



Protocol C4891026

An Interventional Safety and Efficacy Phase 1b/2, Open-Label Study to Investigate Tolerability, PK, and Antitumor Activity of Vepdegestrant (ARV-471/PF-07850327), an Oral Proteolysis Targeting Chimera, in Combination with PF-07220060 in Participants Aged 18 Years and Older with ER+/HER2 Advanced or Metastatic Breast Cancer

Statistical Analysis Plan (SAP)

Version: 1

Date: January 30, 2024

TABLE OF CONTENTS

LIST OF TABLES	5
1. VERSION HISTORY	7
2. INTRODUCTION	7
2.1. Study Objectives, Endpoints, and Estimands.....	7
2.1.1. Primary Estimands.....	10
2.2. Study Design	11
2.2.1. Phase 1b	11
2.2.2. Phase 2	13
3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS	13
3.1. Primary Endpoints.....	13
3.1.1. Phase 1b	13
3.1.2. Phase 2	13
3.2. Secondary Endpoints.....	14
3.2.1. Safety endpoints.....	14
3.2.2. Efficacy endpoints	14
3.3. Pharmacokinetic Endpoints.....	14
3.4. Biomarker Endpoints.....	15
3.5. Exploratory Endpoints.....	15
3.5.1. Patient-Reported Outcomes (Phase 2 Only)	15
3.5.2. Biomarker Endpoints	16
3.6. Baseline Variables.....	16
3.6.1. Study drug, study treatment and baseline definitions.....	16
3.6.2. Baseline characteristics.....	17
3.7. Safety Endpoints	17
3.7.1. Adverse Events	17
3.7.2. Laboratory Data	17
4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS).....	17
5. GENERAL METHODOLOGY AND CONVENTIONS.....	18
5.1. Hypotheses and Decision Rules	18
5.1.1. Hypotheses and sample size determination	18

5.1.2. Decision rules	20
5.2. General Methods	20
5.2.1. Data Handling after the cut-off date	21
5.2.2. Pooling of centers	21
5.2.3. Presentation of continuous and qualitative variables.....	21
5.2.4. Definition of study day	22
5.2.5. Definition of start of new anti-cancer therapy	22
5.2.6. Definition of on-treatment period.....	23
5.2.7. Standard derivations and reporting conventions.....	23
5.2.8. Unscheduled visits	23
5.2.9. Adequate baseline tumor assessment.....	23
5.2.10. Adequate post-baseline tumor assessment.....	24
5.3. Methods to Manage Missing Data	24
5.3.1. Missing data.....	24
5.3.1.1. Pharmacokinetic concentrations.....	24
5.3.1.2. Pharmacokinetic parameters	25
5.3.2. Handling of incomplete dates	25
5.3.2.1. Missing Dates	25
5.3.2.2. Adverse events	25
5.3.2.3. Prior and concomitant medications	27
5.3.2.4. Exposure.....	27
5.3.3. Imputation rules for date of last contact and efficacy assessments	28
5.3.3.1. Date of last contact.....	28
5.3.3.2. Death date.....	28
5.3.3.3. Tumor assessments.....	28
5.3.3.4. Date of start of new anti-cancer therapy	29
6. ANALYSES AND SUMMARIES	30
6.1. Primary Endpoint(s)	30
6.1.1. DLT for Phase 1b.....	30
6.1.1.1. Primary Analysis	30
6.1.2. Phase 2	30
6.1.2.1. Objective Response	30

6.1.2.2. Primary Analysis	30
6.2. Secondary Endpoint(s)	30
6.2.1. Safety endpoints.....	30
6.2.2. Efficacy endpoints	31
6.2.2.1. Objective response as assessed by derived investigator assessment per RECIST V1.1.....	31
6.2.2.2. Tumor shrinkage from baseline.....	32
6.2.2.3. Clinical benefit response	33
6.2.2.4. Duration of response	33
6.2.2.5. Progression-free survival.....	35
6.2.2.6. Overall Survival	38
6.2.3. Pharmacokinetic endpoints.....	39
6.2.4. Population pharmacokinetic endpoints.....	40
6.2.5. Biomarker endpoints.....	40
6.3. Other Endpoints: Patient-reported outcomes	40
6.4. Subset Analyses.....	41
6.5. Baseline and Other Summaries and Analyses.....	41
6.5.1. Baseline Summaries.....	41
6.5.1.1. Demographic characteristics	41
6.5.1.2. Medical history.....	42
6.5.1.3. Disease characteristics.....	42
6.5.1.4. Prior anti-cancer therapies.....	43
6.5.2. Study Conduct and Participant Disposition.....	43
6.5.2.1. Participant disposition.....	43
6.5.2.2. Protocol deviations.....	44
6.5.3. Study Treatment Exposure	44
6.5.3.1. Exposure to vepdegestrant	44
6.5.3.2. Exposure to PF-07220060.....	45
6.5.3.3. Dose reductions	46
6.5.3.4. Dose interruptions	46
6.5.4. Concomitant Medications and Nondrug Treatments.....	46
6.5.5. Subsequent anti-cancer therapies.....	47

6.6. Safety Summaries and Analyses	47
6.6.1. Adverse Events	47
6.6.1.1. All adverse events	48
6.6.1.2. Adverse events leading to dose reduction	49
6.6.1.3. Adverse events leading to discontinuation of study treatment	49
6.6.2. Deaths	49
6.6.3. Serious adverse events	50
6.6.4. Laboratory Data	50
6.6.4.1. Hematology and chemistry parameters	50
6.6.4.2. Other laboratory parameters	53
6.6.5. Electrocardiograms	53
6.6.6. Vital Signs	54
7. INTERIM ANALYSES	54
8. REFERENCES	54
9. APPENDICES	56

LIST OF TABLES

Table 1.	Summary of Changes	7
Table 2.	Provisional Dose Levels	12
Table 3.	Analyses Sets	18
Table 4	Exact 95% CIs for ORR for 25 participants	19
Table 5.	Summary of Tabulations	20
Table 6	Outcome and Event Dates for DoR Analyses	34
Table 7	DoR Censoring Reasons and Hierarchy	35
Table 8	Outcome and Event Dates for PFS Analyses	36
Table 9	PFS Censoring Reasons and Hierarchy	37
Table 10	OS Censoring Reasons and Hierarchy	38
Table 11	Summary Statistics for PK Parameters	39

APPENDICES

Appendix 1. Data Derivation Details56

Appendix 2. Statistical Methodology Details – SAS code.....56

Appendix 3 Statistical methodology details – Phase 1b dose determination rules58

Appendix 4 List of Abbreviations60

Appendix 5 Response Evaluation Criteria in Solid Tumors: RECIST version 1.162

1. VERSION HISTORY

Table 1. Summary of Changes

Version/ Date	Associated Protocol Amendment	Rationale	Specific Changes
30Jan2024	Protocol Amendment 1 20Nov2023	N/A	N/A

2. INTRODUCTION

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C4891026. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

Statistical analyses will be performed using cleaned eCRF data as well as non-CRF data such as pharmacokinetic (PK) concentration and biomarker data. The primary analysis will include all data up to a cut-off date when the last evaluable participant for the primary endpoint of Phase 2 has been on treatment for at least 1 year (in the absence of disease progression) or reported disease progression (whichever occurs first), or at sponsor discretion if the data support ending the study.

Additional analyses of the data may be performed for publication or regulatory reporting purposes.

2.1. Study Objectives, Endpoints, and Estimands

Phase 1b

Objectives	Endpoints	Estimands
Primary:	Primary:	Primary:
To assess safety and tolerability of vepdegestrant in combination with PF-07220060 in participants with ER+/HER2- A/MBC to select up to 2 RDEs for the combination.	DLTs during DLT observation period (Cycle 1).	DLT rate estimated based on data from DLT-evaluable participants during the DLT observation period (Cycle 1).
Secondary:	Secondary:	Secondary:
To evaluate the overall safety profile.	AEs as characterized by type, frequency, intensity (as graded by NCI CTCAE version 5.0), seriousness, and relationship to	Not applicable.

Objectives	Endpoints	Estimands
	vepedegestrant in combination with PF-07220060.	
To evaluate antitumor activity of vepdegestrant in combination with PF-07220060.	- Confirmed OR (CR or PR) by investigator assessment. - DoR by investigator assessment. - CBR (confirmed CR or PR at any time, or SD $\geq$ 24 weeks) by investigator assessment. - PFS by investigator assessment.	Not applicable.
To evaluate the plasma exposure of vepdegestrant, ARV-473, and PF-07220060 when vepdegestrant and PF-07220060 are given in combination.	- Plasma concentrations of vepdegestrant, ARV-473, and PF-07220060. - Steady-state C_{max}, T_{max}, and AUC_{last} AUC_{tau} and of vepdegestrant, ARV-473, and PF-07220060. - If data permit and as appropriate CL/F, V_z/F, and $t_{1/2}$ of vepdegestrant, ARV-473, and PF-07220060.	Not applicable.

Phase 2

Objectives	Endpoints	Estimands
Primary:	Primary:	Primary:
To assess the clinical antitumor activity of vepdegestrant in combination with PF-07220060.	Confirmed OR (CR or PR) determined by investigator assessment.	The treatment effect of vepdegestrant in combination with PF-07220060 assessed by the OR rate based on Investigator assessment per RECIST v1.1. in the analysis population. Participants who do not have a post-baseline disease assessment due to early PD, who receive anticancer therapies other than the study intervention prior to achieving an objective response, or who die, experience documented PD, or stop disease assessments for any reason prior to achieving an objective response will be counted as non-responders.

Objectives	Endpoints	Estimands
Secondary:	Secondary:	Secondary:
To determine additional antitumor activity outcomes of vepdegestrant in combination with PF-07220060.	- DoR by investigator assessment. - CBR (confirmed CR or PR at any time or SD $\geq$ 24 weeks) by investigator assessment. - PFS by investigator assessment. - OS.	Not applicable.
To further characterize the overall safety profile and tolerability of vepdegestrant in combination with PF-07220060.	AEs as characterized by type, frequency, intensity (as graded by NCI CTCAE version 5.0), seriousness, and relationship to vepdegestrant in combination with PF-07220060.	Not applicable.
To evaluate the plasma exposure of vepdegestrant, ARV-473 and PF-07220060 when vepdegestrant and PF-07220060 are given in combination.	Plasma concentrations of vepdegestrant, ARV-473, and PF-07220060.	Not applicable.
To assess changes from baseline levels in plasma ctDNA with treatment.	ctDNA plasma quantitative changes from pre-treatment to evaluate potential predictability of their associations with clinical outcomes.	Not applicable.

090177e19fe3e49a\Final\Final On: 08-Feb-2024 19:23 (GMT)

2.1.1. Primary Estimands

Phase 1b

Primary Estimand: DLT (dose limiting toxicity) rate estimated based on data from DLT-evaluable participants during the DLT observation period. The estimand has the following 5 attributes:

- **Treatment:** Vepdegestrant (ARV-471/PF-07850327) plus PF-07220060
- **Population:** DLT-evaluable participants with ER+/HER2- A/MBC, defined by the inclusion and exclusion criteria to reflect the targeted population of the treatment, who receive at least 1 dose of the study interventions in Phase 1b and either experience DLT(s) during the DLT observation period or complete the DLT observation period without DLT. Participants without DLTs who receive less than CC% of the planned dose of each study drug during DLT observation period for reasons other than treatment-related toxicity are not evaluable for DLT. All participants deemed non-evaluable for DLT may be replaced.
- **Variable:** Occurrence of DLTs during the DLT observation period.
- **Intercurrent events:**
 - Discontinue treatment during the DLT observation period without DLT
 - Taking less than CC% of study drug during the DLT observation period
- **Population-level summary:** DLT rate defined as the number of DLT-evaluable participants with DLTs in the DLT observation period divided by the number of DLT-evaluable participants in the DLT observation period.

Phase 2

Primary Estimand: The treatment effect of Vepdegestrant (ARV-471/PF-07850327) in combination with PF-07220060 on ORR per RECIST v1.1. response criteria as determined by derived investigator assessment. The estimand has the following 5 attributes:

- **Treatment:** Vepdegestrant (ARV-471/PF-07850327) plus PF-07220060
- **Population:** participants with ER+/HER2- A/MBC, as defined by the inclusion and exclusion criteria to reflect the targeted population of the treatment, who received at least 1 dose of study intervention.
- **Variable:** OR is defined as confirmed CR or PR per RECIST v1.1. response criteria as determined by derived investigator assessment, from the date of first dose of treatment combination until the first documentation of progressive disease (PD), death, or start of new anticancer therapy. Both CR and PR must be confirmed by repeat assessments performed no fewer than 28 days after the criteria for response are first met.
- **Intercurrent events:** All data collected after an intercurrent event of subsequent anticancer therapy will be excluded. All response assessments regardless of gaps in disease assessments will be considered. Participants who do not have a post-baseline disease assessment due to early PD, who receive anticancer therapies other than the

study intervention prior to achieving an objective response, or who die, experience documented PD, or stop disease assessments for any reason prior to achieving an objective response will be counted as non responders.

- **Population-level summary:** overall response rate (ORR) defined as the proportion of participants in the analysis population with an OR and 2-sided 95% CI for ORR.

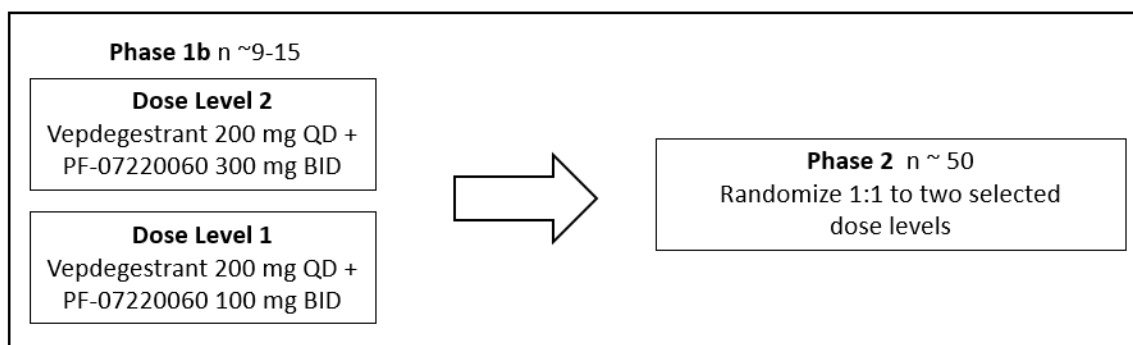
2.2. Study Design

C4891026 is a prospective, open-label, multicenter, Phase 1b/2 study to evaluate the safety, tolerability, antitumor activity, and PK of vepdegestrant (ARV-471) with PF-07220060 in the treatment of participants with ER+/HER2- A/MBC.

The combination will be assessed in 2 parts in the C4891026 study:

- **Phase 1b** will evaluate the safety and tolerability of vepdegestrant in combination with PF-07220060 and select up to two recommended doses for dose expansion for the combination (combination RDEs) via the mTPI-2 dose-finding design.
- **Phase 2** If at least 2 dose levels from Phase 1b are considered tolerable, a randomized dose optimization Phase 2 will be conducted to further evaluate the safety, tolerability, PK and antitumor activity of vepdegestrant in combination with PF-07220060 to determine the RP2D of the combination. If only one dose level from Phase 1b is considered tolerable, a single arm Phase 2 dose expansion will be conducted.

Schema



2.2.1. Phase 1b

A participant is classified as DLT-evaluable if they either 1) receive at least cc% of the planned doses of vepdegestrant and PF-07220060 and have received all scheduled significant safety assessments during the DLT observation period or 2) have experienced a DLT. The percentage of drug received is calculated as the percentage of the actual total amount of

drugs administered during the DLT observation period in relation to the planned total amount of drugs to be administered for the DLT observation period. If a participant fails to meet these criteria, they may be replaced. The sole exception to criterion (1) is if the participant has missed a minority of safety assessments due to emergency situations (eg, site accessibility issues, inability to go to an external laboratory). In such cases, the sponsor and investigators will have regular teleconferences to evaluate the participant evaluability for DLT, depending on the abundance of the available data.

For dose-finding, the DLT observation period will be during the first cycle of treatment in each participant. Participants who withdraw for any reason after Cycle 1 will not be replaced.

In Phase 1b the starting dose level of vepdegestrant will be 200 mg QD and the starting dose level of PF-07220060 will be 100 mg BID. Up to 2 starting dose levels may be investigated for vepdegestrant (100 mg QD and 200 mg QD), and for PF-07220060 (100 mg BID and 300 mg BID). The decision to escalate the starting dose level of vepdegestrant and PF-07220060 will be using mTPI-2 decision criteria based on the number of DLT-evaluable participants and the number of DLTs in those participants during the DLT observation period. The provisional dose levels to be evaluated are listed in Table 2. Alternate doses may be evaluated based on clinical findings.

Table 2. Provisional Dose Levels

Dose level	Vepdegestrant	PF-07220060
Dose Level 1	200 mg QD	100 mg BID
Dose Level 2	200 mg QD	300 mg BID

In the Phase 1b, based on the mTPI-2 decision criteria and depending on the occurrence of DLTs and need for dose reduction, two additional dose levels may be investigated: vepdegestrant CCI mg QD in combination with PF-07220060 CCI mg BID and/or vepdegestrant CCI mg QD in combination with PF-07220060 CCI mg BID.

A minimum of 3 DLT-evaluable participants will be required in each cohort at a given dose level before a decision is made of de-escalating, escalating, or staying at the same dose level. Additional participants will be enrolled in a specific cohort to replace participants who are not considered DLT-evaluable, where required. When all DLT-evaluable participants treated in a cohort have completed the DLT observation period or experienced a DLT, whichever occurs first, the posterior probabilities for the risk of DLT for new participants at different dose levels for the combination of interest will be evaluated. If the toxicity rate of the currently used dose level is far smaller than the target probability of toxicity (pT) of 0.30, the mTPI-2 will recommend escalating the dose level; if it is close to pT, the mTPI-2 will recommend continuing at the current dose; if it is far greater than pT, the mTPI-2 will recommend de-escalating the dose level. If the current dose is the highest dose and the rule indicates dose-escalation, new participants will be treated at the highest dose. If the current dose is the lowest dose and the rule indicates dose de-escalation, new participants will be treated at the lowest dose unless the number of DLTs reaches the elimination boundary, at which point the study would either be amended to add a lower dose level or terminated for safety. The mTPI-2 dose escalation/de-escalation recommendation will be based on the

estimated toxicity rate and 3 intervals (underdosing, target toxicity, and excessive toxicity) as shown below:

- If current dose is in the underdosing [CCI] toxicity interval: escalate to next higher dose;
- If current dose is in the target toxicity [CCI] toxicity interval: stay at current dose;
- If current dose is in the excessive toxicity or overdosing [CCI] toxicity interval: de-escalate to a lower dose.

Phase 1b will use the mTPI-2 dose finding design with the target toxicity rate of [CCI] with target toxicity interval of [CCI].

2.2.2. Phase 2

In the Phase 2 portion of the study up to two dose levels will be further explored.

The RP2D of vepdegestrant and PF-07220060, identified as the dose that can be used in further large-scale testing, will be determined based on an overall assessment of safety, tolerability, PK and early signals of efficacy. The RP2D may be the same as the MTD, or may be lower. The RP2D will be determined in this study after the Phase 2 expansion cohort(s) have provided sufficient data on which to make such a determination.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

3.1.1. Phase 1b

The primary endpoint in Phase 1b is DLT during the DLT-observation period.

Analyses of DLTs are based on the DLT Evaluable Set. The occurrence of DLTs and AEs constituting DLTs during the DLT observation period will be summarized and listed by dose level for participants enrolled in the Phase 1b part.

3.1.2. Phase 2

The primary efficacy endpoint in Phase 2 is confirmed OR defined as confirmed CR or PR, determined by derived investigator assessment.

ORR is defined as the proportion of participants in the SAS with confirmed OR per RECIST v1.1 criteria assessed by investigator. Both CR and PR must be confirmed by repeat assessments performed no fewer than 28 days after the criteria for response are first met. Participants who do not have a post-baseline tumor assessment due to early progression of disease, who receive anti-cancer therapies other than the study interventions prior to reaching

a CR or PR, or who die, have documented PD, or stop tumor assessments for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of ORR.

3.2. Secondary Endpoints

3.2.1. Safety endpoints

- AEs as characterized by type, frequency, intensity as graded by NCI CTCAE, version 5.0, timing, seriousness, and relationship to vepdegestrant in combination with PF-07220060.

AEs will be graded by the investigator according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 and coded using the Medical Dictionary for Regulatory Activities (MedDRA).

- Laboratory test abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE v5.0), and timing.
- Incidence of ECG abnormalities

3.2.2. Efficacy endpoints

- DoR is defined for participants by derived investigator assessment according to RECIST v1.1 for confirmed OR (CR or PR) as the time from the first documentation of OR to the first documentation of PD or to death due to any cause, whichever occurs first.
- CBR is defined as the proportion of participants with best overall response (BOR) of confirmed CR or PR at any time, or SD ≥ 24 weeks by derived investigator assessment according to RECIST v1.1.
- PFS is defined as the time from the date of first dose of treatment combination to the date of first documentation of PD or death due to any cause, whichever occurs first.
- OS is defined as the time from the date of first dose of treatment combination to the date of death due to any cause.

3.3. Pharmacokinetic Endpoints

Phase 1b and 2

- Blood samples will be collected for PK assessment of vepdegestrant, its epimer ARV-473 (same sample draw), and PF-07220060 according to the PK sampling schedule in the study protocol. Plasma concentrations of vepdegestrant, ARV-473, and PF-07220060 will be used to evaluate changes in drug concentration of each drug overtime and throughout treatment.
- The plasma concentrations will be used to estimate the steady-state C_{max} , T_{max} , AUC_{tau} and AUC_{last} of vepdegestrant, ARV-473, and PF-07220060.

- If data permit and as appropriate, the plasma concentrations will also be used to estimate CL/F, V_z/F , and $t_{1/2}$ of vepdegestrant, ARV-473, and PF-07220060.

3.4. Biomarker Endpoints

- ctDNA of plasma will be analyzed for potential genomic alterations involved in cancer cells to evaluate potential mechanisms of resistance and/or sensitive to testing treatment, which include, but not limited to, cell cycle and cell proliferation related gene alterations. The associations between genomic alterations and clinical outcome will be assessed.
- Evaluate if quantitative changes of ctDNA on treatment from baseline could be a predictor of participant treatment response (i.e. participant with decrease in level of ctDNA during treatment could have longer PFS).

3.5. Exploratory Endpoints

3.5.1. Patient-Reported Outcomes (Phase 2 Only)

- General cancer related symptoms as measured by the EORTC- QLQ-C30

The EORTC QLQ-C30 consists of 30 questions, which can be grouped into 5 functional scales (physical, role, cognitive, emotional, and social); a global health status/global quality of life scale; 3 symptom scales (fatigue, pain, nausea and vomiting); and 6 single items that assess additional symptoms (dyspnea, appetite loss, sleep disturbance, constipation, and diarrhea); and financial impact. All scales and single item measures range in score from 0 to 100. Higher scores on the global health status/quality of life scale represent higher health status/quality of life. Higher scores on the functional scales represent higher levels of functioning. Higher scores on the symptom scales/items represent a greater presence of symptoms.

- General health status as measured by the EQ-5D-5L

The EQ-5D-5L is a 5-item participant-completed questionnaire designed to assess health status in terms of a single index value or utility score. There are 2 components, a Health State Profile which has individuals rate their level of problems (none, slight, moderate, severe, extreme/unable) in 5 areas (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), and a VAS in which participants rate their overall health status from 0 (worst imaginable) to 100 (best imaginable). Published weights are available that allow for the creation of a single summary score. Overall scores range from 0 to 1, with lower scores representing a higher level of dysfunction (Rabin & de Charro, 2001).

3.5.2. Biomarker Endpoints

- ESR1 and other tumor related gene alterations will be examined to test if such gene alterations lead to resistance or sensitivity to the treatment. Associations of tumor gene/genomic alterations with participant treatment responses and disease biology will be assessed.

3.6. Baseline Variables

3.6.1. Study drug, study treatment and baseline definitions

In this study, ‘**study drug**’ refers to vepdegestrant in combination with PF-07220060, and ‘**study treatment**’ (or ‘**treatment group**’ or ‘**treatment combination**’) refers to:

- Phase 1b Dose Level 1: vepdegestrant 200mg + PF-07220060 100mg
- Phase 1b Dose Level 2: vepdegestrant 200mg + PF-07220060 300mg
- Phase 2: To be determined from Phase 1b.

Start and end dates of study treatment:

The start date/time of first dose of study treatment in a combination group is the earliest date/time of the first non-zero dose date/time for the study drugs in the combination.

The end date/time of last dose of study treatment in a combination group is the latest date/time of the last non-zero dose date/time for the study drugs in the combination.

Definition of baseline:

Definition of baseline for efficacy and PRO analyses and for safety analyses

The last available assessment prior to the start of study treatment is defined as ‘baseline’ value or ‘baseline’ assessment for safety and efficacy analyses. If an assessment is planned to be performed prior to the first dose of study treatment combination in the protocol and the assessment is performed on the same day as the first dose of study treatment combination, it will be assumed that it was performed prior to study treatment administration, if assessment time point is not collected or is missing. If assessment time points are collected, the observed time point will be used to determine pre-dose on study day 1 for baseline calculation. Unscheduled assessments will be used in the determination of baseline. However, if time is missing, an unscheduled assessment on study day 1 will be considered to have been obtained after study treatment combination administration.

Participants who start treatment and discontinue from the study on the same day may have two different sets of data collected on study day 1 (one during study and one in the End of Treatment (EOT) visit). Data reported at the EOT visit are not eligible for baseline selection.

If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be treated and analyzed similar to an unscheduled post-dose measurement.

Baseline for RR and QT/QTc interval assessments will be derived from the visit where both RR and QT are not missing. Triplicate ECGs are collected in the study and the baseline for each ECG measurement is the average of the pre-dose replicate measurements on the baseline day.

Unscheduled assessments will not be included in the calculation of the average. QTcF will be derived based on RR and QT. The average of the replicate measurements will be determined after the derivation of the individual parameter at each time point.

Definition of baseline for biomarker analyses

The last assessment prior to first dose of study treatment will serve as the baseline assessment for biomarker analyses. For biomarkers that are planned to be assessed on Cycle 1 Day 1, it will be assumed that the assessment was performed prior to study treatment administration, if the assessment time point is not collected or is missing.

3.6.2. Baseline characteristics

Baseline characteristics (including demographics, disease history and prior anti-cancer therapies) are described in Section 6.5.1.

3.7. Safety Endpoints

3.7.1. Adverse Events

Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are those events with onset dates occurring during the on-treatment period.

On-treatment period is defined as the time from the first dose of study treatment through the last dose of study treatment + 28 days.

3.7.2. Laboratory Data

Hematology and chemistry results will be programmatically graded according to the NCI CTCAE version 5.0 for relevant parameters. Parameters which cannot be graded will be summarized relative to the normal range (ie, normal range high or normal range low). Additional details are provided in Section 6.6.4.

4. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per standard operating procedures.

Only participants who signed informed consent will be included in the analysis sets below.

Table 3. Analyses Sets

Population	Description
Safety Analysis Set (SAS)	All enrolled participants assigned to study intervention and who take at least 1 dose of either vepdegestrant or PF-07220060 study intervention.
DLT Evaluable Set	All enrolled participants who received at least 1 dose of study interventions in the Phase 1b and either experienced DLT during the DLT observation period or complete the DLT observation period without DLT. Participants without DLTs who receive less than 80% of the planned dose of each study drug during the DLT observation period for reasons other than treatment-related toxicity are not evaluable for DLT. All participants deemed non-evaluable for DLT may be replaced.
Biomarker Analysis set	The biomarker analysis set for each biomarker is a subset of the Safety Analysis Set and includes all participants with at least 1 biomarker evaluated predose.
PK Concentration Set	All participants who are in the Safety Analysis Set and have at least 1 concentration of either ARV-471, ARV-473, or PF-07220060 above the lower limit of quantitation.
PK Parameter Set	All participants who are in the Safety Analysis Set and have at least 1 of the PK parameters of interest calculated for either ARV-471, ARV-473, or PF-07220060.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

5.1.1. Hypotheses and sample size determination

Phase 1b

A sufficient number of participants will be screened to achieve required number of participants enrolled to study intervention. Approximately 15 participants will be enrolled in Phase 1b to evaluate DLTs during the first cycle of treatment to confirm the safety of the combination and identify the RDE. However, the total number of participants will depend on the observed safety profile, which will determine the number of participants at each dose level and the number of dose levels explored. Beginning with the first dose level, participants will be enrolled, treated, and monitored in cohorts of 3-4 participants during the DLT

observation period (Cycle 1). At least 6 participants will be treated at the RDE of the combination before proceeding to Phase 2.

If any participant is deemed non-evaluable for DLT, additional participants may be enrolled to ensure there are a sufficient number of evaluable participants in Phase 1b.

Phase 2

In Phase 2, approximately 25 participants per group will be enrolled. Participants will receive the study drugs in combination at the RP2D (as determined from Phase 1b). The sample size is not based on statistical hypothesis test; rather, it is based on clinical consideration that the stated sample size will provide sufficient evidence of preliminary antitumor activity of vepdegestrant in combination with PF-07220060.

The primary goal of Phase 2 is to estimate the ORR with standard error ≤ 0.11 . A total of 20 participants will provide the specified precision of ORR estimates.

Table 4 provides the 2-sided 95% CIs for ORR based on different observed responders.

Table 4 Exact 95% CIs for ORR for 25 participants

N	Number of Responders	Observed ORR	95% CIs for ORR*
25	1	4%	(0.7% – 19.5%)
25	2	8%	(2.2% – 25.0%)
25	3	12%	(4.2% – 30.0%)
25	4	16%	(6.4% – 34.7%)
25	5	20%	(8.9% – 39.1%)
25	6	24%	(11.5% – 43.4%)
25	7	28%	(14.3% – 47.6%)
25	8	32%	(17.2% – 51.6%)
25	9	36%	(20.2% – 55.5%)
25	10	40%	(23.4% – 59.3%)
25	11	44%	(26.7% – 62.9%)
25	12	48%	(30.0% – 66.5%)
25	13	52%	(33.5% – 70.0%)
25	14	56%	(37.1% – 73.3%)
25	15	60%	(40.7% – 76.6%)
25	16	64%	(44.5% – 79.8%)
25	17	68%	(48.4% – 82.8%)
25	18	72%	(52.4% – 85.7%)
25	19	76%	(56.6% – 88.5%)
25	20	80%	(60.9% – 91.1%)

* Using Wilson score method (not corrected).

5.1.2. Decision rules

Phase 1b

A safe dose will be determined using the adaptive mTPI-2 design (Guo et al, 2017) described in Appendix 3.

5.2. General Methods

As described in Section 3.6, in this study ‘**treatment group**’ refers to:

- Phase 1b Dose Level 1: vepdegestrant 200mg + PF-07220060 100mg
- Phase 1b Dose Level 2: vepdegestrant 200mg + PF-07220060 300mg
- Phase 2: To be determined from Phase 1b

Table 5 provides an overview of the summaries and the tabulations for this study.

Table 5. Summary of Tabulations

Summaries	Analysis Population	Phase 1b	Phase 2	Phase 1b, and Phase 2 combined
Baseline characteristics and disposition	Safety analysis set	- by treatment group - for all treatment groups combined	For all Phase 2 participants	For all participants treated at RP2D
DLTs	DLT-evaluable set	- by treatment group - for all treatment groups combined	Not applicable	Not applicable
Efficacy data	Safety analysis set and response evaluable analysis set (OR only)	- by treatment group - for all treatment groups combined	For all Phase 2 participants	For all participants treated at RP2D
Other safety data, exposure	Safety analysis set	- by treatment group	For all Phase 2 participants	For all participants

data, concomitant medications, non-drug treatment		- for all treatment groups combined		treated at RP2D
PK data for ARV-471 and ARV-473	PK concentration set	- by treatment group - by cycle, day and nominal time	- by cycle, day and nominal time	For all participants
PK data for PF-07220060	PK concentration set	- by treatment group - by cycle, day and nominal time	- by cycle, day and nominal time	For all participants treated at RP2D
Biomarker data	Biomarker analysis set		For all Phase 2 participants	For all participants treated at RP2D
PRO data	Safety analysis set		For all Phase 2 participants	For all participants treated at RP2D

5.2.1. Data Handling after the cut-off date

Data after the cut-off date may not undergo the cleaning process and will not be displayed in any listings or used for summary statistics, statistical analyses or imputations.

5.2.2. Pooling of centers

In order to provide overall estimates of treatment effects, data will be pooled across centers. The 'center' factor will not be considered in statistical models or for subgroup analyses due to the high number of participating centers in contrast to the anticipated small number of participants treated at each center.

5.2.3. Presentation of continuous and qualitative variables

Continuous variables will be summarized using descriptive statistics i.e., number of non-missing values and number of missing values [i.e., n (missing)], mean, median, standard deviation (SD), minimum, maximum and first and third quartile (Q1 and Q3).

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise specified, the calculation of proportions will include the missing category. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

In case the analysis refers only to certain visits, percentages will be based on the number of participants still present in the study at that visit, unless otherwise specified.

5.2.4. Definition of study day

Start day of study treatment is the day of the first dose of study treatment.

The study day for assessments occurring on or after the start of study treatment (e.g., adverse event onset, tumor measurement) will be calculated as:

$$\text{Study day} = \text{Date of the assessment/event} - \text{start of study treatment} + 1.$$

The study day for assessments occurring prior to the first dose of study treatment (e.g., baseline characteristics, medical history) will be negative and calculated as:

$$\text{Study day} = \text{Date of the assessment/event} - \text{start of study treatment}.$$

The study day will be displayed in all relevant data listings.

5.2.5. Definition of start of new anti-cancer therapy

Start date of new anti-cancer therapy (drug, radiation, surgery) is used for censoring in efficacy analyses (see Section 6.1.2 and Section 6.2.2).

The start date of new anti-cancer therapy is the earliest date after the first dose of study treatment amongst the following:

- Start date of anti-cancer drug therapy recorded in the 'Cancer Therapy' eCRF pages
- Start date of radiation therapy recorded in 'Radiation Treatment eCRF pages with 'Treatment Intent' = 'Curative in intent'.
- Surgery date recorded in 'Cancer Surgery' eCRF pages when 'Surgery Outcome' = 'Resected' or 'Partially Resected'.

When start date of anti-cancer therapy is missing or partially missing, the imputation rules described in Section 5.3.3.4 should be applied using 'Cancer Therapy', 'Radiation Treatment', 'Cancer Surgery' eCRF pages.

5.2.6. Definition of on-treatment period

Safety endpoints will be summarized based on the on-treatment period unless otherwise specified.

On-treatment period is defined as the time from the first dose of study treatment through the last dose of study treatment + 28 days.

Safety data collected outside the on-treatment period as described above will be listed and flagged in listings but not summarized.

5.2.7. Standard derivations and reporting conventions

The following conversion factors will be used to convert days into weeks, months or years: 1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

Demographics and physical measurements:

Age [years]: (year of given informed consent - year of birth)

The integer part of the calculated age will be used for reporting purposes.

For reporting conventions, mean and median should generally be displayed in one more decimal place than the raw data and standard deviation should be displayed to two more decimal places than the raw data. Percentages will be reported to one decimal place. The rounding will be performed to closest integer / first decimal using the common mid-point between the two consecutive values. E.g., 5.1 to 5.4 will be rounded to an integer of 5, and 5.5 to 5.9 will be rounded to an integer of 6.

5.2.8. Unscheduled visits

Generally, data collected at unscheduled visits will be included and analyzed for both safety and efficacy analyses in the same fashion as the data collected at scheduled visits except where otherwise noted in the sections that follow. Descriptive statistics (mean, SD, median, minimum, maximum, quartiles) by nominal visit or time point for safety endpoints such as laboratory measurements, ECGs and vital signs will include only data from scheduled visits (if such analyses are performed).

PK concentration and PK parameter data from unscheduled visits will be listed and may also be summarized and analyzed. Unscheduled PK concentrations and PK parameters may be summarized with other participants during scheduled visits if the time/date of collection and dosing history is comparable. Any summaries and analyses with unscheduled samples will also be repeated with same outputs without unscheduled samples.

5.2.9. Adequate baseline tumor assessment

Adequate baseline is defined using the following criteria:

- All baseline assessments must be within 28 days prior to and including the start date of study treatment.
- All documented lesions must have non-missing assessments (i.e. non-missing measurements for target lesions and non-missing lesions assessment status at baseline for non-target lesions).

5.2.10. Adequate post-baseline tumor assessment

An adequate post-baseline assessment is defined as an assessment where a response of CR, PR, SD, non-CR/non-PD, or PD can be determined (see Section 6.2.2.5). Time points where the response is not evaluable (NE) or no assessment was performed will not be used for determining the censoring date.

5.3. Methods to Manage Missing Data

5.3.1. Missing data

Unless otherwise specified, all data will be evaluated as observed, and no imputation method for missing values will be used.

In all participant data listings imputed values will be presented. In all listings imputed information will be flagged.

Missing statistics, eg when they cannot be calculated, should be presented as 'ND' or 'NA'. For example, if $N=1$, the measure of variability (SD) cannot be computed and should be presented as 'ND' or 'NA'.

5.3.1.1. Pharmacokinetic concentrations

Concentrations Below the Limit of Quantification

For all calculations, figures and estimation of individual pharmacokinetic parameters, all concentrations assayed as below the level of quantification (BLQ) will be set to zero. In log-linear plots these values will not be represented. The BLQ values will be excluded from calculations of geometric means and their CIs. A statement similar to 'All values reported as BLQ have been replaced with zero' should be included as a footnote to the appropriate tables and figures.

Deviations, Missing Concentrations and Anomalous Values

In summary tables and plots of median profiles, concentrations will be set to missing if one of the following cases is true:

- A concentration has been reported as ND (ie, not done) or NS (ie, no sample);
- A deviation in sampling time is of sufficient concern or a concentration has been flagged as anomalous by the clinical pharmacologist. If this is performed, a footnote or other documentation in the clinical study report (CSR) detailing the exclusions will be provided.

Summary statistics will not be presented at a particular time point if less than 3 evaluable measurements are available. For analysis of pharmacokinetic concentrations, no values will be imputed for missing data.

5.3.1.2. Pharmacokinetic parameters

Actual sampling time will be used for the derivation of PK parameters. If a PK parameter cannot be derived from a participant's concentration data, the parameter will be coded as NC (ie, not calculated). NC values will not be generated beyond the day that a participant discontinues.

In summary tables, statistics will be calculated by setting NC values to missing. Statistics will not be presented for a particular treatment if less than 3 evaluable measurements are available. For statistical analyses (ie, analysis of variance), PK parameters coded as NC will also be set to missing.

If an individual participant has a known biased estimate of a PK parameter (due for example to a deviation from the assigned dose level), this will be footnoted in summary tables or flagged in the listings and will not be included in the calculation of summary statistics or statistical analyses.

5.3.2. Handling of incomplete dates

5.3.2.1. Missing Dates

For purposes of data listings, dates will reflect only the information provided by the investigator on the CRF.

If start dates for adverse events or concomitant medications are completely missing a worst case approach will be taken whereby the events will be considered treatment emergent and the medications will be considered concomitant. If only partial information is available (e.g. only a month and year or only a year) and the partial information provide sufficient information to indicate the dates are prior to the start of study treatment (e.g. month/year less than month/year of first dose) then these will be considered to have started prior to treatment; otherwise a similar worst case approach will apply, and these will be considered to have started after treatment.

5.3.2.2. Adverse events

AE Onset Date

If the date of first exposure to study drug treatment is earlier than AE stop date, the missing AE onset date is set to the date of first exposure to study drug. Otherwise if the date of first exposure to study drug is after AE stop date, then the missing AE onset date is set to the earliest of non-missing AE stop date or informed consent date.

AE Stop Date

Ongoing events will have the AE Stop Date set to one of the following values:

- Date of Death, if the participant died and a date of death exists.
- Maximum of (Participant Withdraw date, AE Onset Date, AE Collection Date) if the participant withdrew from the study and a date of withdraw exists.
- Maximum of (AE Onset Date, Participant Summary Collection Date, AE Collection Date) if the Disposition CRF page for the long-term follow-up phase exists but a date of completion/discontinuation does not exist.
- Maximum of (Last Treatment Date, AE Onset Date) if no Disposition CRF page for the long-term follow-up phase exists.

Imputation will only occur if event is unique for the participant, or it is the last of a series of similar events; otherwise, the Stop Date will not be imputed. Adverse Events are deemed similar if they have the same verbatim term.

Missing Toxicity Grade of Adverse Events

Prior to Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event prior to the first study treatment, then Grade 1 will be assumed for purposes of defining a baseline grade for assessing if further occurrences are treatment emergent. However, if the participant experiences multiple episodes of the same AE prior to study treatment, then the maximum toxicity grade observed prior to study treatment will be utilized in assessing if further occurrences are treatment emergent.

During Study Treatment: If no toxicity grade is available or the grade is reported as unknown for an adverse event during the study treatment, then the event will be considered treatment emergent without toxicity grade unless a baseline event was reported as Grade 4.

In summaries which present maximum toxicity grade, the maximum of non-missing grades will be displayed. Missing grade will only be displayed for cases where a participant reported only one event and the grade is missing.

Missing ECG Data

For ECG analyses, no values will be imputed for missing data (except RR values that can be derived from HR values, if present). In case of missing RR value, RR (msec) will be derived as $(60/\text{HR (bpm)}) * 1000$, if HR is collected.

If one or two of the triplicate measurements for an ECG parameter are missed, the average of the remaining two measurements or the single measurement can be used in the analyses. If all triplicate measurements are missing at a timepoint for an ECG parameter, no values will be imputed for this timepoint. If the triplicate needs to be repeated because of an artifact, then the repeated triplicate will be reported on an unscheduled CRF page. Based on a review of the data these unscheduled assessments may be used in place of the assessments at the nominal time. Data review and consultation with the study team is required to flag these cases.

Other Missing or Partial Dates

Imputation methods for other partial dates as follows:

DMB02-GSOP-RF02 7.0 Statistical Analysis Plan Template 31-Jan-2022

PFIZER CONFIDENTIAL

TMF Doc ID: 98.03

Page 26 of 66

- If the day of the month is missing for a start date used in a calculation, the first day of the month will be used to replace the missing date.
- If both the day and month are missing for a start date, the first day of the year is used.
- For stop dates, the last day of the month, or last day of the year is used if the day or day and month are missing, respectively. Provided the imputed date is not after the date of death.
- If the date is completely missing, no imputation will be performed.

5.3.2.3. Prior and concomitant medications

Incomplete prior/concomitant medication dates will be imputed as follows:

- If the medication date is missing completely, then no date imputation is made. If the day of medication date is missing, Pfizer general imputation rule will be applied. Like missing start month and day will be imputed as 01JAN and missing stop month and day will be imputed as 31DEC. Incomplete stop date will be replaced by the last day of the month (if day is missing only) or 31DECYYYY (if both month and day are missing), if not resulting in a date later than the date of participant's death. In the latter case the date of death will be used to impute the incomplete stop date.
- In all other cases the incomplete medication stop date will not be imputed.

5.3.2.4. Exposure

No imputation will be done for first dose date. Date of last dose of study treatment, if unknown or partially unknown, will be imputed as follows:

- If the last date of study treatment is completely missing and there is no Disposition CRF page for the treatment phase and no death date, the participant should be considered to be ongoing and use the data cutoff date for the analysis as the last dosing date; or
- If the last date of study treatment is completely or partially missing and there is EITHER a Disposition CRF page for the treatment phase OR a death date available (on or prior to the data cutoff date), then impute this date as the last dose date:

- = 31DECYYYY, if only Year is available and Year < Year of min (Date of Completion/Discontinuation from the Disposition CRF page for the treatment phase, death date),
- = Last day of the month, if both Year and Month are available and Year = Year of min (Date of Completion/Discontinuation from the Disposition CRF page for the treatment phase, death date) and Month < the month of min (EOT date, death date), or
- = min (Date of Completion/Discontinuation from the Disposition CRF page for the treatment phase, death date), for all other cases.

5.3.3. Imputation rules for date of last contact and efficacy assessments

5.3.3.1. Date of last contact

The date of last contact will be derived for participants not known to have died at the analysis cut-off using the latest complete date among the following:

- All participant assessment dates (eg, blood draws (laboratory, PK), vital signs, performance status, ECG, tumor assessments)
- Start and end dates of anti-cancer therapies after study treatment discontinuation
- Start and end dates of concomitant therapies including non-drug treatment or procedures.
- AE start and end dates
- Last date of contact were collected on the ‘Survival Follow-up’ eCRF (do not use date of survival follow-up assessment unless status is ‘alive’)
- Study drug start and end dates
- Withdrawal of consent date
- Date of discontinuation on disposition eCRF pages (do not use if reason for discontinuation is lost to follow-up or death).

Only dates associated with actual examinations of the participant will be used in the derivation. Dates associated with a technical operation unrelated to participant status such as the date a blood sample was processed will not be used. Assessment dates after the cut-off date will not be applied to derive the last contact date.

5.3.3.2. Death date

Missing or partial death dates will be imputed based on the last contact date:

- If the entire date is missing it will be imputed as the day after the date of last contact (see derivation of date of last contact in section 5.2.5.); or
- If the day or month is missing, death will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the following:

Missing day: 1st day of the month and year of death, or

Missing day and month: January 1st of the year of death.

5.3.3.3. Tumor assessments

All investigation dates (eg, X-ray, CT scan) must be completed with day, month and year.

If there are multiple scan dates associated with an evaluation, ie, radiological assessments occur over a series of days rather than the same day, the choice of date of assessment could

impact the date of progression and/or date of response. If there are multiple scan dates associated with an evaluation, the earliest of the scan dates associated with the evaluation will be used as the date of assessment.

If one or more investigation dates for an evaluation are incomplete but other investigation dates are available, the incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the earliest of all investigation dates (eg, X-ray, CT-scan).

If all measurement dates for an evaluation have no day recorded, the 1st of the month is used.

If the month is not completed for any of the investigations for an evaluation, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both the previous and following assessments are not available, this assessment will not be used for any calculations.

5.3.3.4. Date of start of new anti-cancer therapy

Incomplete dates for new anti-cancer therapy will be imputed as follows and will be used to determine censoring dates for efficacy analyses:

- The end date of new anti-cancer therapy will be included in the imputation for start date of new anti-cancer therapy. If the end data of new anti-cancer therapy is:
 - completely missing then it will be ignored in the imputations below,
 - partially missing with only year available then the imputations below will consider 31DECYYYY as the end date of the new anti-cancer therapy, or
 - partially missing with only month and year available then the imputations below will consider the last day of the month for MMMYYYY as the end date of the new anti-cancer therapy.
- For participants who have not discontinued study treatment at the time of the data cutoff date, last dose of study treatment is set to the data cutoff date in the imputations below.
- If the start date of new anti-cancer therapy is completely or partially missing then the imputed start date of new anti-cancer therapy is:

= 31DECYYYY, if only Year is available and Year < Year of min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

= Last day of the month, if both Year and Month are available and

Year = Year of min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

Month < Month of min [max (PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]

= min [max (PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy], for all other cases.

6. ANALYSES AND SUMMARIES

Refer to Section 4 for definitions of analysis sets and Section 5 for general methodology.

6.1. Primary Endpoint(s)

6.1.1. DLT for Phase 1b

6.1.1.1. Primary Analysis

DLT rate is calculated as the number of DLT-evaluable participants with DLTs in the DLT evaluation period divided by the number of DLT-evaluable participants in the DLT-observation period.

DLTs will be listed and summarized by treatment group and for all treatment groups combined, including data from Phase 1b only.

6.1.2. Phase 2

6.1.2.1. Objective Response

6.1.2.2. Primary Analysis

ORR is defined as the proportion of participants in the SAS with confirmed OR per RECIST v1.1 criteria assessed by derived investigator assessment. Both CR and PR must be confirmed by repeat assessments performed no fewer than 28 days after the criteria for response are first met. Participants who do not have a post-baseline tumor assessment due to early progression of disease, who receive anti-cancer therapies other than the study interventions prior to reaching a CR or PR, or who die, have documented PD, or stop tumor assessments for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of ORR.

ORR will be calculated along with the 2-sided 95% CI using the Wilson score method (Newcombe, 1998; Wallis, 2013).

Refer to Section 6.2.2.1 for assessment of response.

ORR will be presented as summarized in Section 5.2.

6.2. Secondary Endpoint(s)

6.2.1. Safety endpoints

Refer to Section 6.6.

6.2.2. Efficacy endpoints

6.2.2.1. Objective response as assessed by derived investigator assessment per RECIST V1.1

BOR will be assessed based on derived overall lesion responses at different evaluation time points from the date of first dose of study treatment combination until the first documentation of PD, according to the following rules. Only tumor assessments performed on or before the start date of any further anti-cancer therapies will be considered in the assessment of BOR. Clinical deterioration will not be considered as documentation of disease progression.

- CR = at least two determinations of CR at least 28 days apart and before first documentation of PD
- PR = at least two determinations of PR or better (PR followed by PR or PR followed by CR) at least 28 days apart and before first documentation of PD (and not qualifying for a CR)
- SD (applicable only to participants with measurable disease at baseline) = at least one SD assessment (or better) ≥ 6 weeks after the date of first dose of study treatment combination and before first documentation of PD (and not qualifying for CR or PR).
- Non-CR/non-PD (applicable only to participants with non-measurable disease at baseline) = at least one non-CR/non-PD assessment (or better) ≥ 6 weeks after the date of first dose of study treatment combination and before first documentation of PD (and not qualifying for CR or PR). Note that, enrolling a non-measurable disease is a major protocol violation for this study.
- PD = first documentation of PD after the date of first dose of study treatment and not meeting the criteria for “PD too late”.
- NE: all other cases.

An objective status of PR or SD cannot follow one of CR. SD can follow PR only in the rare case that tumor increases by less than 20% from the nadir, but enough that a previously documented 30% decrease from baseline no longer holds. If this occurs, the sequence PR-SD-PR is considered a confirmed PR. A sequence of PR – SD – SD – PD would be a best response of SD if the window for SD definition has been met.

Objective Response (OR) is defined as confirmed BOR of CR or PR according to RECIST v1.1.

Participants who do not have a post-baseline radiographic tumor assessment due to early progression, who receive anti-cancer therapies other than the study treatments prior to reaching a CR or PR, or who die, progress, or drop out for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of OR. Each participant will have an objective response status (0: no OR; 1: OR). OR rate (ORR) is the proportion of participants with OR in the analysis set.

A supplementary analysis of ORR in the Response Evaluable Set will also be performed using the same methods as noted above.

ORR by treatment group will also be calculated along with the 2-sided 95% CI using the Wilson score method, see Appendix 4.

In addition, the frequency (number and percentage) of participants with a confirmed BOR of CR, PR, SD, non-CR/non-PD (applicable only to participants with non-measurable disease at baseline), PD, and NE will be tabulated. Participants with confirmed BOR of NE will be summarized by reason for having NE status. The following reasons will be used:

- No baseline assessment
- No post-baseline assessments due to death
- No post-baseline assessments due to other reasons
- All post-baseline assessments have overall response NE
- New anti-cancer therapy started before first post-baseline assessment
- SD of insufficient duration (< 6 weeks after the date of first dose of study treatment combination without further evaluable tumor assessments)
- PD too late (after two inadequate assessments or after missing two consecutive planned assessments)

Special and rare cases where BOR is NE due to both SD of insufficient duration and late PD will be classified as ‘SD too early’ (i.e., SD of insufficient duration).

6.2.2.2. Tumor shrinkage from baseline

Tumor shrinkage will be summarized as the percent change from baseline in target lesions (sum of longest diameter for non-nodal lesion and short axis for nodal lesion) per time point. It will be derived as:

- $((\text{Sum of target lesions at week XX} - \text{sum of target lesions at baseline}) / \text{sum of target lesions at baseline}) \times 100$

The maximum reduction in target lesions from baseline will be derived across all the post-baseline assessments until documented disease progression, excluding assessments after start of subsequent anti-cancer therapy, as:

- Minimum of $((\text{sum of target lesions at week XX} - \text{sum of target lesions at baseline}) / \text{sum of target lesions at baseline}) \times 100$

An enhanced waterfall plot proposed by Moon et al. (JCO, 2021) will be generated. The waterfall will include all the participants with and without measurable disease:

- For participants with measurable disease at baseline and had at least one post-baseline tumor assessment, maximum percent reduction in the sum of longest diameter for non-nodal lesions and short axis for nodal lesions from baseline will be calculated.

The best percentage change from baseline in the sum of the diameters of all target lesions will be presented.

- For participants with no measurable disease at baseline who subsequently achieved complete response of non-target lesions, the best percentage change set to -100.
- For participants with no measurable disease at baseline who subsequently had best overall response of non-CR/non-PD, the best percent change set to 0.
- For participants whose best overall response was progression due to new lesions or unequivocal worsening of non-target lesions at the first post-baseline assessment, the best percent change set to +21.
- For participants whose best overall response was non-evaluable due to treatment failure because of no post-baseline assessments due to early death, global deterioration of health status or start of new anti-cancer therapy, the best percent change set to +21.

6.2.2.3. Clinical benefit response

CBR rate is defined as the proportion of participants in the SAS with BOR of confirmed CR or PR at any time, or SD $\geq$ 24 weeks by derived investigator assessment according to RECIST v1.1. CBR will be summarized by frequency counts and percentages, along with 2-sided 95% CI using the Wilson score method.

6.2.2.4. Duration of response

Duration of Response (DoR) is defined, for participants with OR, as the time from the first documentation of objective response (CR or PR) to the date of first documentation of PD or death due to any cause. If a participant has not had an event (PD or death), DoR is censored at the date of last adequate tumor assessment. The censoring rules for DoR are described in Table 6.

$$\text{DoR (months)} = [\text{date of event or censoring} - \text{first date of OR} + 1] / 30.4375$$

Table 6 Outcome and Event Dates for DoR Analyses

Scenario	Date of event/censoring	Outcome
PD or death After at most one missing or inadequate post-baseline tumor assessment	Date of PD or death	Event
PD or death - After 2 or more missing or inadequate post-baseline tumor assessments	Date of last adequate tumor assessment documenting no PD before missed tumor assessments	Censored
No PD and no death	Date of last adequate tumor assessment documenting no progression	Censored
Treatment discontinuation due to 'Disease progression' without documented progression	Not applicable	Information is ignored. Outcome is derived based on documented progression only.
New anti-cancer therapy given	Date of last adequate tumor assessment documenting no PD before new anti-cancer therapy is given	Censored

Kaplan-Meier estimates (product-limit estimates) (Kaplan-Meier, 1958) will be presented together with a summary of associated statistics including the median DoR time with 2-sided 95% CIs. In particular, the DoR rates at 3, 6, 9, and 12 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to (Brookmeyer & Crowley, 1982) and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to (Kalbfleisch & Prentice, 2002) (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

DoR will be displayed graphically and analyzed using Kaplan-Meier methodology. If the number of participants with OR is small, the Kaplan-Meier method may not provide reliable estimates. In this case, only descriptive statistics or listings will be provided.

Frequency (number and percentage) of participants with each event type (PD or death) and censoring reasons will be presented. Reasons for censoring will be summarized according to the categories in Table 7 following the hierarchy shown.

Table 7 DoR Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	Start of new anti-cancer therapy	Start of new anti-cancer therapy
2	Event after 2 or more missing or inadequate post-baseline tumor assessments	Event after 2 or more missing assessments ^a
3	No event and [withdrawal of consent date $\geq$ date of first dose of study treatment combination]	Withdrawal of consent
4	No event and lost to follow-up in any disposition page	Lost to follow-up
5	No event and none of the conditions in the prior hierarchy are met	Ongoing without an event

^a 2 or more missing or inadequate post-baseline tumor assessments.

6.2.2.5. Progression-free survival

Progression-Free Survival (PFS) is defined as the time from the date of first dose of study treatment combination to the date of the first documentation of PD or death due to any cause, whichever occurs first.

PFS data will be censored on the date of the last adequate tumor assessment for CCI

In this study, antitumor activity will be assessed through radiological tumor assessments conducted at screening, every 8 weeks (± 7 days) for 1 year from start of study treatment, and then every 12 weeks (± 7 days) thereafter until disease progression regardless of initiation of subsequent anti-cancer therapy.

The censoring and event date options to be considered for the PFS analysis are presented in Table 8.

$$\text{PFS (months)} = [\text{date of event or censoring} - \text{date of first dose of study treatment combination} + 1] / 30.4375$$

Table 8 Outcome and Event Dates for PFS Analyses

Scenario	Date of event/censoring	Outcome
CCI		
PD or death - After at most one missing or inadequate post-baseline tumor assessment, or - ≤ 16 weeks after the date of first dose of study treatment combination	Date of PD or death	Event
CCI		
Treatment discontinuation due to 'Disease progression' without documented progression	Not applicable	Information is ignored. Outcome is derived based on documented progression only.
CCI		

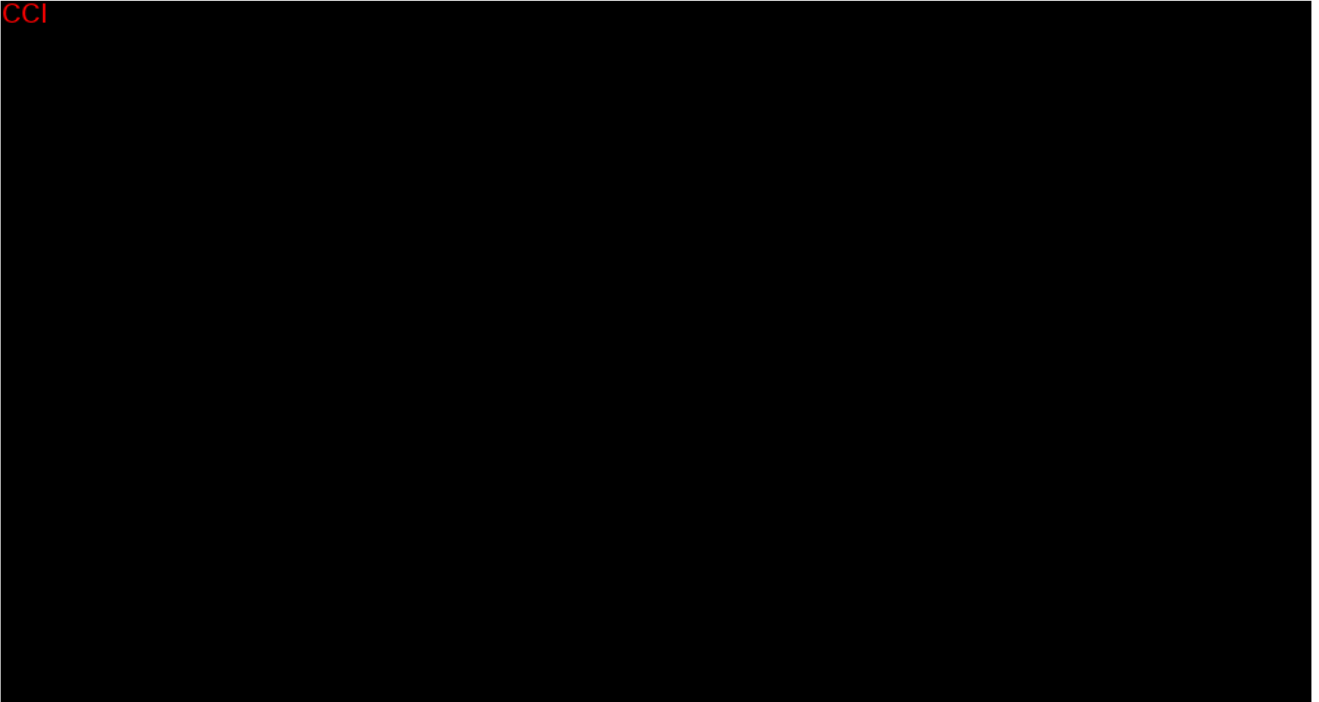
Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including the median PFS time with 2-sided 95% CIs. In particular, the PFS rates at 3, 6, 9, and 12 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to (Brookmeyer & Crowley, 1982) and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to (Kalbfleisch & Prentice, 2002) (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of participants with each event type (PD or death) and censoring reasons will be presented. Reasons for censoring will be summarized according to the categories in Table 9 following the hierarchy shown.

The PFS time or censoring time and the reasons for censoring will also be presented in a participant listing.

Time of Follow-Up for PFS

A plot will be generated to compare planned and actual relative day of tumor assessments. A Kaplan-Meier plot for PFS follow-up duration will also be generated to assess the follow-up time in the treatment groups reversing the PFS censoring and event indicators, including the median time of follow-up for PFS with 2-sided 95% CI.



6.2.2.6. Overall Survival

Overall survival (OS) is defined as the time from the date of first dose of study treatment combination to the date of death due to any cause. Participants last known to be alive will be censored at the date of last contact.

$$\text{OS (months)} = [\text{date of death or censoring} - \text{date of first dose of study treatment combination} + 1] / 30.4375$$

Kaplan-Meier estimates⁶ (product-limit estimates) will be presented together with a summary of associated statistics including the median OS time with 2-sided 95% CIs. In particular, the OS rates at 3, 6, 9, 12, 15, and 18 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (1982) and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (2002) (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of participants with an event (death) and censoring reasons will be presented. Reasons for censoring will be summarized according to the categories in Table 10 following the hierarchy shown.

Table 10 OS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No event and [withdrawal of consent date $\geq$ date of first dose of study treatment or End of study (EOS) = Participant refused further follow-up]	Withdrawal of consent
2	No event and [lost to follow-up in any disposition page or data cut-off date – last contact date > 12 weeks]	Lost to follow-up
3	No event and none of the conditions in the prior hierarchy are met	Alive

The OS time or censoring time and the reasons for censoring will also be presented in a participant listing.

Time of Follow-Up for OS

A Kaplan-Meier plot for OS follow-up duration will also be generated to assess the follow-up time in the treatment groups reversing the OS censoring and event indicators, including the median time of follow-up for OS with 2-sided 95% CIs.

6.2.3. Pharmacokinetic endpoints

Participants will have pharmacokinetic sampling performed for vepdegestrant, ARV-473, and PF-07220060 as described in section 3.3. For periods with serial PK sampling, pharmacokinetic parameters ($C_{\max}$, $T_{\max}$, AUC_{last} , AUC_{tau} , CL/F , $t_{1/2}$, and V_z/F) for all three analytes will be estimated using noncompartmental methods if applicable and if data permits. For periods of sparse PK sampling, only descriptive statistics of plasma concentrations will be summarized.

The following summaries are planned for the PK analysis in each phase of the study.

Phase 1b Assessment Cohort(s):

- Listing of individual PK concentrations by nominal and/or actual sampling time.
- Descriptive statistics (n, mean, SD, %CV, median, minimum, maximum) of plasma concentrations for vepdegestrant, ARV-473 and PF-07220060 will be presented in tabular form by treatment cohort, cycle, day and nominal time.
- For periods with serial PK sampling, linear and semi-log plots of median and mean plasma concentrations by nominal time for vepdegestrant, ARV-473, and PF-07220060 will be prepared by treatment cohort, cycle, and day. Similar plots will be prepared for each individual participant plasma concentrations.
- Pharmacokinetic parameters for vepdegestrant, ARV-473, and PF-07220060 will be listed and summarized by treatment cohort, cycle, and day using descriptive statistics as shown in Table 11. PK parameters with zero values will be set to 0.0001 for the calculation of geometric mean and geometric CV. If an intra participant dose escalation or reduction occurs that may affect PK analysis, data for that participant will only be included in descriptive statistics and summary plots up to the time of the dose change.

Table 11 Summary Statistics for PK Parameters.

Parameter	Summary Statistics
$C_{\max}$, AUC_{last} , AUC_{tau} , CL/F , and V_z/F ,	N, arithmetic mean, median, SD, %CV, minimum, maximum, geometric mean and geometric %CV
$T_{\max}$,	N, median, minimum, maximum
$t_{1/2}$	N, arithmetic mean, median, SD, %CV, minimum, maximum

Phase 2:

In the Phase 2 portion of the study, sparse PK samples of vepdegestrant, ARV-473, and PF-07220060 will be collected as described in Section 3.3. The following summaries are planned for the PK analysis in Phase 2:

- Table of individual PK concentrations by nominal time.
- Descriptive statistics (n, mean, SD, %CV, median, minimum, maximum) of plasma concentrations for vepdegestrant, ARV-473, and PF-07220060 will be presented in tabular form, cycle, day and nominal time.

6.2.4. Population pharmacokinetic endpoints

Pharmacokinetic and/or pharmacodynamic data from this study may be analyzed using modeling approaches and may also be pooled with data from other studies to investigate any association between vepdegestrant and PF-07220060 exposure and biomarkers or significant safety/efficacy endpoints. The results of these analyses, if performed, may be reported separately.

6.2.5. Biomarker endpoints

Secondary endpoints in the study are to evaluate quantitative changes of ctDNA levels of participant plasma from baseline and its association with treatment outcome.

6.3. Other Endpoints: Patient-reported outcomes

Analyses of the patient-reported outcomes will be based on the SAS. Analyses of EORTC QLQ-C30 and EQ-5D-5L will be done Phase 2 only.

Descriptive statistics (n, mean, SD, median, minimum, maximum) of absolute scores over time will be reported for each of the total and subscales of the EORTC QLQ-C30, and the EQ-5D-5L.

The mean change of absolute scores from baseline (and 95% CI) over time will also be reported.

Line charts depicting the mean changes with standard error (SE) bars over time will be provided overall and for each cohort.

For the EQ-5D-5L health state profiles, the proportions of participants reporting having “none”, “slight”, “moderate”, “severe”, or “extreme/unable” problems at each time point will be reported.

The completion rate of the questionnaires EORTC QLQ-C30 and EQ-5D-5L will be summarized by item.

6.4. Subset Analyses

OR and DoR (if meaningful) will also be summarized for all participants in the Phase 1b treated with vepdegestrant and PF-07220060 at the RP2D and in the Phase 2 combined:

- ECOG status: 0 vs. ≥ 1
- Prior CDK4/6i treatment: <12 vs ≥ 12 months
- Prior fulvestrant treatment: Yes/No

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline Summaries

6.5.1.1. Demographic characteristics

Demographic characteristics will be summarized using the following information from the 'Screening/Baseline Visit' eCRF pages.

- Demographic characteristics
 - Gender: Male, Female
 - Race: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Not Reported
 - Ethnic origin: Hispanic/Latino (Yes/No/Not Reported)
 - Menopausal status: Yes/No
 - Age (years): summary statistics
 - Age categories :
 - < 65 years, ≥ 65 years
 - < 65 , $65-<75$, $75-<85$, ≥ 85 years
 - Pooled Geographical Region (as applicable):
 - North America
 - Europe
 - Asia
 - Rest of the World (Australasia, Latin America, Africa and/or Middle East will be included as additional pooled geographical regions if including $> 10\%$ of the overall randomized population)

- Geographic Region (as applicable):
 - North America
 - Latin America
 - Western Europe
 - Eastern Europe
 - Middle East
 - Australasia
 - Asia
 - Africa
- Eastern Cooperative Oncology Group (ECOG) Performance Status: 0, 1, 2, 3, and 4

Center codes will be used for the determination of the participant's geographic region.

The listing of demographics and baseline characteristics will include the following information: participant identifier, treatment group, age, sex, race, ethnicity, menopausal status and ECOG performance status.

6.5.1.2. Medical history

Medical history will be coded using the most current available version of Medical Dictionary for Regulatory Activities (MedDRA) and will be summarized from the 'General Medical History' eCRF page. Medical history will be summarized as the numbers and percentages of participants by MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) as summary category. Each participant will be counted only once within each PT and SOC.

Medical history will be displayed in terms of frequency tables: ordered by primary SOC and PT in alphabetical order.

6.5.1.3. Disease characteristics

Information on disease characteristics collected on 'Primary Diagnosis' (initial and current) and RECIST eCRF pages will be summarized for the following.

From the 'Primary Diagnosis' eCRF pages:

- Primary diagnosis (summarize all categories collected in the 'Primary Diagnosis' eCRF page)
- Time since initial diagnosis to start date (months), defined as (start date – date of initial diagnosis)/30.4375

From the RECIST eCRF page:

- Involved tumor sites at baseline

Listing of disease history will be provided with all relevant data (as collected on the ‘Primary Diagnosis’ eCRF page) and derived variables as above.

6.5.1.4. Prior anti-cancer therapies

The prior anti-cancer therapies are collected under the ‘Prior Cancer Therapy’, ‘Radiation Treatment - Prior’ and ‘Prior Cancer Surgery’ eCRF pages.

The number and percentage of participants in each of the following anti-cancer therapy categories will be tabulated:

- Participants with at least one type of prior anti-cancer therapy
- Participants with at least one prior anti-cancer drug therapy
- Participants with at least one prior anti-cancer radiotherapy
- Participants with at least one prior anti-cancer surgery, excluding biopsy procedure.

Prior anti-cancer drug therapy will be summarized as follows based on the number and percentage of participants with the following:

- At least one prior anti-cancer drug therapy
- Number of prior anti-cancer drug therapy regimens: missing, 1, 2, 3, ≥ 4
- Intent of Drug Therapy: Neo-Adjuvant, Adjuvant, Advanced – Metastatic

Best response: CR, PR, No CR/No PD, SD, PD, Unknown, and NE. Best response is derived from the last treatment regimen.

The prior anti-cancer drugs will also be summarized based on the number and percentage of participants by the drug class and preferred term. A participant will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. The summary will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

Prior anti-cancer drug therapies will be included in the listing that follow with a flag to identify prior therapies. These will include the participant identification number, and all the relevant collected data-fields on the corresponding eCRF pages.

- Listing of anti-cancer drug therapies

6.5.2. Study Conduct and Participant Disposition

6.5.2.1. Participant disposition

The percentages below will be calculated based on the number of participants in the SAS.

- Total number of participants screened overall

- Number of participants who discontinued from the study prior to treatment with study drug overall and by the main reason for discontinuation
- Number and percentage of treated participants in each of the analysis sets defined in Section 4
- Number and percentage of participants with study drug ongoing (separately for each study drug administered in combination)
- Number and percentage of participants who discontinued study drug overall and by the main reason for discontinuation of study drug (separately for each study drug administered in combination)
- Number and percentage of participants who entered safety follow-up
- Number and percentage of participants who discontinued safety follow-up overall and by the main reason for discontinuation

In addition, the following will be summarized:

- Number and percentage of treated participants overall, by region (Europe, EEA (required by EudraCT), North America, Latin America, Middle East, Asia, Australasia, Africa), by country within region
- Number and percentage of treated participants by center

In addition, a cross tabulation of participants who have discontinued/are ongoing treatment with vepdegestrant vs participants who have discontinued/are ongoing treatment with PF-07220060 will also be provided.

6.5.2.2. Protocol deviations

All protocol violations that impact the safety of the participants and/or the conduct of the study and/or its evaluation will be summarized and reported by category for SAS population.

6.5.3. Study Treatment Exposure

6.5.3.1. Exposure to vepdegestrant

The dose level for vepdegestrant is calculated as actual dose administered (mg/day).

Intended duration of treatment with vepdegestrant (weeks) = (end date – date of first dose of vepdegestrant + 1)/7,

where end date = start date of last cycle with non-zero dose of vepdegestrant + 28 – 1

Actual duration of exposure to vepdegestrant (weeks) =

(last dose date of vepdegestrant – first dose date of vepdegestrant + 1)/7, regardless of unplanned intermittent discontinuations.

Overall cumulative dose (mg) is the sum of the actual doses of vepdegestrant received in the study.

Actual Dose Intensity (DI)

- Overall actual DI (mg/week) = [overall cumulative dose (mg)] / [intended treatment duration (weeks)]

Relative Dose Intensity (RDI)

- $RDI (\%) = 100 \times [\text{overall cumulative dose}] / [\text{intended cumulative dose per week} \times \text{number of weeks from first dose to last dose of vepdegestrant}]$

$$= 100 \times [\text{overall cumulative dose}] / [7 \times d \times \text{duration of exposure to vepdegestrant in weeks}]$$

where d = 100 or 200.

vepdegestrant is given 28 days per 28 days cycle.

6.5.3.2. Exposure to PF-07220060

The dose level for PF-07220060 is calculated as actual dose administered (mg/day).

Intended duration of treatment with PF-07220060 (weeks) = (end date – date of first dose of PF-07220060 + 1)/7,

where end date = start date of last cycle with non-zero dose of PF-07220060 + 28- 1.

Actual duration of exposure to PF-07220060 (weeks) =

$$(\text{last dose date of PF-07220060} - \text{first dose date of PF-07220060} + 1)/7$$

Overall cumulative dose (mg) is the sum of the actual doses of PF-07220060 received in the study.

Actual Dose Intensity (DI)

- Overall actual DI (mg/week) = [overall cumulative dose (mg)] / [intended treatment duration (weeks)]

Relative Dose Intensity (RDI)

- $RDI (\%) = 100 \times [\text{overall cumulative dose}] / [\text{intended cumulative dose per week} \times \text{number of weeks from first dose of PF-07220060 to last dose of PF-07220060}]$

$$= 100 \times [\text{overall cumulative dose}] / [7 \times d \times \text{duration of exposure to PF-07220060 in weeks}]$$

where d = 100 or 300.

PF-07220060 is given 28 days per 28 days cycle.

6.5.3.3. Dose reductions

Applicable to PF-07220060 and vepdegestrant. Dose reduction is defined as a change to a non-zero dose level lower than the prior dose level.

The number and percentage of participants with at least one dose reduction as well as a breakdown of number of dose reductions (e.g. 1, 2) and reason for dose reduction will be summarized.

6.5.3.4. Dose interruptions

Applicable to vepdegestrant and PF-07220060.

An interruption is defined a 0 mg dose administered on one or more days for vepdegestrant and PF-07220060. What follows defines how dose interruptions will be counted in the case of multiple dose interruptions.

- If an interruption occurs consecutively for at least two days, then it will be counted only once (example: If the actual non-zero dose on days 1-3 and the actual dose on days 4-5 is 0 mg, then the total number of dose interruptions is 1).
- If an interruption occurs for more than one day, but the days are not consecutive, i.e. there is at least one dosing day in between, then each dose interruption will be counted as a different occurrence (example: If the actual non-zero dose on days 1, 3 and 5, and the actual dose on days 2 and 4 is 0 mg, the total number of dose interruptions is 2).

A dose interruption is not considered a dose reduction.

The number and percentage of participants with dose interruptions and the corresponding reasons will be summarized. A breakdown of the number of dose interruptions (1, 2, 3, ≥ 4) will be summarized.

6.5.4. Concomitant Medications and Nondrug Treatments

The following analyses will be based on the safety analysis set by treatment group as shown in Table 5.

Concomitant medications are medications, other than study drugs, which started prior to first dose date of study treatment and continued during the on-treatment period as well as those started during the on-treatment period. **Prior medications** are medications, other than study drugs and pre-medications for study drug, which are started before the first dose of study treatment and ended prior to the on-treatment period.

Concomitant medications will be summarized from the 'Concomitant Treatment eCRF page.

Summary of concomitant medications will include the number and percentage of participants by Anatomical Therapeutic Chemical (ATC) Classification level 2 and preferred term. A participant will be counted only once within a given drug class and within a given drug

name, even if he/she received the same medication at different times. If any concomitant medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. The summary tables will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under 'Unavailable ATC classification' category.

6.5.5. Subsequent anti-cancer therapies

The following analyses will be based on the SAS by treatment group.

Anti-cancer drug treatment will be provided in a data listing with data retrieved from 'Cancer Therapy' eCRF page.

Number and percentage of participants with any anti-cancer therapy after discontinuation will be tabulated overall and by type of therapy based on the data collected from the 'Cancer Therapy', 'Radiation Treatment' and 'Cancer Surgery' eCRF pages.

6.6. Safety Summaries and Analyses

6.6.1. Adverse Events

Treatment-emergent adverse events (TEAEs) are those events with onset dates occurring during the on-treatment period as defined in Section 3.7.1.

All analyses described will be based on TEAEs (started during the on-treatment period) if not otherwise specified. The AE listings will include all AEs (whether treatment-emergent or not). TEAEs outside the on-treatment period will be flagged in the listings.

- **Related Adverse Events:** adverse events with relationship to study treatment (as recorded on the AE eCRF page, Relationship with study treatment = Related) reported by the investigator and those of unknown relationship (ie, no answer to the question 'Relationship with study treatment'). Related AEs are those related to any study drug (ie, at least one of the study drugs).
- **Serious Adverse Events (SAE):** serious adverse events (as recorded on the AE eCRF page, Serious Adverse Event = Yes).
- **Adverse Events Leading to Dose Reduction:** adverse events leading to dose reduction of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = Dose reduced).
- **Adverse Events Leading to Interruption of Study Treatment:** adverse events leading to interruption of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = Drug interrupted).
- **Adverse Events Leading to Permanent Treatment Discontinuation:** adverse events leading to permanent discontinuation of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = Drug withdrawn).

- **Adverse Events Leading to Death:** adverse event leading to death (as recorded on the AE eCRF page, Outcome = Fatal, as well as AEs of Grade 5).

Unless otherwise specified, AEs will be summarized by number and percentage of participants with the AE in the category of interest as described above, by treatment group, PT in decreasing frequency based on the frequencies observed for the group with all participants in the Phase 1b treated with vepdegestrant and PF-07220060 at the RP2D and in the Phase 2 combined.

Each participant will be counted only once within each SOC or PT. If a participant experiences more than one AE within a SOC or PT for the same summary period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity.

6.6.1.1. All adverse events

Adverse events will be summarized by worst severity (according to NCI-CTCAE version 5.0) per participant, using the latest version of MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) body term as Body System category.

In case a participant has events with missing and non-missing grades, the maximum of the non-missing grades will be displayed. Imputation of missing grades is described in Section 5.3.2.2.

The following tables will be created:

- The overall summary of AEs table will include the frequency (number and percentage) by descending order by MedDRA PT of participants with each of the following:
 - TEAEs
 - TEAEs, Grade ≥ 3
 - Related TEAEs
 - Related TEAEs, Grade ≥ 3
 - TEAEs leading to temporary or permanent discontinuation of any study drugs
 - Related TEAEs leading to temporary or permanent discontinuation of any study drug
 - Serious TEAEs
 - Related Serious TEAEs
 - TEAEs leading to death
 - Related TEAEs leading to death

- TEAEs by SOC and PT or Cluster Term, and NCI CTCAE Max Grade: Any Grade, Grade 3, and Grade 4
- TEAEs related to any study drug by SOC and PT or Cluster, and NCI CTCAE Max Grade: Any Grade, Grade 3, and Grade 4

6.6.1.2. Adverse events leading to dose reduction

The frequency (number and percentage) of participants will be presented for TEAEs leading to dose reduction of vepdegestrant or/and PF-07220060 by SOC, PT and treatment group.

- TEAEs leading to dose reduction of vepdegestrant
- TEAE leading to dose reduction of PF-07220060

6.6.1.3. Adverse events leading to discontinuation of study treatment

The frequency (number and percentage) of participants will be presented for TEAEs leading to temporary or permanent discontinuation of each study drug and study treatment, by treatment group:

- TEAEs leading to discontinuation of any study drug by SOC and PT.
- Related TEAEs leading to discontinuation of any study drug by SOC and PT.
- TEAEs leading to discontinuation of vepdegestrant
- TEAE leading to discontinuation of PF-07220060

The listing of all AEs leading to treatment discontinuation will also be provided with the relevant information.

6.6.2. Deaths

The frequency (number and percentage) of participants in the safety analysis set who died early within 30 days after the first dose and who died within 28 days after last dose of study treatment as well as the reason for death, will be tabulated based on information from the 'Death Detail' and 'Survival Follow-Up' eCRFs, by treatment group.

- All deaths
- Deaths within 30 days after the first dose
- Deaths within 28 days after last dose of study treatment
- Reason for Death
 - Disease progression

- Study treatment toxicity
- AE not related to study treatment
- Unknown
- Other.

In addition, date and cause of death will be provided in individual participant data listing together with selected dosing information (study treatment received, date of first / last administration, dose) and will include the following information:

- AEs with fatal outcome (list preferred terms of AEs with outcome=Fatal, as well as AEs of Grade 5),
- Flag for death within 28 days of last dose of study treatment.

6.6.3. Serious adverse events

The frequency (number and percentage) of participants with each of the following will be presented for treatment-emergent SAEs by treatment group:

- SAEs by SOC and PT or Cluster Term and NCI CTCAE Maximum Grade: Any Grade, Grade 3, and Grade 4
- Related SAEs by SOC and PT or Cluster Term and NCI CTCAE Maximum grade: Any Grade, Grade 3, and Grade 4

The listings of all SAEs will also be provided with the relevant information with a flag for SAEs with onset outside of the on-treatment period.

6.6.4. Laboratory Data

6.6.4.1. Hematology and chemistry parameters

Laboratory results will be classified according to the NCI-CTCAE criteria version 5.0. Non-numerical qualifiers (with the exception of fasting flags) will not be taken into consideration in the derivation of CTCAE criteria (eg, hypokalemia Grade 1 and Grade 2 are only distinguished by a non-numerical qualifier and therefore Grade 2 will not be derived). Additional laboratory results that are not part of NCI-CTCAE will be presented according to the categories: below normal limit, within normal limits and above normal limit (according to the laboratory normal ranges).

Abnormalities classified according to NCI-CTCAE toxicity grading v5.0 will be described using the worst grade. For those parameters which are recorded with two distinct abnormalities such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (eg, hypokalemia) grades at baseline and post baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (eg, hyperkalemia), and vice versa.

For **WBC differential counts** (total neutrophil [including bands], lymphocyte, monocyte, eosinophil, and basophil counts), the absolute value will be used when reported. When only

percentages are available (this is mainly important for neutrophils and lymphocytes, because the CTCAE grading is based on the absolute counts), the absolute value is derived as follows:

$$\text{Derived differential absolute count} = (\text{WBC count}) \times (\text{Differential \%value} / 100)$$

If the range for the differential absolute count is not available (only range for value in % is available) then Grade 1 will be attributed to as follows:

- Lymphocyte count decreased:
 - o derived absolute count does not meet Grade 2-4 criteria, and
 - o % value < % LLN value, and
 - o derived absolute count $\geq 800/\text{mm}^3$
- Neutrophil count decreased
 - o derived absolute count does not meet Grade 2-4 criteria, and
 - o % value < % LLN value, and
 - o derived absolute count $\geq 1500/\text{mm}^3$

For **calcium**, CTCAE grading is based on Corrected Calcium and Ionized Calcium (CALCIO). Corrected Calcium is calculated from Albumin and Calcium as follows

$$\text{Corrected calcium (mmol/L)} = \text{measured total Calcium (mmol/L)} + 0.02 (40 - \text{serum albumin [g/L]})$$

Liver function tests: Alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin (TBILI) are used to assess possible drug induced liver toxicity. The ratios of test result over upper limit of normal (ULN) will be calculated and classified for these three parameters during the on-treatment period.

Summary of liver function tests will include the following categories. The number and percentage of participants with each of the following during the on-treatment period will be summarized by treatment group:

- ALT $\geq 3 \times \text{ULN}$, ALT $\geq 5 \times \text{ULN}$, ALT $\geq 10 \times \text{ULN}$, ALT $\geq 20 \times \text{ULN}$
- AST $\geq 3 \times \text{ULN}$, AST $\geq 5 \times \text{ULN}$, AST $\geq 10 \times \text{ULN}$, AST $\geq 20 \times \text{ULN}$
- (ALT or AST) $\geq 3 \times \text{ULN}$, (ALT or AST) $\geq 5 \times \text{ULN}$, (ALT or AST) $\geq 10 \times \text{ULN}$, (ALT or AST) $\geq 20 \times \text{ULN}$
- TBILI $\geq 2 \times \text{ULN}$
- Concurrent ALT $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$
- Concurrent AST $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$
- Concurrent (ALT or AST) $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$
- Concurrent (ALT or AST) $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$ and ALP $> 2 \times \text{ULN}$
- Concurrent (ALT or AST) $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$ and (ALP $\leq 2 \times \text{ULN}$ or missing)

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, ie, a participant with an elevation of AST $\geq 10 \times \text{ULN}$ will also appear in the categories $\geq 5 \times \text{ULN}$ and $\geq 3 \times \text{ULN}$. Liver function elevation and possible Hy's Law cases will be summarized using frequency counts and percentages.

In addition, a listing of all TBILI, ALT, AST and ALP values for participants with concurrent (ALT or AST) $\geq 3 \times \text{ULN}$ and TBILI $\geq 2 \times \text{ULN}$ and (ALP $\leq 2 \times \text{ULN}$ or missing) and with an ALP $\leq 2 \times \text{ULN}$ or missing will be provided.

Parameters with NCI-CTC grades available:

Shift tables will be provided, to present a summary of participants baseline CTCAE Grade with their worst result on the study grades. The laboratory toxicities will be tabulated using descriptive statistics (number of participants and percentages) during the on-treatment period. The denominator to calculate percentages for each laboratory parameter is the number of participants evaluable for CTCAE grading (ie those participants for whom a Grade 0, 1, 2, 3 or 4 can be derived). Laboratory abnormalities will be listed.

- Table of shift summary of hematology laboratory test results during the on-treatment period by maximum CTCAE grade by each grade
- Table of shift summary of hematology laboratory test results during the on-treatment period by maximum CTCAE grade showing only those with shift from Grade ≤ 2 at Baseline to Grade 3/4 during the on-treatment period
- Table of shift summary of chemistry laboratory test results during the on-treatment period by maximum CTCAE grade by each grade
- Table of shift summary of chemistry laboratory test results during the on-treatment period by maximum CTCAE grade showing only those with shift from Grade ≤ 2 at Baseline to Grade 3/4 during the on-treatment period

The above analyses apply to hematology and chemistry evaluations (if available in database) which can be graded per CTCAE, if data is available in database:

- Hematology:
 - Anemia, Leukocytes (white blood cell decreased), Lymphocytes (lymphocyte count increased/decreased), Neutrophils / Absolute Neutrophils Count (ANC) (neutrophil count decreased), Platelet Count (PLT) (platelet count decreased).
- Serum Chemistry:
 - Albumin (hypoalbuminemia), Alkaline Phosphatase (alkaline phosphatase increased), Alanine Aminotransferase (ALT) (ALT increased), Amylase (serum amylase increased), Aspartate Aminotransferase (AST) (AST increased), Total Bilirubin (blood bilirubin increased), Creatinine (creatinine increased), Creatine Kinase (CPK increased), Potassium

(hypokalemia/ hyperkalemia), Sodium (hyponatremia/ hypernatremia), Magnesium (hypomagnesemia/hypermagnesemia), Calcium (hypocalcemia/ hypercalcemia), Glucose (hypoglycemia/hyperglycemia), Gamma Glutamyl Transferase (GGT) (GGT increased), Lipase (lipase increased), Phosphates (hypophosphatemia/hyperphosphatemia), Triglycerides (hypertriglyceridemia).

Parameters with NCI-CTC grades not available:

- Hematology and chemistry evaluations which cannot be graded per CTCAE criteria will be categorized as low, normal, high and a shift table will be provided.

6.6.4.2. Other laboratory parameters

All other parameters collected on the eCRF will be listed in dedicated listings presenting all corresponding collected information on the eCRF.

- Coagulation: activated partial thromboplastin time (aPTT) and PT or prothrombin time (INR) collected at screening.

The listings of laboratory results will be provided for all laboratory parameters. The listings will be sorted by parameters and assessment dates or visits for each participant. Laboratory values that are outside the normal range will also be flagged in the data listings, along with corresponding normal ranges. A listing of CTCAE grading will also be generated for those laboratory tests.

6.6.5. Electrocardiograms

QTc prolongation is identified as a potential safety risk. The QTc prolongation is currently under investigation for vepdegestrant, since there have been mostly non-serious cases reported in the ongoing Phase 1/2 ARV-471-mBC-101 study and other studies. Early ECG QTc monitoring, along with time matched PK sampling at C1D8 were added.

QTcF will be derived based on RR and QT (see below). The average of the replicate measurements should be determined after the derivation of the individual parameter at each time point.

Selecting Primary QT Correction for Heart Rate

The analysis of QT data is complicated by the fact that the QT interval is highly correlated with heart rate. Because of this correlation, formulas are routinely used to obtain a corrected value, denoted QTc, which is independent of heart rate. This QTc interval is intended to represent the QT interval at a standardized heart rate. Several correction formulas have been proposed in the literature. For this study, the QT interval corrected for heart rate by the Fridericia's formula, QTcF, will be used for analysis and is defined as

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

where RR represents the RR interval of the ECG, in seconds, and can be estimated as 60/Heart Rate.

ECG Summaries

The following analyses will be performed for each applicable ECG parameters (RR, PR, QRS, QT, and QTc) by treatment group, during the on-treatment period. The denominator to calculate percentages for each category is the number of participants evaluable for the category.

- Frequency (number and percentage) of participants with notable ECG values according to the following categories:
 - QT/QTc increase from baseline < 30 ms, 30-60 ms, and > 60 ms
 - Maximum post-dose QT/QTc > (<450 ms, 450- 480 ms, 481-<501 ms, and $\geq$ 501 ms)
 - PR changes from baseline $\geq$ 50% if absolute baseline value was <200 ms, and $\geq$ 25% if absolute baseline value was $\geq$ 200 ms
 - QRS changes from baseline $\geq$ 50% if absolute baseline value was <100 ms, and $\geq$ 25% if absolute baseline value was $\geq$ 100 ms

6.6.6. Vital Signs

Height will be measured at screening only and body weight will be measured at screening and Day 1 of each cycle. Vital signs will not be summarized.

7. INTERIM ANALYSES

There is no formal interim analysis planned for this study.

8. REFERENCES

1. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. *Biometrics*. 38: 29-41, 1982.
2. Eisenhauer E, Therasse P, Bogaerts J, Schwartz LH, et al. New response evaluation criteria in solid tumors: Revised RECIST guideline version 1.1. *Eur J Can* 45: 228-47, 2009.
3. Guo W, Wang SJ, Yang S, et al. A Bayesian interval dose-finding design addressing Ockham's razor: mTPI-2. *Contemp Clin Trials*. 2017;58(23-33).

4. Ji Y, Wang S. Modified Toxicity Probability Interval Design: A Safer and More Reliable Method Than the 3+3 Design for Practical Phase I Trials. *Journal of Clinical Oncology* 31:1785-2792, 2013.
5. Kalbfleisch JD, Prentice, RL. *Statistical Analysis of Failure Time Data*, 2nd Edition. Hoboken, Wiley Interscience. 2002.
6. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 53: 457-81, 1958.
7. Moon J, LeBlanc M, Othus M. Accounting for All Patients in Waterfall Plots. *JCO Clinical Cancer Informatics* 5:414-420, 2021
8. Newcombe, R. G. (1998). "Two-sided confidence intervals for the single proportion: comparison of seven methods". *Statistics in Medicine.* **17** (8): 857–872.
9. Rosalind Rabin & Frank de Charro (2001) EQ-SD: a measure of health status from the EuroQol Group, *Annals of Medicine*, 33:5, 337-343
10. Wallis, Sean A. (2013). Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods (PDF). *Journal of Quantitative Linguistics.* **20** (3): 178–208.
11. Wilson, EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927; 22: 209–212.

9. APPENDICES

Appendix 1. Data Derivation Details

Enrollment	Date of assignment of the enrollment number
Lead-in period -7Day to -1Day	Vepdegestrant (ARV-471) only
Study Day 1 (Phase 1b portion)	Enrollment day to Phase 1b (escalating dose)
Study treatment start date	First date when a non-zero dose of study drug of respective treatment combination is administered
Day 1 (cycle start date; Phase 2 portion)	Day 1 of a cycle is every 28 days unless there is a dosing delay.
Cycle length (all but last cycle)	Cycle length is 28 days (previous cycle length may exceed planned length if there is a delay in study treatment administration).
Cycle X treatment delayed	If start of study treatment administration is delayed for Cycle X then Cycle X-1 is extended.
Last cycle	For participants off treatment, from Day 1 of final cycle to the last dose of study treatment. For participants on treatment, from Day 1 of most recent cycle start to protocol specified cycle length.
On-treatment period for TEAEs and other safety endpoints	From the date that the participant took the first dose of study treatment + 28 calendar days after the last dose of study treatment (including lead-in period)
Baseline value/assessment for safety (Phase 1b and Phase 2 portions)	Last available assessment prior to first dose of study treatment (including lead-in period)

Appendix 2. Statistical Methodology Details – SAS code

SAS code for 95% CI Wilson score method, (Wilson, 1927).

```
data Wilson;
r = x ;      /* x = input number of responses */
n = y;       /* y = input number of participants */
alpha = 0.05;
```

```
p = r / n;  
q = 1-p;  
z = probit (1-alpha/2);  
L = 100*(( 2*r+z**2 - (z*sqrt(z**2+4*r*q)) ) / (2*(n+z**2)));  
U = 100*(( 2*r+z**2 + (z*sqrt(z**2+4*r*q)) ) / (2*(n+z**2)));  
put r= L= U=;  
run;
```

090177e19fe3e49a\Final\Final On: 08-Feb-2024 19:23 (GMT)

Appendix 3 Statistical methodology details – Phase 1b dose determination rules

The mTPI-2 design uses a Bayesian statistics framework and a beta/binomial hierarchical model to compute the posterior probability of 3 dosing intervals (corresponding to under, proper, and overdosing in terms of toxicity) that reflect the relative difference between the toxicity rate of each dose level to the pT rate ($pT=0.30$). If the toxicity rate of the currently used dose level is far smaller than pT, the mTPI-2 will recommend escalating the dose level; if it is close to pT, the mTPI-2 will recommend continuing at the current dose; if it is far greater than pT, the mTPI-2 will recommend deescalating the dose level. These rules are conceptually similar to those used by the 3+3 design, except the decisions of an mTPI-2 design are based on posterior probabilities calculated under a coherent probability model. As shown by Ji and Wang, the mTPI design is more efficient and safer than the 3+3 design (Ji & Wang, 2013). They considered 42 scenarios to cover a wide range of practical dose response shapes and concluded that the 3 + 3 design was more likely to treat participants at toxic doses above the MTD and less likely to identify the true MTD than the mTPI design. For example, the 3 + 3 design exhibited a lower overall toxicity percentage than the mTPI-2 design in only 1 of 42 scenarios. Sometimes the decision rules in the original mTPI can be debated in practice. The mTPI-2, an extension of the mTPI, helps to solve these undesirable issues.

Decision rules are based on calculating the Unit Probability Mass (UPM) of 3 dosing intervals corresponding to under, proper, and overdosing in terms of toxicity. Specifically, the underdosing interval is defined as $[0, pT-e1]$, the overdosing interval $[pT+e2, 1]$, and the proper-dosing interval $[pT-e1, pT+e2]$, where $e1$ and $e2$ are small fractions. Based on the projected safety profile, $e1$ is selected as $\frac{CC1}{CC1+CC2}$, and $e2$ is selected as $\frac{CC1}{CC1+CC2}$. Therefore, the target interval for the DLT rate is $\frac{CC1}{CC1+CC2}$. The 3 dosing intervals are associated with 3 different dose-escalation decisions. The underdosing interval corresponds to a dose escalation (E), overdosing corresponds to dose de-escalation (D), and proper dosing corresponds to staying at the current dose (S). Given a dosing interval and a probability distribution, the UPM of that dosing interval is defined as the probability of a participant belonging to that dosing interval divided by the length of the dosing interval. The mTPI-2 design calculates the UPMs for the 3 dosing intervals, and the one with the largest UPM informs the corresponding dose finding decision, which is the dose level to be used for future participants. For example, if the underdosing interval has the largest UPM, the decision will be to escalate, and the next 3-4 participants will be treated at the next higher dose level. Simulations have demonstrated that the decision based on UPM is optimal in that it minimizes a posterior expected loss (ie, minimizes the chance of making a wrong dosing decision).

Select escalation/de-escalation algorithms for total number of participants treated at the current dose level (current and previous cohorts):

For example based on mTPI-2 decision rules in Appendix 2, with 6 participants treated at current dose level, and the current dose is the highest dose

0-1 DLT → based on the observed safety profile, Phase 1b enrollment may stop, and the highest dose will be declared as the MTD/potential RP2D;

2-3 DLTs → de-escalate;

4-6 DLTs → deescalate; consider current dose as intolerable.

With 6 participants treated at current dose level, and the current dose is the lowest dose:

0-1 DLT → escalate if the higher dose has not previously been considered as intolerable;

2-3 DLTs → new participants will be treated at the lowest dose unless the number of DLTs reaches the elimination boundary;

4-6 DLTs → consider current dose as intolerable

mTPI-2 Decision Rules

Number of Participants

	3	4	5	6	7	8	9	10	11	12	13	14	15
0	E	E	E	E	E	E	E	E	E	E	E	E	E
1	D	S	E	E	E	E	E	E	E	E	E	E	E
2	D	D	D	D	S	S	E	E	E	E	E	E	E
3	DU	DU	D	D	D	D	D	S	S	S	E	E	E
4		DU	DU	DU	D	D	D	D	D	D	S	S	S
5			DU	DU	DU	DU	DU	D	D	D	D	D	D
6				DU	DU	DU	DU	DU	DU	D	D	D	D
7					DU	DU	DU	DU	DU	DU	DU	D	D
8						DU	DU	DU	DU	DU	DU	DU	DU
9							DU	DU	DU	DU	DU	DU	DU
10								DU	DU	DU	DU	DU	DU
11									DU	DU	DU	DU	DU
12										DU	DU	DU	DU
13											DU	DU	DU
14												DU	DU
15													DU

E = Escalate to the next higher dose level

S = Stay at the current dose level

D = De-escalate to the next lower dose level

DU = The current dose level is unacceptably toxic and should be eliminated from further dose finding

***DLTs** = Dose Limiting Toxicities; Targeted DLT rate = 30%

Appendix 4 List of Abbreviations

Abbreviation	Term
A/MBC	Advanced or metastatic breast cancer
Abs	Absolute
AE	adverse event
ATC	Anatomic Therapeutic Chemical
AUC	area under the curve
AUC _{inf}	area under the concentration-time curve from time zero to infinity
AUC _{last}	area under the plasma concentration vs. time curve
AUC _τ	area under the concentration-time curve at steady state over the dosing interval
AUC _{ss,τ}	area under the concentration-time curve during a dosing interval τ at steady state
BLQ	below the limit of quantitation
BOR	best overall response
BP	blood pressure
CBR	Clinical benefit response
CDARS	Clinical Data Analysis and Reporting System (of US Food and Drug Administration)
CI	confidence interval
CL	clearance
C _{max}	maximum observed concentration
CR	complete response
CRF	case report form
CSR	clinical study report
C _{ss,max}	maximum observed plasma concentration during a dosing interval at steady state
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	circulating tumor deoxyribonucleic acid
DL	dose level
DI	Dose intensity
DLT	dose limiting toxicity
DoR	duration of response
eCRF	electronic case report form
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Quality of Life Instrument
EQ-5D-5L	EuroQol 5-dimension scale 5-level version
ESR1	Estrogen receptor 1
FSH	follicle stimulating hormone

FDA	Food and Drug Administration (United States)
GCP	Good Clinical Practice
GMC	geometric mean concentration
GMFR	geometric mean fold rise
GMR	geometric mean ratio
GMT	geometric mean titer
HR	Heart rate
ICD	informed consent document
ICH	International Council for Harmonisation
LLOQ	lower limit of quantitation
LOD	limit of detection
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum tolerated dose
mTPI	Modified toxicity probability interval
NA	not applicable
NCI	national cancer institute
ND	Not done
NE	Not evaluable
NS	No sample
OR	objective response
ORR	objective response rate
OS	overall survival
PD	pharmacodynamic(s)
PD	progressive disease
PFS	progression-free survival
PK	pharmacokinetic(s)
PR	Partial Response or time interval from the upper to lower chamber of the heart Receptor (depending on context)
PRO	patient-reported outcome
PT	prothrombin time
PT	preferred term
QD	once daily
QRS	QRS interval start at the end of PR interval (beginning of Q wave) followed to the end of the down ward (S wave), normal QRS interval is 0.08 to 0.10
QT	Time between the start of the Q wave and the end of the T wave in the heart
QTc	QT interval corrected for heart rate
QTcF	QT interval corrected using Fridericia formula
RDE	Recommended dose for expansion
RECIST	Response Evaluation Criteria in Solid Tumors
RP2D	recommended Phase 2 dose
RR	Interval between two successive ventricular depolarizations or between two
RDI	Relative dose intensity
SAE	serious adverse event

SAP	statistical analysis plan
SAS	Safety Analysis Set
SD	standard deviation or stable disease
SOC	System organ class
SOP	standard operating procedure
TEAE	treatment-emergent adverse event
T _{max}	maximum plasma concentration time
ULN	upper limit of normal

Appendix 5 Response Evaluation Criteria in Solid Tumors: RECIST version 1.1

Adapted from (Eisenhauer et al, 2009).

CATEGORIZING LESIONS AT BASELINE

All sites of disease must be assessed at screening as close as possible to the start of study intervention date (≤ 28 days prior to the start of study intervention unless otherwise specified in Section). If screening assessment is inadequate, subsequent statuses generally should be not evaluable.

Lesions are selected at screening and followed as either target or non-target lesions. All target lesions must be measurable. Any lesions or sites of disease not classified as target lesions including those that are non-measurable are classified as non-target. Definitions of measurable, non-measurable, target, and non-target lesions are provided below.

Measurable Lesions

Measurable lesions (including both nodal and non-nodal) are lesions that can be accurately measured as follows:

- Lesions with longest diameter ≥ 10 mm in the axial plane (bone lesions are not included except for soft tissue expansile masses arising from bone) when assessed by CT or MRI
- Lesions with longest diameter ≥ 20 mm when assessed by chest X-ray.
- Superficial lesions with longest diameter ≥ 10 mm when assessed by caliper
- Malignant lymph nodes with the short axis ≥ 15 mm. Nodal size is normally reported as 2 dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the smaller of the axial, sagittal or coronal will be considered the short axis).

NOTE: Normal nodes: Nodes with short axis < 10 mm are considered normal and should not be recorded or followed either as measurable or non-measurable disease.

Non-measurable disease

All other disease including lesions too small to be considered measurable, pleural or pericardial effusions, ascites, bone disease, inflammatory breast disease, leptomeningeal disease, lymphangitic involvement of skin and lung, clinical lesions that cannot be accurately measured with calipers, abdominal masses identified by physical exam that are not measurable by reproducible imaging techniques, nodes with short axis ≥ 10 mm but < 15 mm, disease documented by indirect evidence only (eg, by lab values), or previously radiated lesions that have not progressed, are considered non-measurable.

Target Lesions

All measurable lesions up to a maximum of 2 lesions per organ, 5 lesions in total, representative of all involved organs, should be identified as target lesions at baseline. Target lesions should be selected on the basis of size (lesions with longest diameter for non-lymph nodes ≥ 10 mm and nodes with longest short axis ≥ 15 mm) and suitability for accurate repeated measurements. Measurements are recorded in mm and must be provided for target lesions.

Non-Target Lesions

All non-measurable lesions are non-target. All measurable lesions not identified as target lesions are also included as non-target lesions.

OBJECTIVE STATUS AT EACH EVALUATION

Response is assessed by the investigators, with objective status at each assessment recorded on the Investigator Overall Objective Tumor Assessment (IOTA) CRF.

Objective Status is to be recorded at each post-baseline evaluation.

Sites of disease should be assessed using the same technique as baseline, including consistent administration of contrast agents (see Section). If a change needs to be made the case should be discussed with the radiologist and the sponsor to determine if substitution is possible. If not, subsequent objective statuses are not evaluable.

If target lesions were present at screening, screening sum of diameters is calculated as the total sum of the longest diameters for non-nodal lesions and short axis for nodal lesions. Sum of diameters at subsequent assessments can be calculated in the same manner using the same set of target lesions.

Response evaluation of target lesions

- **CR:** Complete disappearance of all target lesions, with the exception of nodal disease. All nodes, must decrease to normal (short axis < 10 mm). All target lesions must be assessed.

- **PR:** At least a 30% decrease from baseline in the sum of diameters of all target lesions. The short diameter is used in the sum for nodal target lesions, while the longest diameter is used in the sum for non-nodal target lesions. All target lesions must be assessed.
- **PD:** At least a 20% increase in the sum of diameters of target measurable lesions above nadir (smallest sum observed considering baseline and all assessments prior to the timepoint under evaluation), with a minimum absolute increase of 5 mm relative to nadir. If only a subset of target lesions is assessed and the sum of the non missing lesion diameters results in an increase of at least 5 mm and at least 20% above the nadir then the progression criteria will have been met.
- **Stable Disease (SD):** Does not qualify for CR, PR, PD. All target lesions must be assessed.
- **Not Evaluable (NE):** PD has not been documented, and at least one of the following is true:
 - one or more target lesions have not been assessed
 - assessment methods used were inconsistent with those used at baseline (unless considered “interchangeable”)
 - one or more target lesions cannot be measured accurately unless due to being too small to measure
 - one or more target lesions were excised or irradiated and have not reappeared or increased.

General considerations on Response Evaluation of Target lesions:

- A target lesion that is irradiated or excised post-screening is no longer measurable and generally should be assessed as Not Evaluable. Increase or reappearance of such lesions should constitute progression.
 1. A target lesion biopsied during the screening period will NOT be considered as a measurable lesion.
 2. Participants with bone disease only may have measurable disease if the following applies: lytic or mixed lytic-blastic lesion that can be accurately assessed by CT or MRI (ie, with soft tissue component that can be evaluated by CT or MRI and meet the definition of measurability as per RECIST v1.1).
- For lesions that become necrotic in the center, the longest diameter should still be used in the sum of diameters.

- Measurements for target lesions that become small should continue to be recorded;
- if the lesion is considered to have disappeared, 0 mm should be recorded;
- if a lesion is determined to be present but too small to measure (ie, lesion less than 10 mm with an exact measurement that cannot be performed), the lesion status will indicate “too small to measure” and 5 mm will be used in the calculation of the sum of the diameters.
- If 2 target lesions coalesce, the measurement of the coalesced mass is used. If a large target lesion splits, the sum of the parts is used.
- If a subset of the target lesions identified at baseline were not measured post-baseline, the objective status of target lesions at the visit is Not Evaluable, unless the partial sum of diameters for the subset of target lesions assessed at the visit already meets the PD criteria, ie, at least 5 mm and 20% increase above the smallest sum of diameters observed, in which case the objective status of target lesions at the visit is PD.
- If a lesion disappears and reappears at a subsequent time point, it should continue to be measured. The participant’s response at the point in time when the lesion reappears will depend upon the status of his/her other lesions and previous response. For example, if the participant’s tumor had reached a CR status and a non-lymph node target lesion reappeared, then the participant’s response would be considered PD at the time of reappearance.
- In contrast, if the tumor status was a PR or SD and 1 lesion which had disappeared then reappears, its maximal diameter should be added to the sum of the remaining lesions for a calculated response.

Response evaluation of non-target lesions

- **CR/Absent:** Complete disappearance of all non-target lesions. All nodes, must decrease to normal (short axis <10 mm). All disease sites must be assessed.
- **Non-CR/Non-PD/Present without unequivocal progression:** Persistence of 1 or more non-target lesion(s).
- **PD/Unequivocal progression:** Unequivocal progression of existing non-target lesions.
- **Not All Evaluated/Inevaluable:** Progression has not been documented and non-target lesions are not all assessable or disease assessed with inconsistent methods (unless considered “interchangeable”) or one or more target lesions were excised or irradiated and have not reappeared or increased.

Note: there is no standard RECIST method of using non-target disease assessments to determine unequivocal progression in non-target disease; by the nature of non-measurable disease, this is not possible. Unequivocal progression is based on clinical judgment.

New lesions evaluation criteria

If a lesion is identified in a location not scanned previously, this will be considered progression of disease (new lesion).

If a new lesion is equivocal at one assessment (classified as “Equivocal”) and verified at the next, (classified as “Unequivocal”), the earlier date is used as the progression date.

OR Status at each assessment

Assessment is based on the target and non-target lesions response as described above and in the table below.

Objective Response Status at Each Assessment for Participants With Measurable Disease at Baseline

Target Lesions	Non-Target Lesions	New Lesions	Objective Status
CR	CR	No	CR
CR	Non-CR/Non-PD or not all evaluated	No	PR
PR	Non-PD* or not all evaluated	No	PR
SD	Non-PD* or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes**	PD

*Non-PD includes CR and Non-CR/Non-PD

** New lesions must be unequivocal

Note: cases for which target disease is assessed as SD or better but non-target assessments are missing, must be reviewed carefully. By considering the whole sequence, it may be reasonable to assume missed non-target assessments were not progression.