBM (glioblastoma), Lung (squamous non-small cell lung cancer), HN (squamous cell head and neck cancers), Gastric (gastric cancer), Breast (breast cancer), Cholangio (Cholangiocarcinoma), Endometrial (endometrial cancer), Glioma (glioma), Ovarian (ovarian cancer), and STS (soft tissue sarcoma), RCC (renal cell cancer), Cervical (cervical cancer), Esophageal (Esophageal cancer), and Hepatocellular (hepatocellular cancer). Table A2 lists the simulation scenarios. The percentage number under each tumor histology is the proportion of the tumor histology in the trial assumed for enrollment (see Table A1 for details). The scenarios are based on the assumed true ORR for each tumor histology. Six scenarios are shown, with different true ORRs for each histology and the group Other. The last column of Table A2 shows the (prevalence) weighted average ORR as the overall truth assumed on the original tumor agnostic pool (the Broad Panel Cohort before possible futility on any tumor histologies).

More specifically, the simulation scenarios include:

- **Global Null**: the true ORRs for all tumor histologies and the group Other are 15% at the boundary of the null hypothesis (i.e., ORR ≤ 15%). Under this scenario, all tumor histologies and the original tumor agnostic pool are nonactive and should not succeed as a pool after futility analysis at the primary efficacy analysis.
- **Alternative**: the true ORRs for all tumor histologies and the group Other are 35%, at the boundary of the alternative hypothesis (i.e., ORR ≥ 35%). Under this scenario, all tumor

histologies and the original tumor agnostic pool are active and expected to succeed as a pool after futility analysis at the primary analysis.

- **Target**: the true ORRs for all tumor histologies and the group Other are 25%, the target ORR for BHM posterior probability. Under this scenario, all tumor histologies and the original tumor agnostic pool are active and expected to succeed as a pool after futility analysis at the primary analysis.
- **Half-Half Other35**: the most prevalent half of the tumor histologies are nonactive with ORR at 10%, the rest half are active with ORR at 35%, and the group Other is active with ORR at 35%. Under this scenario, the original tumor agnostic pool is nonactive with a weighted ORR of approximately 15%. However, the pool for the primary efficacy analysis with tumor histologies that didn't meet the futility criteria is possible to be active.
- **Half-Half Other15**: the most prevalent half of the tumor histologies are nonactive with ORR at 10%, the rest half are active with ORR at 35%, and the group Other is nonactive with ORR at 15%. Under this scenario, the original tumor agnostic pool is nonactive with a weighted ORR of 14%. However, the pool for the primary efficacy analysis with tumor histologies that didn't meet the futility criteria is possible to be active.
- **Cholangio35**: Cholangiocarcinoma is the only active tumor histology with ORR at 35%, GBM and Lung have ORR of 10%, and all others are with 15% ORR. Under this scenario, the pool is nonactive with weighted ORR 14%. However, the pool for the primary efficacy analysis with tumor histologies that didn't meet the futility criteria is possible to be active.

The Global Null scenario is to evaluate the false positive rate (FPR) for the primary pooled efficacy analysis and the family-wise error rate (FWER) with BHM analysis. The scenarios Alternative and Target are for power assessment. The scenarios Half-Half Other35, Half-Half Other15, and Cholangio35, with their original tumor agnostic pooled ORR at the null, are designed to assess the impact of futility and to evaluate the FPR for the primary pooled efficacy analysis. Note that scenarios with original pooled ORR > 15% can also be designed but are of less interest to be investigated for power assessment thus are not discussed here.

Table A2: Simulation Scenarios Based on ORR for Individual Tumor Histologies

Scenario	GBM	Lung	HN	Gastric	Breast	Cholangio	Endometrial	Glioma	Ovarian	STS	RCC	Cervical	Esophageal	Hepatocellular	Other	Pooled
Global Null	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Alternative	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**
Target	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**	**0.25**
Half-Half Other35	0.10	0.10	0.10	0.10	0.10	0.10	0.10	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	0.15
Half-Half Other15	0.10	0.10	0.10	0.10	0.10	0.10	0.10	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	**0.35**	0.15	0.14
Cholangio35	0.10	0.10	0.15	0.15	0.15	0.35	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.14

Note: Bold indicates the active (ORR>15%) tumor histologies; no bold represents the nonactive (ORR≤15%).

A.3.2. Definitions of Operating Characteristics

The following terminology will be used to discuss operating characteristics of the BHM analysis.

Selection Power

Selection power is the proportion of simulations that selected at least a given percentage of active tumor histologies (e.g. 80%) to be pooled for the primary efficacy analysis.

Proportion of Success

Proportion of success is the proportion of simulations in which BHM claimed at least one tumor histology as a success. Similarly, the success criteria (Pr(ORR > 25%) > 80%) is used for BHM to claim a success (efficacy) for a given tumor histology.

Family-wise Error Rate (FWER)

FWER represents the proportion of simulations in which BHM claimed false-positive efficacy in one or more of the nonactive tumor histologies. Note that the success criteria (Pr(ORR > 25%) > 80%) is used for BHM to claim a success (efficacy) for a given tumor histology.

The following terminology is used to discuss the exact binomial test in the pooled primary efficacy analysis.

Power

Power is calculated as the proportion of simulations claiming a success (efficacy) on the pooled data of primary efficacy analysis.

False Positive Rate (FPR)

FPR is defined as the proportion of simulations claiming a false-positive success (efficacy) on the pooled data of the primary efficacy analysis. Note that a success is false-positive only when the weighted average of true ORRs is less than 15% for the remaining histologies that didn't meet the futility criteria in the pool for primary efficacy analysis. The weights are the renormalized assumed prevalence in Table A1 for each tumor histology that ended up in the pool.

Mixing Nonactive

Mixing nonactive is defined as the proportion of claiming a success (efficacy) in the pooled data at the primary efficacy analysis that is mixed with at least one nonactive histology.

A.3.3. Operating Characteristics Results

20000 simulation runs were conducted to evaluate the operating characteristics of the proposed design.

Selection Power under The Alternative Scenario (Sample Size Determination)

The selection power under the Alternative scenario was used for sample size determination. As mentioned in section 1.4., the sample size of 200 response-evaluable subjects for the Broad Panel Cohort is justified based on extensive simulations with various Bayesian hierarchical model (BHM) designs to achieve approximately 80% power to select at least 80% tumor histologies (e.g., ≥ 12 of 15 tumor histologies) in the Alternative scenario (true ORR at 35% for all histologies). The selection power for Alternative scenario under the proposed design is shown in Figure A3. The probability of selecting more than 12 active histologies for the pooled primary analysis is approximately 83%; the probability of selecting at least 10 active histologies is approximately 95%; the probability of selecting all 15 active histologies groups dropped to 29.4%.

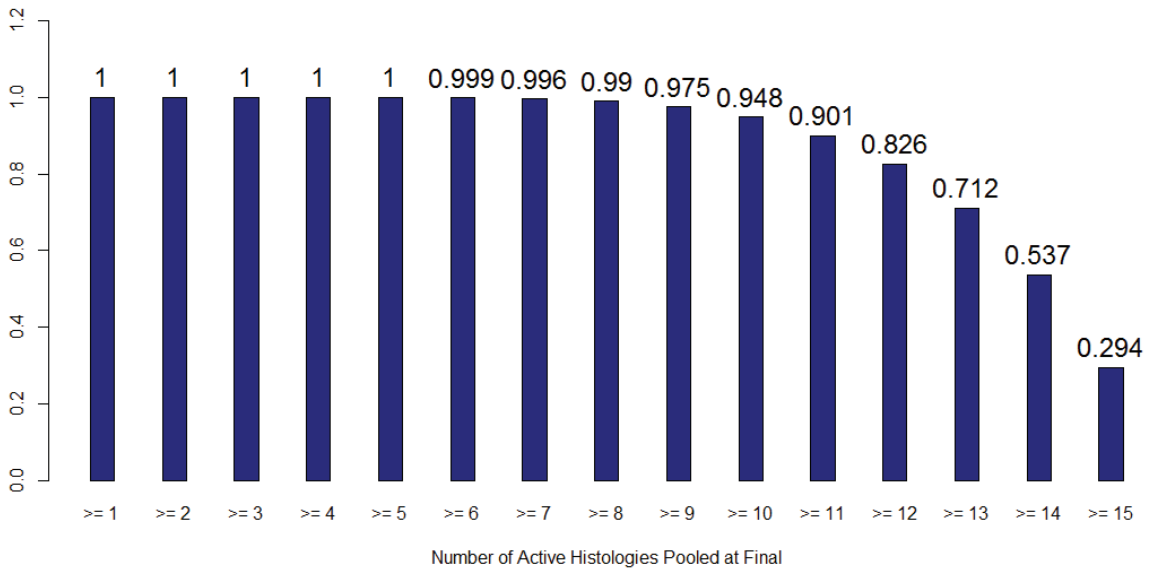


Figure A3: Selection Power for Alternative Scenario

Power and Proportion of Success

The power based on pooled efficacy analysis and the proportion of success based on BHM are listed in Table A3 for all scenarios. Note that the proportion of success and the power in the first row of Global Null represent the FWER and FPR and will be discussed the next. The design is achieving above 99% power and proportion of success for scenario Alternative; For scenario Target, it achieves above 95% while has a moderate proportion of success of 70%; The scenario Half-Half Other35 achieves 99% power and 94% proportion of success; The power and proportion of success for the scenario Half-Half Other15 are approximately 98% and 87%; The power and proportion of success for scenario Cholangio35 are lower than 50%, note that if the ORR for Cholangio is 50% with other ORRs remaining the same, the power and proportion of success will increase to approximately 83% and 77% respectively (not listed as one scenario).

Duration of the Trial

The expected trial duration listed in Table A3 varies among different scenarios mainly based on whether the relatively high prevalent tumor histologies are active or not. In the Alternative scenario when all tumor histologies are active, it is unlikely that the high prevalent tumor histologies will stop enrolment early, thus these histologies reach the cap easily and the maximum sample size of 200 is fulfilled in a shorter duration. Similar to the Alternative scenario, the Target scenario has a relatively short duration with all active tumor histologies. In comparison, for the other three scenarios, tumor histologies with high prevalence are assumed nonactive thus likely to be stopped early. Since the remaining low prevalent tumor histologies will enroll subjects at a much slower rate, longer trial durations will result. Note that the duration numbers (in weeks) are for comparison among scenarios to study the impact of early futility on accrual, and are not intended to be a definitive reference on actual trial duration.

FWER and FPR

The FPR based on pooled efficacy analysis and FWER based on BHM are listed in Table A4 for all scenarios.

- **FWER**

The FWERs for all scenarios are well controlled at 5% level. Specifically, the FWERs are 3.53% under Global Null, 0.24% under Half-Half Other35, 0.60% under Half-Half Other15, and 4.15% under Cholangio35.

As mentioned in section 5.2.3.2., the posterior probability of ORR > 25% for each individual tumor histology will be evaluated as a supplementary analysis. In the simulation results, the posterior probability of ORR > 25% is not reported. Instead, the probability of declaring efficacy on each individual tumor histology is listed in Table A4 (Columns 2-11) based on success criteria (Pr(ORR > 25%) > 80%) as an evaluation of the BHM as well as the possible supplementary analysis of the study.

- **FPR and Mixing Nonactive**

It is shown that the FPR for pooled analysis is around 18.57% for Global Null scenario. Among the false positive cases, there are 16.34% false positive cases that include group Other that did not have interim futility analyses, and 36.7% false positive cases that include at least 1 group with less than 5 subjects. For the Half-Half Other35 scenario, the FPR is controlled at 0.08%. For the Half-Half Other15 scenario, the FPR is controlled at 0.71%. For the Cholangio35 scenario, FPR is at 7.42%. The reasonable FPR control on scenarios Half-Half Other35, Half-Half Other15, and Cholangio35 are aided by the interim and final futility analysis before pooled efficacy analysis, thus helping to prune original tumor agnostic pool with ORR < 15% to the final tumor pool with ORR > 15% by excluding those tumor histologies that met the futility criteria from final pooling.

Another aspect to consider is the proportion of mixing nonactive tumor histologies for success on pooled data. For the scenarios Half-Half Other35, Half-Half Other15 and Cholangio35, the proportions of mixing nonactive are 19.8%, 23.5%, and 36.1% respectively. Noticing that theses scenarios have a considerable number of nonactive tumor histologies, the pooled efficacy analysis may not be performed at the primary analysis if a large amount (e.g. 50%) of tumor histologies were stopped for enrollment.

Note that the FPR for scenario Global Null is inflated as 18.57% and the scenario Alternative is overpowered with power 1, different alpha levels for testing are investigated in Table A5. An alpha level of 0.005 will lower the FPR under 10% while maintaining strong power for both Alternative and Target scenarios. An alpha level of 0.001 will lower the FPR under 5% while moderately sacrificing the power for scenario Target.

Summary

Based on the operating characteristics of different scenarios, the proposed design outlined in this study will have a control of FPR for Half-Half Other35 (0.08%) and Half-Half Other15 (0.71%), and a moderate control of FPR for Cholangio35 at 7.42%, while the FPR for the Global Null is 18.57%. In addition, it achieves a satisfactory performance in terms of testing power (e.g., above 95% power and proportion of success for scenario Alternative) and selection power (approximately 85% power to select at least 80% tumor histologies for scenario Alternative).

Table A3: Operating Characteristics for Duration of the Trial, Ppn of Success, and Power

Scenario	Expected Duration of the Trial (weeks)	Ppn of Success	Power
Global Null	251	0.0353	0.1857
Alternative	58	0.9982	1
Target	100	0.7075	0.9763
Half-Half Other35	157	0.9352	0.9966
Half-Half Other15	163	0.8660	0.9816
Cholangio35	219	0.3206	0.4974

Note: Proportion of success is the proportion of simulations in which BHM claimed at least one tumor histology as a success.
Power is calculated as the proportion of simulations claiming a success (efficacy) on the pooled data of the primary efficacy analysis.

Table A4: Operating Characteristics for FWER and FPR

Scenario	Probability Erda Will be Declared Successful in Each Histology by BHM																Pooled Analysis	
	GBM	Lung	HN	Gastric	Breast	Cholangio	Endometrial	Glioma	Ovarian	STS	RCC	Cervical	Esophageal	Hepatocellular	Other	FWER	FPR	Mixing Nonactive
Global Null	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.004	0.003	0.004	0.004	0.003	0.003	0.0353	0.1857	-
Alternative	**0.608**	**0.611**	**0.580**	**0.572**	**0.547**	**0.477**	**0.450**	**0.451**	**0.415**	**0.364**	**0.359**	**0.285**	**0.280**	**0.281**	**0.500**	-	-	-
Target	**0.116**	**0.121**	**0.108**	**0.107**	**0.103**	**0.101**	**0.104**	**0.101**	**0.097**	**0.086**	**0.083**	**0.060**	**0.057**	**0.060**	**0.121**	-	-	-
Half-Half Other35	0.000	0.000	0.000	0.000	0.000	0.000	0.001	**0.378**	**0.351**	**0.313**	**0.310**	**0.219**	**0.215**	**0.217**	**0.507**	0.0024	0.0008	0.1976
Half-Half Other15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	**0.360**	**0.334**	**0.295**	**0.290**	**0.200**	**0.202**	**0.200**	0.004	0.0060	0.0071	0.2351
Cholangio35	0.000	0.000	0.002	0.002	0.002	**0.299**	0.003	0.003	0.003	0.004	0.005	0.006	0.005	0.005	0.003	0.0415	0.0742	0.3606

Note: the bold numbers represent the power to declare efficacy in a given active tumor histology; no bold numbers represent the false-positive rate to declare efficacy in a nonactive tumor histology.
FWER represents the proportion of simulations in which BHM claimed false-positive efficacy in one or more of the nonactive tumor histologies.
FPR is defined as the proportion of simulations claiming a false-positive success (efficacy) on the pooled data of the primary efficacy analysis.

Table A5: FPR and Power across Different Alpha Level

Alpha	FPR under Scenario Global Null	Power under Scenario Alternative	Power under Scenario Target
0.025	0.1857	1	0.9763
0.02	0.1635	1	0.9721
0.015	0.1444	1	0.9678
0.01	0.1225	1	0.9604
0.005	0.0850	1	0.9456
0.001	0.0373	1	0.8878

A.4. SEPARATE EFFICACY ANALYSIS FOR JAPAN SUBMISSION

Based on the discussion with PMDA, for Japan submission, different endpoints and analysis are specified. The timing of interim, primary and final efficacy analysis is the same with the interim, primary and final analysis of global study.

Primary analysis

Regarding the primary endpoint for Japan submission, the same definition of ORR assessed by IRC will be used as described in Section 5.2.1, but will only include ORR as per RECIST in subjects with solid tumors. ORR according to RANO in subjects with primary CNS tumors is defined as secondary endpoint. Efficacy analysis will be performed for FGFR gene mutation and fusion separately.

The primary analysis will be performed on the pooled data of all tumor histologies that did not meet the futility for the Treated Population in the Broad Panel Cohort as same approach described in section 5.1.3. The futility decision for each tumor type will be made based on the global analysis plan.

The population as the component of estimand is “subjects with advanced solid tumors and target FGFR mutation” for FGFR mutation subgroup and “subjects with advanced solid tumors and target FGFR fusion” for FGFR fusion subgroup. Other components of estimand are the same as Section 5.2.2

The primary endpoint will be calculated with 95% 2-sided exact CI. If the confidence interval excludes 15%, this study will be considered success. Multiplicity will not be adjusted between FGFR mutation and fusion since these are independent and target for different indications. Statistical inference for Core Panel Cohort will not be performed for Japan.

Supplementary analysis

The same supplementary analysis in 5.2.3.2 will be performed for subjects with solid tumors other than primary CNS tumors and will be performed for FGFR mutation and fusion separately.

Major Secondary analysis

For subjects with primary CNS tumors whose ORR is assessed by RANO, the ORR and its 95% 2-sided exact CI will be calculated.

Same as ORR assessed by IRC, DCR, CBR, DoR and PFS will be analyzed in the same manner with Section 5.3.1, 5.3.2, 5.3.3, 5.3.4 and 5.3.5 except the analysis will be performed for the following subgroups; 1. subjects with FGFR mutation and solid tumors whose ORR is assessed as per RECIST, 2. subjects with FGFR fusion and solid tumors whose ORR is assessed as per RECIST, 3. subjects with FGFR mutation and primary CNS tumors whose ORR is assessed as per RANO, 4. subjects with FGFR fusion and primary CNS tumors whose ORR is assessed as per RANO).

OS will be analyzed in the same manner with Section 5.3.6 except the analysis is performed by FGFR mutation and fusion separately.

Pediatric cohort will be analyzed in the same manner as Broad Panel Cohort.