Statistical analysis plan	
Study code	<<D5161R00037>>
Version	<<2.0>>
Date	<<17-05-2024>>

Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multi-center Study (GPS)

Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multi-center Study (GPS)

A□ Study Statistician PPD

PPD Date

Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multi-center Study (GPS)

A□ Medical Advisor

PPD

PPD

Date

Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multi-center Study (GPS)

A□ Medical Advisor

PPD

PPD

Date

Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multi-center Study (GPS)

A□ Global Statistician PPD

PPD Date

TABLE OF CONTENTS PAGE

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

Abbreviation or special term	Explanation
CI	Confidence Interval
CRF	Case report form
CSP	Clinical Study Protocol
CSR	Clinical Study Report
CT	Computed Tomography
ctDNA	circulating tumour DNA
DCO	Data Cutoff
DNA	Deoxyribonucleic Acid
DSUR	Development Safety Update Report
EC	Ethics Committee, synonymous to Institutional Review Board (IRB) and Independent Ethics Committee (IEC)
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
EGFR	Epidermal Growth Factor Receptor
FAS	Full Analysis Set
FDA	Food and Drug Administration
FFPE	Formalin Fixed and Paraffin Embedded
FISH	Fluorescent in Situ Hybridization
GCP	Good Clinical Practice
HIV	Human Immunodeficiency Virus
HR	Hazard Ratio
IB	Investigator's Brochure
ICF	Informed Consent Form
ICH	International Council for Harmonisation
IEC	Independent Ethics Committee
IHC	Immunohistochemistry
IRB	Institutional Review Board
MRI	Magnetic Resonance Imaging
NE	Not Evaluable
NL	New Lesion
NSCLC	Non-Small Cell Lung Cancer
NTL	Non-Target Lesion
ORR	Objective Response Rate
OS	Overall Survival
PD	Progression of Disease
PET	Positron Emission Tomography
PFS	Progression-Free Survival
PR	Partial Response

Abbreviation or special term	Explanation
RECIST 1.1	Response Evaluation Criteria in Solid Tumors, Version 1.1
SAP	Statistical Analysis Plan
SD	Stable Disease
SoA	Schedule of Activities
TKI	Tyrosine Kinase Inhibitor
TL	Target Lesion
TOC	Table of Contents
WHO	World Health Organisation
ddPCR	Droplet Digital PCR

AMENDMENT HISTORY

Date	Brief description of change
16-11-2023	Initial version
17-05-2024	Updated to V2.0 with following changes: 1. Added overall percent agreement as a measurement of concordance of NGS in tissue and plasma for secondary endpoint analysis. [redacted]

1. OBJECTIVES

1.1 Objectives

This statistical analysis plan is for AstraZeneca "Gene Profile in EGFRm Locally Advanced or Metastatic NSCLC patients post Osimertinib 1L treatment failure: A real-world, multicenter Study (GPS)" (Study code: D5161R00037) and will give a detailed description of the planned analysis. This version of SAP is for interim analysis and final analysis. Tigermed is responsible for baseline characteristics, Geneplus is responsible for gene related analysis. CCI [redacted] CCI [redacted]

This statistical analysis plan is based on the 2.0 version of the research protocol on January 20, 2023, the 3.0 version of the Case Report Form on Dec 27, 2023.

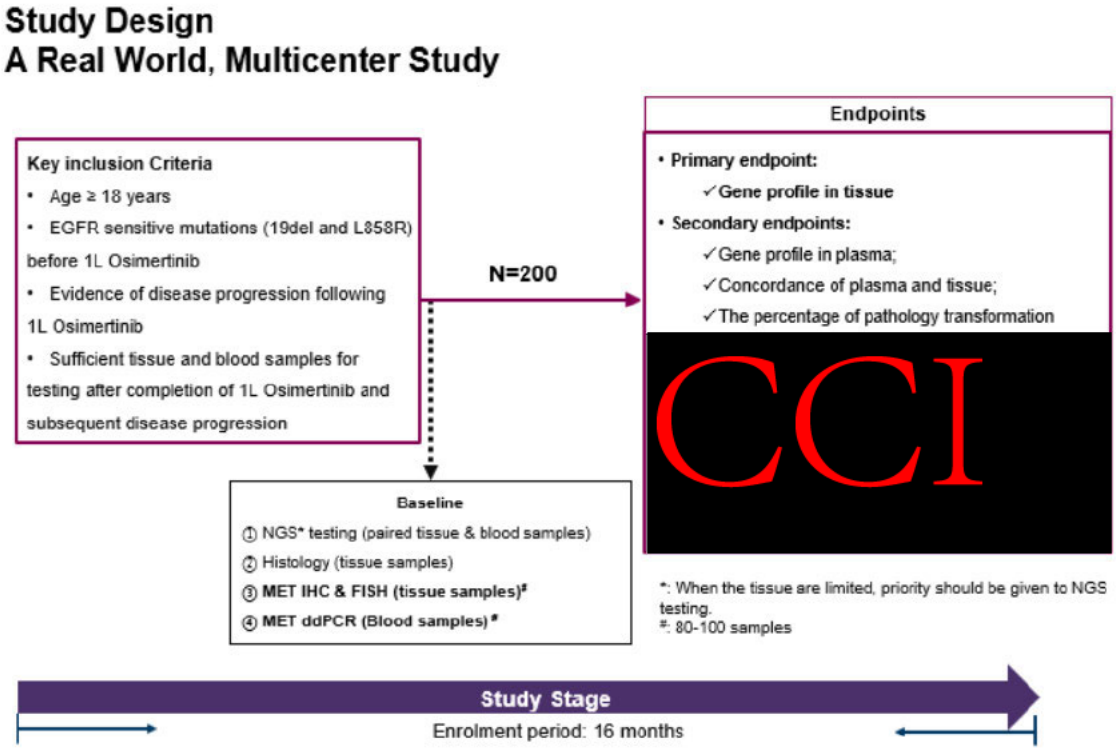
Objectives and Endpoints

Objectives	Endpoints
Primary	
To describe the gene profile in *EGFRm* advanced NSCLC patients post Osimertinib 1L treatment.	● Gene profile in tissue Gene profile defined as gene alternation detected by NGS and frequency of every gene alternation. All gene alterations include gene mutation, copy number variation(CNV), fusion, etc. The frequency of all gene alteration detected by NGS(%) = (number of patients with every gene alteration detected by NGS)/(total number of patients in the FAS)×100%.
Secondary	
To assess the concordance of plasma and tissue.	● Gene profile in plasma Gene profile defined as gene alternation detected by NGS and frequency of every gene alternation. ● Concordance of plasma and tissue Using selected representative genes (for example EGFR, MET, PIK3CA and etc.) to measure the concordance Concordance of Gene X in plasma and tissue is defined as sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) between plasma and tissue results., Tissue sample is the reference standard. - sensitivity=(number of patients with positive result in both plasma and tissue)/(total number of patients with positive result in tissue samples)×100%; - specificity=(number of patients with negative result in both plasma and tissue)/(total number of patients with negative result in tissue samples)×100%. - PPV (%)=(number of patients with positive result in both plasma and tissue)/(total number of patients with positive result in plasma samples)×100%; - NPV (%)=(number of patients with negative result in both plasma and tissue)/(total number of patients with negative result in plasma samples)×100%; ● The percentage of pathology transformation

	Pathology transformation is defined as those transformation from non-small-cell lung cancer to small-cell lung cancer or from adenocarcinoma to squamous carcinoma, can be observed by IHC
CCI	
CCI	CCI

2. STUDY DESIGN AND SAMPLE SIZE

2.1 Schema



RECIST 1.1=Response Evaluation Criteria in Solid Tumors, Version 1.1.
NSCLC, non-small cell lung cancer; *EGFR*, epidermal growth factor receptor.

Figure 1 Study Design

2.2 Sample size

Approximately 200 participants in EGFRm locally advanced or metastatic NSCLC with the evidence of disease progression following 1L Osimertinib will be enrolled and 15 sites will be involved to elucidate the gene profile in paired tissue and plasma so that approximately 180 participants whom would be get gene profile, assuming a 10% drop-out.

Assuming some common gene alteration with an incidence of 20%, then with 200 subjects the 95% CI will be around 14.7% - 26.2% based on an exact method (Clopper-Pearson).

Table below summarizes the 95% CIs for the target incidences ranging from 10% to 50% with sample size of 200.

Assumed incidence of gene mutation	95% CI
10%	6.2%, 15.0%
20%	14.7%, 26.2%
30%	23.7%, 36.9%
40%	33.2%, 47.1%
50%	42.9%, 57.1%

In addition, 200 subjects are enough to have 86% probability to observe at least one case if the incidence of an gene alteration is 1%.

Note: "Enrolled" means a participant's, or their legally acceptable representative's, agreement to participate in a clinical study following completion of the informed consent process. Potential participants who are screened for the purpose of determining eligibility for the study, but are not randomly assigned/assigned in the study, are considered "screen failures", unless otherwise specified by the protocol.

Human Biological Sample Analysis:
Participants will be required to provide paired tissue and whole blood after disease progression following 1L Osimertinib. 200 tissue samples and 200 whole blood samples will be used to detect gene alteration by NGS, respectively. 200 tissue samples will be used to detect pathological transformation by IHC. Approximately 80-100 tissue samples will be used to test MET overexpression by MET IHC and MET amplification by FISH respectively. Approximately 80-100 whole blood samples will be used to test MET amplification by ddPCR.

3. ANALYSIS SETS

3.1 Definition of analysis sets

The following populations are defined:

Table 1 Populations for Analysis

Population/Analysis Set	Description
Enrolled	All participants who sign the ICF
FAS	All participants who meet the required inclusion/exclusion criteria and have valid Gene data from both plasma and tissue. CCI [redacted]

3.2 Violations and Deviations

The list of protocol deviations will be finalized and documented before the database lock.

Protocol deviations will be collected, reviewed and reconciled throughout the study. Important protocol deviations (IPDs) will be identified from the complete set of protocol deviations. IPDs are those which may significantly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect a patient's rights, safety, or wellbeing.

A set of pre-determined IPDs are listed in the section 'Protocol Deviation Categories' in the protocol deviations tracking log.

The IPDs are grouped in to the following PD categories, where full details of the individual IPDs within each IPD category are provided in the protocol deviations plan:

- Informed consent problem.
- Inclusion/exclusion criteria.
- Lack of labs/other procedures.
- Visit schedule.
- Others

4. EXPOSURE(S) AND OUTCOMES

4.1 Exposures

Not applicable.

4.1.1 Definition of primary drug exposure

Not applicable.

4.1.2 Definition of comparison drug exposure

Not applicable.

4.2 Outcomes

4.2.1 Safety

Not applicable.

4.2.2 Primary outcome(s)

Gene profile in tissue
Gene profile is defined as gene alteration detected by NGS and frequency of every gene alteration. All gene alterations include gene mutation, copy number variation(CNV), fusion, etc.

4.2.3 Secondary outcome(s)

Gene profile in plasma
Gene profile is defined as gene alternation detected by NGS and frequency of every gene alternation.

Proportion of Gene X alteration(%) = (number of patients with Gene X alteration, which is an example of a certain gene)/(total number of patients in the FAS)×100%.

Concordance of plasma and tissue

Using selected representative genes (for example EGFR, MET, PIK3CA and etc.) to measure the concordance of plasma and tissue mutation for each gene. Concordance of Gene X in plasma and tissue is defined as sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) between plasma and tissue results.

The percentage of pathology transformation

Pathology transformation is defined as those transformation from non-small-cell lung cancer to small-cell lung cancer or from adenocarcinoma to squamous carcinoma, can be observed by IHC.

CCI [redacted]

4.3 Other variables and covariates

4.3.1 Baseline characteristics

Baseline information consists of :

Demography:

Age, sex, race, ethnicity,

Vital signs:

Weight, height, BMI

Smoking history:

Never, current, former usage, number of pack years

Surgical History:

Surgical history will be coded using MedDRA codes of the version currently in effect at the time of database lock. The number (percentage) of patients reporting a surgical history will be summarized by system organ class and preferred term.

Cancer Therapy – Osimertinib:

Duration of Osimertinib treatment, dose, dose adjustment, cause of dose adjustment, dose interruption, cause of dose interruption, dose termination, cause of dose termination, best response.

ECOG Performance Status:

Normal activity, restricted activity, in bed less than or equal to 50% of the time, in bed more than 50% of the time, 100% bedridden, death.

Pathology At Time of Diagnosis:

Histology type, primary tumour, regional lymph nodes, distant metastases.

Extent of Disease At Time of Diagnosis:

Metastatic/Locally Advanced, location, tumor number

Pathology Upon Entry to Study:

Histology type, primary tumour, regional lymph nodes, distant metastases.

Extent of Disease Upon Entry to Study:

Metastatic/Locally Advanced, location, tumor number

Multi Gene Status At Time of Diagnosis:

Sample type, test type, gene name and variant classification

Medical history:

Medical history will be coded using MedDRA codes of the version currently in effect at the time of database lock. The number (percentage) of patients reporting a medical history will be summarized by system organ class and preferred term.

Nicotine Use

Classification and pack years

5. ANALYSIS METHODS

5.1 General Aspects

All statistical analyses will be completed using SAS version 9.4 or later, and R 4.0 or later. Descriptive statistics will be provided for all variables, as appropriate. Continuous variables will be summarized by the number of observations, mean, standard deviation, median, interquartile range (Q1, Q3), minimum, and maximum. Categorical variables will be summarized by frequency counts and percentages for each category. The 95% confidence interval (CI) will be calculated as appropriate.

The analyses will be based on Full analysis set (FAS), which will include all patients who meet the required inclusion/exclusion criteria and have valid Gene data from both plasma and tissue.

Patients missing tissue or plasma NGS data or MET data will not be used for concordance analysis, but only for individual tissue or plasma analysis.

5.1.1 Handling of Missing Data

Missing data will not be imputed unless otherwise specified elsewhere in the SAP.

5.2 Subject disposition

The subject disposition will be summarized. Screened, enrolled, screen failure, study completion, study discontinuation will be summarized and the reasons for failure and study discontinuation will be listed.

5.3 Baseline characteristics

Analysis will be performed based on FAS.

Descriptive statistics will be provided for all variables, as appropriate.

Baseline information consists of demography(age, sex, race, ethnicity, weight, height, BMI), smoking history , medical history, ECOG/WHO performance status, histology, clinical stage with lymph node and distant metastasis at diagnosis and post 1L Osimertinib, duration of 1L Osimertinib treatment, best response of 1L Osimertinib.

Age will be divided to two categories: <65 years, >= 65 years.

Duration of Osimertinib treatment:

If date is complete, then Duration of Osimertinib treatment (Months) = (End Date – Start Date)/365.25*12, keep one decimal places.

If either day of start date or end date is missing, then Duration of Osimertinib treatment (Months) = (Year of End Date – Year of Start Date)*12 + (Month of End Date – Month of Start Date), where the Year and Month of Start Date and End Date will be extracted from the corresponding date.

If missing day and month, then Duration of Osimertinib treatment (Months) will be considered as missing.

5.4 Biomarker analysis

5.4.1 Primary Endpoint

The primary objective is to describe the gene profile in EGFRm advanced NSCLC patients after progression on 1L Osimertinib treatment.

Gene profile in tissue will be analyzed by NGS method. All gene alterations in tissue(e.g., gene mutation, amplification, deletion, fusion, etc.) detected by NGS will be observed. Proportion of Gene X alteration(%) = (number of patients with Gene X alteration, which is an example of a certain gene)/(total number of patients in the FAS)×100%. Some cases which patients have more than one gene alteration need to be considered, therefore percentages may not add up to 100%.

The OncoPrint plot will be used to visualize the gene alterations across all cases and reveal trends, such as co-occurrence and mutual exclusivity.

5.4.2 Secondary Endpoint

The secondary objective is to assess gene profile in plasma, concordance of plasma and tissue and the percentage of pathology transformation will be analysed by NGS method and IHC. All gene alterations in plasma(e.g., gene mutation, amplification, deletion, fusion, etc.) detected will be observed. Proportion of Gene X alteration(%) = (number of patients with Gene X alteration, which is an example of a certain gene)/(total number of patients in the FAS)×100%. Some cases which patients have more than one gene alteration need to be considered, therefore percentages may not add up to 100%.

Using selected representative genes to measure the concordance. The representative genes should include the top 20 mutated genes observed in tissue samples. Concordance of Gene X in plasma and tissue is defined as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall percent agreement between plasma and tissue results., Tissue sample is the reference standard.

Sensitivity=(number of patients with positive result in both plasma and tissue)/(total number of patients with positive result in tissue samples)×100%;

Specificity=(number of patients with negative result in both plasma and tissue)/(total number of patients with negative result in tissue samples)×100%.

PPV (%)=(number of patients with positive result in both plasma and tissue)/(total number of patients with positive result in plasma samples)×100%;

NPV (%)=(number of patients with negative result in both plasma and tissue)/(total number of patients with negative result in plasma samples)×100%.

Overall percent agreement(%) =(number of patients with same result <positive or negative> in both plasma and tissue)/(total number of patients in FAS)×100%

Pathology transformation defines as pathological type of lung cancer change, which will be confirmed by pathologists. Number and percentage of patients with pathology transformation will be summarized. Proportion of pathology transformation(%) = (number of patients with pathology transformation)/(total number of patients in the FAS)×100%.

CCI [redacted]

CCI [redacted]

[redacted]

[redacted]

		[redacted]		[redacted]
		[redacted]	[redacted]	
[redacted]	[redacted]			
	[redacted]			
[redacted]				

[redacted]

[redacted]

[redacted]

- [redacted]
- [redacted]

CCI

[redacted]

[redacted]

CCI

□. TIMING OF ALL ANALYSES

Two analyses are planned. The first interim analysis data cut-off (DCO) point will happen when approximately 100 patients have completed their gene testing in tissue and whole blood and got their testing results. The first interim analysis will focus on the primary outcome and secondary outcomes. The final analysis is planned when the last patient has completed their gene testing in tissue and whole blood and got their testing results. The primary outcome, secondary outcomes CCI will be updated using the data of all patients.

□. CHANGES TO PROTOCOL PLANNED ANALYSES

Changes to protocol planned analyses are summarized as following:

- Added overall percent agreement as a measurement of concordance of NGS in tissue and plasma for secondary endpoint analysis.

[redacted]

□. APPENDIX

□.1 Schedule of Activities

Table 1 Schedule of Activities

Procedure	Screening and Data Collection	Notes	Details in CSP Section or Appendix
Visit	**1**		
Informed consent [a]	X		Sections 5.1
Inclusion and exclusion criteria	X		Section 5.1 and 5.2
Histology	X		
Clinical stage with lymph node and distant metastasis at diagnosis and post 1L Osimertinib	X		
Record documented common EGFR mutation results by an accredited laboratory through ARMS, Super-ARMS, and NGS testing [b]	X		
Demography and smoking history [c]	X		Section 7.2.1
Physical examination and weight(Kgs)	X		Section 7.2.1
Height(CMs)	X		Section 7.2.1
Medical history and comorbid conditions [d]	X		Section 5.1 and 5.2
WHO/ECOG performance status post Osimertinib 1L treatment	X		Section 7.2.2

Duration of 1L Osimertinib treatment [e]	X		
Best response of 1L Osimertinib	X		
Progression pattern post 1L Osimertinib	X		
Mandatory screening tumour tissue sample (archived or newly acquired biopsy) for Central NGS testing and central MET FISH testing; local IHC testing (included MET overexpression by MET IHC and SCLC pathology transformation by IHC)	X		Section 7.3.1
Mandatory screening blood sample (newly acquired; 2*10ml) for Central NGS testing and central ddPCR MET amplification testing	X		Section 7.3.1 and Appendix C

a Written informed consent and any locally required privacy act document authorisation must be obtained prior to performing any protocol-specific procedures, including screening/baseline evaluations.
b Participants are eligible to be considered for inclusion on the basis of local pre-existing EGFR tumour tissue or plasma testing results by molecular testing.
c Include date of birth or age, gender, race, and ethnicity for all participants.
d Include any treatment history before enrolling the study, complication and related history (chronic diseases which require treatment currently).
e Include the start and end dates of 1L Osimertinib treatment; starting dose,
Note: Data collection following study analysis until the end of the study is described in Section 7.
Note: The screening visit will confirm that the patients will be receiving a biopsy (scheduled for after enrolment), but they will not actually receive a biopsy at the screening visit.
CSP=Clinical Study Protocol; ECOG=Eastern Cooperative Oncology Group; ICF=informed consent form; WHO=World Health Organisation

9. REFERENCES

FDA Guidance for Industry (issued July 2009) 'Drug-induced liver injury: Premarketing clinical evaluation':

http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM174090.pdf