NON-INTERVENTIONAL (NI) STUDY PROTOCOL

PASS information

Title	Real-World Incidence and Prevalence of Brain Metastasis and Its Impact on Clinical Outcomes among Anaplastic Lymphoma Kinase (ALK) Positive Metastatic Non-small Cell Lung Cancer (mNSCLC) Patients Receiving ALK Tyrosine Kinase Inhibitors (TKIs) as First-Line (1L) Treatment in the United States (US)
Protocol number	A8081077
Protocol version identifier	1.0
Date	22 November 2024
EU Post Authorization Study (PAS) register number	EUPAS1000000370
Active substance	Crizotinib
Medicinal product	Xalkori®
Joint PASS	No
Research question and objectives	The primary research question of this study is to assess the prevalence and incidence of BM among patients with ALK+ mNSCLC who received 1L treatment with first-generation (1G) or second-generation (2G) ALK TKIs, and the impact of BM on clinical outcomes (overall survival [OS], real-world progression-free survival [rwPFS] and time to treatment discontinuation [TTD]). More specifically, the study objectives include: **Primary objectives:** 1. To describe the demographic and clinical characteristics of patients with ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs, both overall and stratified by baseline BM status 2. To assess the real-world incidence and prevalence of BM among patients with

	ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs 3. To assess the impact of baseline BM and incident BM on OS (mortality) among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs 4. To assess the impact of baseline BM and incident BM on rwPFS among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs 5. To assess the impact of baseline BM and incident BM on TTD among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs **Secondary objective**: 1. To stratify the above analyses by generation of ALK TKI (1G: crizotinib; 2G: ceritinib, alectinib, brigatinib)
Country(ies) of study	United States
Author	Redacted

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
1G	First-generation
1L	First-line
2G	Second-generation
3G	Third-generation
AE	Adverse event
ALK	Anaplastic lymphoma kinase
BM	Brain metastases
CHF	Congestive heart failure
CI	Confidence interval
CIF	Cumulative incidence function
COPD	Chronic obstructive pulmonary disease
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic health record
HIPAA	Health Insurance Portability and Accountability Act
HR	Hazard ratio
ICD	International Classification of Diseases
IEC	Independent ethics committee
IQR	Interquartile range
IRB	Institutional Review Board
LOT	Line of therapy
mNSCLC	Metastatic non-small cell lung cancer
NR	Not reached
NSCLC	Non-small cell lung cancer
OS	Overall survival
PFS	Progression-free survival
PVD	Peripheral vascular disease
rwPFS	Real-world progression-free survival
SD	Standard deviation

TKI	Tyrosine kinase inhibitor
US	United States

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
Redacted	Redacted	Redacted	Redacted

Other Pfizer Authors

Name, degree(s)	Job Title	Affiliation	Address
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted

External Authors

Name, degree(s)	Job Title	Affiliation	Address
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted
Redacted	Redacted	Redacted	Redacted

4. ABSTRACT

Title: Real-World Incidence and Prevalence of BM and Its Impact on Clinical Outcomes among ALK Positive mNSCLC Patients Receiving ALK TKIs as 1L Treatment in the US

Version: 1

Date of Protocol: 22 November 2024

Author: Redacted

Rationale and Background: Lung cancer is the second most frequent malignancy and the leading cause of cancer-related death worldwide. Approximately 85% of lung cancer cases are non-small cell lung cancer (NSCLC), and among these, around 5% of patients have an ALK mutation. To date, five ALK inhibitors have been approved in the US, including 1G crizotinib, 2G ceritinib, alectinib, brigatinib, and 3G lorlatinib. Despite the advancement achieved newer generations of ALK inhibitors, a high unmet need remains among ALK+ metastatic NSCLC (mNSCLC) patients, especially for patients with BM. A US-based study using Redacted data (2011-2018) has reported a dismal prognosis for ALK+ mNSCLC patients, with a rwPFS of less than 8 months for both 1G and 2G ALK inhibitors. The PFS was even shorter for patients with baseline BM. More recent data is needed to understand the current prevalence of BM among ALK+ mNSCLC patients treated with ALK TKIs and the impact of both baseline and incident (newly developed) BM on clinical outcomes of 1G and 2L ALK inhibitors.

Research question and objectives: The primary research question of this study is to assess the prevalence and incidence of BM among patients with ALK+ mNSCLC who received 1G or 2G ALK TKI as 1L treatment, and the impact of BM on clinical outcomes OS, rwPFS and time to treatment discontinuation [TTD]). More specifically, the study objectives include:

- Primary objectives:
  1. To describe the demographics and clinical characteristics of patients with ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs, both overall and stratified by baseline BM status
  2. To assess the real-world incidence and prevalence of BM among patients with ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs
  3. To assess the impact of baseline BM and incident BM on OS (mortality) among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs
  4. To assess the impact of baseline BM and incident BM on rwPFS among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs
  5. To assess the impact of baseline BM and incident BM on TTD among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs
- Secondary objective:

1. To stratify the above analyses by generation of ALK TKI (1G: crizotinib; 2G: ceritinib, alectinib, brigatinib)

Study Design: This is a retrospective, observational study using the Redacted electronic health record (EHR) database. A cohort of adult patients with ALK+ mNSCLC patients who received 1L treatment with 1G or 2G ALK TKIs will be identified. The initiation date of either 1G or 2G ALK TKI as 1L treatment will be defined as the ***index date***. The ***baseline period*** will be defined as either 6 months prior to the index date or the period between mNSCLC diagnosis and the index date, whichever is longer. During this period, patient characteristics and baseline BM will be assessed. The ***follow-up period*** will be defined as the time from the index date to earliest of death, the last confirmed activity date or the data cutoff date (i.e., 30/11/2023). The ***index treatment period*** will be defined as the time from the index date to earliest of index treatment discontinuation or end of follow-up. Incident BM will be assessed during both the index treatment period and the follow-up period. The cumulative prevalence of BM will be assessed across the combined baseline and follow-up periods. OS will be evaluated during the follow-up period, while rwPFS and TTD will be evaluated during the index treatment period.

Population: The study population will include ALK+ mNSCLC patients who received 1L treatment with either 1G or 2G ALK TKIs in the Redacted database.

Variables: Covariates include baseline demographic and clinical characteristics include age, gender, race/ethnicity, smoking history, geographic regions, site of care, index year, histology, tumor stage, Eastern Cooperative Oncology Group performance score (ECOG PS), time from mNSCLC diagnosis to index, site of metastasis and individual comorbidities. Outcomes include incidence of BM during the index treatment period and follow-up period, cumulative prevalence of BM during the follow-up period, OS during the follow-up period, and rwPFS and TTD during the index treatment period. Exposure includes 1L treatment with 1G or 2G ALK TKIs.

Data Sources: Redacted EHR database (data cutoff date: 30 November 2023).

Study Size: All eligible patients available for the analysis will be included.

Data Analysis: For Primary Objective 1, patient baseline characteristics will be described in the overall sample and stratified by baseline BM status. Means, SD, medians, and interquartile range will be estimated for continuous baseline variables; counts and percentages will be estimated for categorical baseline variables.

For Primary Objective 2, prevalence and incidence BM will be evaluated using four metrics: prevalence of BM at baseline, incidence of BM during the index treatment period and the follow-up period, respectively, among those without BM at baseline, and cumulative prevalence during the baseline and follow-up periods. Cumulative incidence function (CIF) will be used to estimate the incidence and cumulative prevalence.

For Primary Objectives 3, 4 and 5, to assess the impact of baseline BM on OS, rwPFS and TTD, Kaplan-Meier analysis will be conducted to describe OS during the follow-up period and rwPFS and

TTD during the index treatment period, overall and by baseline BM status. Adjusted comparisons of OS, rwPFS and TTD between patients with and without baseline BM will be conducted using Cox proportional hazards regression adjusted for key baseline patient characteristics. To assess the impact of incident BM on OS, rwPFS and TTD, Cox regression models with time-varying covariates will be employed, which will incorporate key baseline characteristics as time-fixed covariates and BM status and ECOG PS as time-varying covariates. HR with 95% confidence intervals (CIs) from the Cox models will be reported.

For the Secondary objective, all the analyses above will be repeated among ALK+ mNSCLC patients receiving 1L treatment with 1G ALK TKI (i.e., crizotinib) as one group, and 2G ALK TKIs (i.e., ceritinib, alectinib, brigatinib) as another group, pending sufficient sample size.

Milestones

Milestone	Planned Date
Development of study protocol	22 November 2024
Start of data collection	09 December 2024
End of data collection	31 December 2024
Registration in the HMA-EMA Catalogues of RWD studies	02 December 2024
Completion of statistical analyses	31 January 2025
Final study report	31 March 2025

5. AMENDMENTS AND UPDATES

None.

6. MILESTONES

Milestone	Planned Date
Development of study protocol	22 November 2024
Start of data collection	09 December 2024
End of data collection	31 December 2024
Registration in the HMA-EMA Catalogues of RWD studies	02 December 2024
Completion of statistical analyses	31 December 2024
Final study report	31 March 2025

7. RATIONALE AND BACKGROUND

Lung cancer is the second most frequent malignancy and the leading cause of cancer-related death worldwide. In 2024, there were an estimated 234,580 new cases of lung cancer in the US, and the estimated death due to lung cancer is 125,070.[1] About 85% of patients with lung cancer have NSCLC, and among them roughly 5% have ALK mutation.[2]

Five ALK TKIs have been approved for the treatment of adult patients with ALK+ metastatic NSCLC (mNSCLC) in the US, including first-generation (1G) crizotinib, 2G ceritinib, alectinib, brigatinib, and 3G lorlatinib. A network meta-analysis based on summary data from randomized controlled trials of ALK TKIs has found that all ALK TKIs improved progression-free survival (PFS) compared to chemotherapy.[3] Alectinib and brigatinib have both shown superiority against crizotinib as 1L therapy among patients with ALK+ mNSCLC.[4,5] Another network meta-analysis comparing 1G, 2G and 3G ALK TKIs has found lorlatinib improves PFS compared with other ALK TKIs, with HRs ranging from 0.54 to 0.63.[6] This finding is supported by the recently published 5-year data from the Phase III CROWN trial, which has shown that after a 5-year follow-up, the median PFS has not been reached (NR) with lorlatinib (95% CI, 64.3 to NR), while the median PFS with crizotinib is 9.1 months (95% CI, 7.4 to 10.9; HR, 0.19 [95% CI, 0.13 to 0.27]).[7] These results highlight the increased clinical benefit associated with successive generations of ALK TKIs.

Despite the advancement made with the new generation of ALK TKI, a high unmet need is still present among ALK+ mNSCLC patients, especially for patients with BM in the real-world setting. A US-based study using Redacted data (2011-2018) assessed patients with ALK+ advanced NSCLC who initiated a ALK TKI (crizotinib, ceritinib, alectinib, or brigatinib).The study found median (95% CI) real-world PFS (rwPFS) was significantly shorter among patients with prior BM (4.97 [3.75–5.99] months) months compared with those without prior BM (8.52 [7.57–10.59] months).[8] A recent study using Medicare data (2017-2022) evaluated patients with ALK+ NSCLC receiving 2G ALK TKIs (alectinib or brigatinib) as first line treatment. The study found incident BM, developed post ALK+ TKI initiation, were associated with an increased risk of death (HR, 2.59, [95% CI, 1.98-3.38]).[9]

The burden of brain metastases in patients with ALK-positive metastatic NSCLC can manifest either at the time of NSCLC diagnosis or as new metastases that develop after the diagnosis. However, data on the epidemiology and impact of brain metastases in this population are still limited. A retrospective review of 115 ALK+ patients with stage IV NSCLC at the University of Colorado reported a baseline prevalence of BM of 34%.[10] Another real-world study reported a high 5-year cumulative incidence of BM of 61%, following a baseline prevalence of 24% among patients with ALK+ mNSCLC.[11] The Medicare study reported a baseline BM prevalence of 28% and a 20% cumulative incidence at 5-year follow-up[9] Additional research with recent data and rich clinical information is needed to explore how the presence of BM might impact survival in this population, especially in the context of new therapies that more effectively cross the blood-brain barrier. A post-hoc analysis of the CROWN study has shown that 1L lorlatinib has delayed time to brain progression versus crizotinib.[12] However, real-world evidence for lorlatinib still needs time to mature.

This noninterventional study is designated as a PASS and is conducted voluntarily by Pfizer.

8. RESEARCH QUESTION AND OBJECTIVES

The primary research question of this study is to assess the prevalence and incidence of BM among patients with ALK+ mNSCLC who received 1G or 2G ALK TKI as 1L treatment, and the impact of BM on clinical outcomes OS, rwPFS and time to treatment discontinuation [TTD]).

Primary Objectives

1. To describe the demographics and clinical characteristics of patients with ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs, both overall and stratified by baseline BM status
2. To assess the real-world incidence and prevalence of BM among patients with ALK+ mNSCLC receiving 1L treatment with either 1G or 2G ALK TKIs
3. To assess the impact of baseline BM and incident BM on OS (mortality) among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs
4. To assess the impact of baseline BM and incident BM on rwPFS among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs
5. To assess the impact of baseline BM and incident BM on TTD among ALK+ mNSCLC patients receiving 1L treatment with either 1G or 2G ALK TKIs

Secondary Objectives

1. To stratify the above analyses by generation of ALK TKI (1G: crizotinib; 2G: ceritinib, alectinib, brigatinib)

9. RESEARCH METHODS

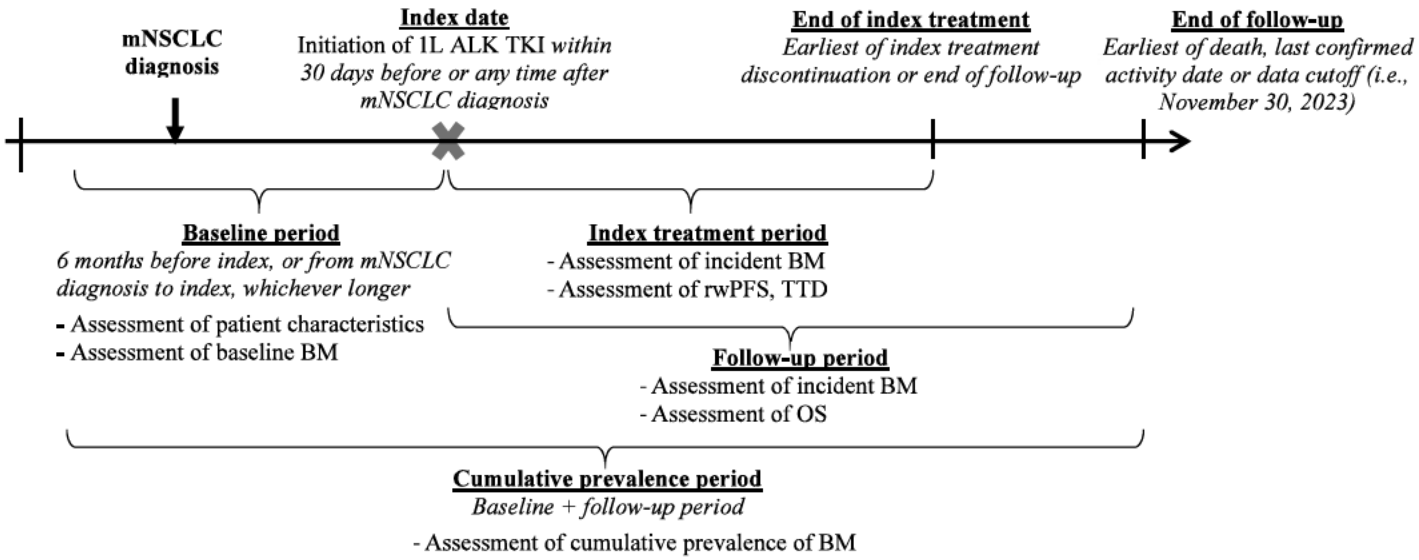
9.1. Study design

A retrospective observational cohort study will be employed for this study. Individual-level data collected retrospectively from oncology practice settings contributing to the Redacted EHR database will be used to characterize the incidence and prevalence of BM and to assess the association between BM and clinical outcomes including OS and rwPFS among patients with ALK+ mNSCLC patients who received 1G or 2G ALK TKIs in the 1L setting. The primary endpoints are demographic and clinical characteristics, incidence and prevalence of BM, OS and TTD. This is a descriptive (non comparative) study designed to generate real-world insights on the incidence and prevalence of brain metastases in the contemporary treatment era of patients with ALK+ mNSCLC. This study will generate real-world evidence to complement data from pivotal clinical trials.

A cohort of adult patients with ALK+ mNSCLC patients who received 1L treatment with 1G or 2G ALK TKIs will be identified from the Redacted database. The initiation date of either 1G or 2G ALK TKI as 1L treatment will be defined as the ***index date***. The ***baseline period*** will be defined as either 6 months prior to the index date or the period between mNSCLC diagnosis and the index date, whichever is longer. During this period, patient characteristics and baseline BM will be assessed. The ***follow-up period*** will be defined as the time from the index date to earliest of death, the last confirmed activity date or the data cutoff date (i.e., 30/11/2023). The ***index treatment period***

will be defined as the time from the index date to earliest of index treatment discontinuation (defined in **Section 0**) or end of follow-up. Incident BM will be assessed during both the index treatment period and the follow-up period. The cumulative prevalence of BM will be assessed across the combined baseline and follow-up periods. OS will be evaluated during the follow-up period, while rwPFS and TTD will be evaluated during the index treatment period. The study design schema is illustrated in **Figure 1.**

Figure 1. Study design schema



9.2. Setting

The study population will include ALK+ mNSCLC patients who received 1L treatment with either 1G or 2G ALK TKIs in the Redacted database. Patients were eligible to be included in the Redacted data if they were diagnosed with metastatic NSCLC between 01/01/16 and 30/11/22, or diagnosed with early-stage NSCLC and subsequently developed metastatic disease between 01/01/16 and 30/11/22. Patients will be followed until November 30, 2023. Inclusion and exclusion criteria are specified below.

9.2.1. Inclusion criteria

Patients included in the Redacted Redacted (as described in **Section 9.4**) must meet all of the following inclusion criteria to be eligible for inclusion in the study:

1. Patients with confirmed ALK+ mNSCLC
    a. For patients who initially diagnosed with metastatic NSCLC, their initial metastatic date will be defined as the initial diagnosis date or the earliest recorded date of any metastatic site, whichever is earlier.

b. For patients who initially diagnosed with early-stage or locally advanced NSCLC, their initial metastatic date will be defined as the earliest recorded date of any metastatic site

2. Initiation of 1G or 2G ALK TKI (crizotinib, ceritinib, alectinib, brigatinib) as 1L treatment within 30 days before or any time after the initial metastatic diagnosis date
   a. The initiation date of ALK TKI as 1L treatment will be defined as the ***index date***
3. Aged ≥18 years at the index date

9.2.2. Exclusion criteria

Patients meeting the following criteria will not be included in the study:

1. Evidence of any other primary cancer (except non-melanoma skin cancer; **Appendix 1**) prior to the index date
2. Evidence of prior ALK TKI treatment use before the index date
3. 1L ALK TKI is initiated in a clinical trial or in combination with other treatments

9.3. Variables

9.3.1. Baseline Characteristics

Demographic and clinical characteristics will be summarized on the index date or during the baseline period, as applicable. If multiple assessments for a variable are available within the baseline period, the last completed assessment closest to or on the index date will be used. Demographic and clinical characteristics of interest are summarized below in **Table 1,** and are contingent on clinical input and their availability in the [redacted] database.

Table 1. Baseline demographics and clinical characteristics

Variable	Role	Data source	Operational definition
Demographics			
Age at index, year	Covariates	[redacted]	Variable will be derived based on date of birth and index date
Gender			Male, Female
Race/ethnicity			Asian, Black or African American, White, Other, Unknown/Missing
Smoking history			History of smoking, No history of smoking, Unknown/Not documented
Geographic region			Northeast, Midwest, South, West, Unknown/Missing
Site of care			Academic, Community, Unknown/Missing
Disease characteristics			
Histology			Non-squamous cell carcinoma, Squamous cell carcinoma, Unknown

Tumor stage at initial NSCLC diagnosis	Covariates	[redacted]	Stage I, Stage II, Stage IIIA, Stage IIIB, Stage IIIC, Stage IV
ECOG PS at index			– 0: Fully active – 1: Restricted in physically strenuous activity – 2: Ambulatory and capable of all self-care – 3: Capable of only limited self-care – 4: Completely disabled – Unknown ECOG PS will be assessed during the window from the start of baseline period to 30 days after the index date. The value closest to the index date will be used • If a patient has multiple ECOG values recorded on the same day, the highest result value will be used • If a patient has no ECOG values recorded in the above time window, that patient will be categorized as "Not available"
Time from mNSCLC diagnosis to the index date, month			Duration from the initial metastatic date to the index date. Patients whose initial metastatic date occurs before the index date will be excluded when summarizing this variable
Site of metastasis			Binary variables (yes/no) indicating whether patients had metastasis prior to the index date in the following sites: - Brain - Liver - Bone - Pleural - Lymph node - Adrenal - Skin - Peritoneum In addition, the total number of distinct metastatic sites prior to the index date will be summarized.
Time from mNSCLC diagnosis to BM, month			Duration from the initial mNSCLC diagnosis date to the date of first BM, among patients developed BM during the baseline period
Comorbidities			

Comorbidities	Covariates	[redacted]	Binary variables (yes/no) indicating whether patients had a diagnosis of the following comorbidities during baseline period prior to the index date (**Appendix 2**). - Chronic obstructive pulmonary disease (COPD) - Chronic pulmonary disease - Diabetes - Peripheral vascular disease (PVD) - Congestive heart failure (CHF) - Hypertension - Myocardial infarction - Atrial fibrillation - Renal disease - Liver disease

9.3.2. Prevalence and Incidence of BM

The following variables will be summarized to evaluate prevalence and incidence of BM.

Table 2. Other variables

Variable	Role	Data source	Operational Definition
Baseline BM	Covariate	[redacted]	Baseline BM will be defined as BM occurring during the baseline period or at index date and will be evaluated among the total sample.
Incident BM	Outcome		Incident BM will be defined as newly developed BM after the index date. It will be assessed during both the index treatment period and the follow-up period. It will be evaluated only among patients who did not have baseline BM.
Cumulative prevalence of BM	Outcome		Cumulative prevalence of BM will be defined as the proportion of patients with BM occurring at any point during the baseline and follow-up periods among the total sample.

9.3.3. Clinical Outcomes

The following outcomes defined in **Table 3** will be summarized during the follow-up period. Real-world clinical outcome definitions will be based on Redacted derived line of therapy (LOT) algorithm.

Table 3. Outcomes during the follow-up period

Variable	Role	Data source	Operational Definition
OS	Outcome	[redacted]	OS will be defined as the time from index date to the date of death due to any cause. Patients without an event will be censored at last confirmed activity date.
rwPFS			rwPFS will be defined as the time from index date to the earliest evidence of disease progression or death due to any cause. Patients without an event will be censored at last clinic note date, or the day before next LOT start date (if initiated next LOT), whichever happened first. Disease progression or death that occurred after the initiation of the next LOT was not considered as an event; for those scenarios, patients were censored at the earlier of their last clinic note date or the day before next LOT start date. ***Additional notes on rwPFS calculation in*** [redacted]: o Progression events that occur within the first 14 days of the start date of the given LOT will be ignored, as they likely will not reflect outcomes related to the current drug regimen. o Patients will not be censored if this occurs and will still have opportunity for progression events beyond 14 days. o Deaths that occur within 14 days of the given LOT start date will still be counted as an event.
TTD			TTD will be defined as time from the index date to the discontinuation date of the index LOT (i.e., latest of all drug episodes across all drugs within the index LOT). Discontinuation will be defined as having a subsequent LOT, having a gap of more than 120 days with no systemic therapy following the last drug episode of 1L, or death. Patients without a discontinuation will be censored at the date of last drug episode activity.

9.4. Data sources

The Redacted EHR database is a longitudinal, de-identified, patient-level, real-world database derived from EHRs collected in US cancer clinics. The database covers more than 2.6 million active cancer patients treated at over 800 sites of care across 4 US census regions. Patient-level data includes both structured (diagnosis, demographics, laboratory values, biomarker information, drug orders, visits, etc.) and unstructured (physician notes, radiology, pathology reports, etc.) sources. Data on death is drawn from structured or unstructured data fields in the EHR, and publicly available sources of mortality including the US Social Security Death Index, and commercial obituary data; this outcome has been validated against the National Death Index as the gold standard.[13] The methods used to curate rwPFS have also been previously evaluated and validated with established frameworks.[14,15]

Pfizer has contracted with Redacted for a custom data extract containing information on patients with ALK+ advanced NSCLC referred to as the Redacted Redacted. The Redacted team completed the data abstraction and curation of the dataset, delivering deidentified structured data for analysis. The data cutoff date is 30 November 2023. To be included within this dataset, patients were required to meet the following criteria:

Redacted ***cohort inclusion criteria:***

- Included in Redacted, which implemented the following criteria:
  - Diagnosis of lung cancer using International Classification of Diseases (ICD) codes (ICD-9 162.x or ICD-10 C34x, or C39.9) from structured data
  - Pathology consistent with NSCLC
  - Have an initial diagnosis with Stage IIIB, IIIC, IVA or IVB NSCLC on or after 1/1/2011, OR diagnosed with early-stage NSCLC and subsequently develops recurrent or progressive disease on or after 1/1/2011
- Additional inclusion criteria for the ALK+ cohort:
  - Have a diagnosis with Stage IV, IVA or IVB NSCLC on or after 01/01/16 up until 30/11/22, OR have a diagnosis with early-stage NSCLC and subsequently developed metastatic disease on or after 01/01/16 up until 30/11/22
  - At least two documented clinical visits, on different days, occurring on or after 01/01/16 up until 30/11/22
  - Has evidence of ALK fusion/rearrangement
  - Has evidence of 1L therapy

Redacted ***cohort exclusion criteria:***

- Lacking relevant prespecified unstructured documents in the Redacted Database for review by the abstraction team

For the present study, a sample of eligible patients receiving 1L treatment with either 1G or 2G ALK TKIs, will be drawn from the Redacted *cohort* by applying the inclusion and exclusion criteria described in **Section 9.2**.

9.5. Study size

As no priori hypotheses are specified, sample size calculations are not applicable. The number of patients eligible for the study will be determined in accordance with the sample selection conducted per the criteria described in **Section 9.2.** Stratified analyses for the secondary objective will occur pending sufficient sample size.

9.6. Data management

Clean, patient-level datasets derived from Redacted data will be generated for use throughout the study. This process will entail basic exploratory checks to ensure data integrity, cleaning and reformatting the raw data as needed, and creating variables for all key study measures. All data will be stored and maintained on a secure encrypted non-cloud-based server and accessed over a secure internal private wide area network. The data will be made accessible only to those working on the current study. No attempt will be made to identify individual patients, hospitals, or physicians. Analyses will be conducted using SAS version 9.4 and/or R version 4.1.0 or later.

9.7. Data analysis

The extent of missing data will be summarized for all variables described in **Section 9.3**. Baseline characteristics, clinical outcomes, and other variables of interest will be evaluated in the subset of patients with complete information on these variables. Given the descriptive nature of this study, no imputation will be conducted for missing data. As only the month and year of death are available in the Redacted database, the 15th of the month will be assumed for patients who die during the observation period. The approach for handling observations with imputed death dates occurring earlier than the index date and within the same month, such as setting the death date to the index date or excluding these observations, will be finalized based on the frequency of this phenomenon in the data and clinical input. Table shells for primary objectives 1-5 and secondary objective 1 are listed in **Appendix 1**.

9.7.1. Data Analysis for Primary Objective 1: Describe baseline characteristics in the ALK+ mNSCLC population, both overall and stratified by baseline BM status

Descriptive analysis will be conducted for baseline demographic and clinical characteristics, as defined in **Table 1**, in the overall sample, and stratified and compared by baseline BM status. Continuous variables will be described using means (with standard deviation [SD]) and medians (with interquartile range [IQR]) and compared using Wilcoxon rank sum tests. Categorical variables will be described using frequencies (with proportions) and compared using Chi-square tests (or Fisher's exact tests).

9.7.2. Data Analysis for Primary Objective 2: Describe real-world prevalence and incidence of BM in the overall ALK+ mNSCLC population

To assess prevalence and incidence of BM, four metrics will be calculated:

- Prevalence of BM at baseline

  - Defined as the proportion of patients with BM occurring at any point during the baseline period or at the index date, calculated among all selected ALK+ mNSCLC patients

- Incidence of BM during the index treatment period
  - Incidence rate of BM (per 100,000 person-years) will be estimated over the index treatment period, defined by earlier of index treatment discontinuation or end of follow-up, among patients without baseline BM
  - CIF will be used to estimate the cumulative incidence of BM among patients without baseline BM. Patients will be censored on discontinuation of index treatment, or end of follow-up. Death will be treated as a competing event.
- Incidence of BM during the follow-up period
  - Incidence rate of BM (per 100,000 person-years) will be estimated over the follow-up period, among patients without baseline BM
  - CIF will be used to estimate the cumulative incidence of BM among patients without baseline BM. Patients will be censored on the end of follow-up. Death will be treated as a competing event.
- Cumulative prevalence of BM until the end of follow-up
  - Defined as composite of prevalent and incident BM: the proportion of patients with BM occurring at any point during the baseline and follow-up period, calculated among the total sample
  - CIF will be used to estimate the cumulative prevalence of BM among all patients. Those with baseline BM will be considered to have an event at day 0 (i.e., the index date). Patients will be censored on end of follow-up. Death will be treated as a competing event.

9.7.3. Data Analysis for Primary Objective 3: Assess the impact of baseline BM and incident BM on OS (mortality) in the overall ALK+ mNSCLC population

To assess the impact of baseline BM on OS, Kaplan-Meier analysis will be conducted to describe OS (defined in **Table 3)** during the follow-up period, overall and by baseline BM status. The number of patients at risk and the number and proportion of patients experiencing death at different time points will be reported. The median time to event and corresponding 95% CIs will be reported. Adjusted comparisons of OS between patients with and without baseline BM will be conducted using Cox proportional hazards regression adjusted for key baseline patient characteristics as described in **Section 9.3.1**. HRs and 95% CIs from the Cox regression will be reported.

To assess the impact of incident BM on OS, a Cox regression model with time-varying covariates will be employed, which will incorporate:

- Event indicator: a binary variable representing whether the outcome (i.e., death) has occurred by a specific time point

- Time-fixed covariates: selected key baseline patient characteristics as described in **Section 9.3.1.**
- Time-varying covariates: BM status and ECOG PS, which will be updated dynamically during the follow-up period

The time-varying covariates (e.g., BM status) will be assessed during the same assessment window as the outcomes. HRs with corresponding 95% CIs will be reported.

For descriptive purposes, Kaplan-Meier curves will be constructed to estimate OS for three groups:

- Patients with baseline BM
- Patients without baseline BM but developed incident BM
- Patients without baseline BM and never developed incident BM

The Cox model setup, covariate selection, and parameterization of time-varying covariates will be finalized during analysis based on model convergence and clinical relevance.

9.7.4. Data Analysis for Primary Objective 4: Assess the impact of baseline BM and incident BM on rwPFS in the overall ALK+ mNSCLC population

The same statistical analyses as described in **Section 9.7.3** will be conducted, with the outcome being rwPFS, as defined in **Table 3**.

9.7.5. Data Analysis for Primary Objective 5: Assess the impact of baseline BM and incident BM on TTD in the overall ALK+ mNSCLC population

The same statistical analyses as described in **Section 9.7.3** will be conducted, with the outcome being TTD, as defined in **Table 3**.

9.7.6. Data Analysis for Secondary Objective 1: Stratify analyses by generation of ALK TKI

All the analyses above will be repeated among ALK+ mNSCLC patients receiving 1L treatment with 1G ALK TKI (i.e., crizotinib) as one group, and 2G ALK TKIs (i.e., ceritinib, alectinib, brigatinib) as another group, pending sufficient sample size.

9.8. Quality control

Best practice guidelines will be followed to ensure project quality, including structured organization of project materials (e.g., data extracts, statistical software programs, output tables) and standard internal audit process. The audit process both confirms the validity of the analytical approach and ensures that all programs and results are accurate.

9.9. Limitations of the research methods

Several limitations should be considered when interpreting results from this real-world study:

- Key conditions that result in inclusion or exclusion of patients from the study sample will be identified by diagnosis codes or lab assessments in the Redacted data, which may be incomplete. Therefore, implementation of certain inclusion criteria dependent on lab assessment may not be feasible, and misclassification due to miscoding of these conditions

may result in the inclusion of patients with exclusionary conditions (e.g., patients diagnosed with other primary malignancies).

- It is expected that measurement error and misclassification due to abstraction errors will be low as the Redacted uses trained abstractors who abstract data according to systematic, modular abstraction guidelines. Additionally, defined Quality assurance/quality control activities aim to reduce biases and data issues. However, given that the underlying data being abstracted are collected in real-world oncology practice, it is possible that errors will have been made at the time of data entry into the patient chart or EHR. The Redacted team is responsible for the curation of the data, Pfizer will receive de-identified structured data for analysis. The Redacted team is responsible for the curation of the data, Pfizer will receive de-identified structured data for ana25lysis.
- Capture of progression is done retrospectively via abstraction of patient charts in the Redacted Redacted. It is possible that abstractors may miss evidence of response and/or progression events during the abstraction process, which may result in an overestimation of rwPFS. In addition, progression events are captured using trained abstractors that review patient charts in a retrospective fashion, which may be subject to bias from variability in tumor assessment interval, tumor assessment modalities, and interpretation of tumor response. Since these data reflect real-world treatment and care practices, the frequency of scans, visits, and assessments will likely not be as regular and systematic as a clinical trial. However, the methods used to curate rwPFS have been previously evaluated and validated with established frameworks.[14,15] The mortality endpoint, however, has been previously validated against a gold standard.[13]
- Results from this study are largely from US community oncology centers, which may limit the generalizability of study findings to patients treated in academic medical centers or outside the US.

9.10. Other aspects

Not applicable.

10. PROTECTION OF HUMAN SUBJECTS

10.1. Patient information

This study involves data that exist in anonymized structured format and contain no patient personal information. Implementation of the protocol does not involve converting unstructured data to structured data, no patient personal data will be accessed throughout the conduct of this study.

10.2. Patient consent

As this study involves anonymized structured data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.

10.3. Institutional review board (IRB)/Independent ethics committee (IEC)

This retrospective database analysis does not involve the collection, use, or transmittal of individually identifiable data. As such, the study falls within the definition of exempt research under

45 CFR 46.104(d)(4)(ii) and IRB approval is not required. Because the dataset does not include individually identifiable health information under 45 CFR 164.514, Health Insurance Portability and Accountability Act (HIPAA) requirements do not apply.

10.4. Ethical conduct of the study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value and rigor and follow generally accepted research practices described in Guidelines for Good Pharmacoepidemiology Practices issued by the International Society for Pharmacoepidemiology.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study involves claims and EHR data that exist as structured data by the time of study start or a combination of existing structured data and unstructured data, which will be converted to structured form during the implementation of the protocol solely by a computer using automated/algorithmic methods, such as natural language processing.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

Based on the analysis results and discussions with Pfizer, a study report summarizing the background, objectives, methods, results, and conclusion of the study will be prepared. In addition, the primary results of this research study will be externally disseminated in a manuscript submitted to a peer-reviewed, scientific journal or in an abstract/presentation at a scientific conference or symposium. Any publication related to the study will be reviewed/approved by Pfizer prior to submitting results externally. Any publication resulting from this work will adhere to the procedures and pre-specified analysis plans within this protocol.
In the event of any prohibition or restriction imposed (e.g., clinical hold) by an applicable competent authority in any area of the world, or if the investigator, Redacted, is aware of any new information which might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

13. REFERENCES

1. Funder JW, Carey RM, Mantero F, et al. The management of primary aldosteronism: case detection, diagnosis, and treatment: an endocrine society clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism.* 2016;101(5):1889-1916.
2. Soda M, Choi Yl Fau - Enomoto M, Enomoto M Fau - Takada S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. (1476-4687 (Electronic)).
3. Shaw AT, Kim Dw Fau - Nakagawa K, Nakagawa K Fau - Seto T, et al. Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. (1533-4406 (Electronic)).
4. Mok T, Camidge DR, Gadgeel SM, et al. Updated overall survival and final progression-free survival data for patients with treatment-naive advanced ALK-positive non-small-cell lung cancer in the ALEX study. (1569-8041 (Electronic)).
5. Camidge DR, Kim HR, Ahn MJ, et al. Brigatinib Versus Crizotinib in Advanced ALK Inhibitor-Naive ALK-Positive Non-Small Cell Lung Cancer: Second Interim Analysis of the Phase III ALTA-1L Trial. (1527-7755 (Electronic)).
6. Shaw AT, Bauer TM, de Marinis F, et al. First-Line Lorlatinib or Crizotinib in Advanced ALK-Positive Lung Cancer. (1533-4406 (Electronic)).
7. Solomon BJ, Bauer TM, Mok TSK, et al. Efficacy and safety of first-line lorlatinib versus crizotinib in patients with advanced, ALK-positive non-small-cell lung cancer: updated analysis of data from the phase 3, randomised, open-label CROWN study. (2213-2619 (Electronic)).
8. Jahanzeb M, Lin HM, Pan X, et al. Real-World Treatment Patterns and Progression-Free Survival Associated with Anaplastic Lymphoma Kinase (ALK) Tyrosine Kinase Inhibitor Therapies for ALK+ Non-Small Cell Lung Cancer. (1549-490X (Electronic)).
9. Bazhenova L, Abrahami D, Ramaswamy K, et al. EP. 17A. 01 The Cumulative Incidence of Brain Metastases in US Medicare Patients with ALK+ mNSCLC Treated with Second-generation ALK TKIs. *Journal of Thoracic Oncology.* 2024;19(10):S711-S712.
10. Patil T, Smith DE, Bunn PA, et al. The Incidence of Brain Metastases in Stage IV ROS1-Rearranged Non-Small Cell Lung Cancer and Rate of Central Nervous System Progression on Crizotinib. (1556-1380 (Electronic)).
11. Schmid S, Cheng S, Chotai S, et al. Real-World Treatment Sequencing, Toxicities, Health Utilities, and Survival Outcomes in Patients with Advanced ALK-Rearranged Non-Small-Cell Lung Cancer. (1938-0690 (Electronic)).
12. Solomon BA-O, Bauer TA-O, Ignatius Ou SA-O, et al. Post Hoc Analysis of Lorlatinib Intracranial Efficacy and Safety in Patients With ALK-Positive Advanced Non-Small-Cell Lung Cancer From the Phase III CROWN Study. (1527-7755 (Electronic)).
13. Curtis MD, Griffith SD, Tucker M, et al. Development and validation of a high‐quality composite real‐world mortality endpoint. *Health services research.* 2018;53(6):4460-4476.
14. Griffith SD, Miksad RA, Calkins G, et al. Characterizing the feasibility and performance of real-world tumor progression end points and their association with overall survival in a large advanced non–small-cell lung cancer data set. *JCO clinical cancer informatics.* 2019;3:1-13.

15. Ton TA-O, Pal N, Trinh H, et al. Replication of Overall Survival, Progression-Free Survival, and Overall Response in Chemotherapy Arms of Non-Small Cell Lung Cancer Trials Using Real-World Data. (1557-3265 (Electronic)).

14. LIST OF TABLES

15. LIST OF FIGURES

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

Appendix 1. Table Shells for Primary Objectives 1-5 and Secondary Objective 1

Redacted

ANNEX 2. ENCEPP CHECKLIST FOR STUDY PROTOCOLS

None.

ANNEX 3. ADDITIONAL INFORMATION

Appendix 2. ICD Diagnosis Codes Used to Identify Other Primary Cancer

Description	Code type	Code
Malignant neoplasm of lip, oral cavity, and pharynx	ICD-9	140-149
Malignant neoplasm of digestive organs and peritoneum	ICD-9	150-159
Malignant neoplasm of respiratory and intrathoracic organs	ICD-9	160-165
Malignant neoplasm of bone and articular cartilage	ICD-9	170
Malignant neoplasm of connective and other soft tissue	ICD-9	171
Malignant melanoma of skin	ICD-9	172
Malignant neoplasm of female breast	ICD-9	174
Malignant neoplasm of male breast	ICD-9	175
Kaposi's sarcoma	ICD-9	176
Malignant neoplasm of genitourinary organs	ICD-9	179-189
Malignant neoplasm of eye	ICD-9	190
Malignant neoplasm of brain	ICD-9	191
Malignant neoplasm of other and unspecified parts of nervous system	ICD-9	192
Malignant neoplasm of thyroid gland	ICD-9	193
Malignant neoplasm of other endocrine glands and related structures	ICD-9	194
Malignant neoplasm of other and ill-defined sites	ICD-9	195
Other malignant neoplasm without specification of site	ICD-9	199.1
Malignant neoplasm associated with transplant organ	ICD-9	199.2
Malignant neoplasm of lymphatic and hematopoietic tissue (excluding Neuroendocrine tumors)	ICD-9	200-208
Malignant carcinoid tumors of the small intestine	ICD-9	209.0
Malignant carcinoid tumors of the appendix, large intestine, and rectum	ICD-9	209.1
Malignant carcinoid tumors of other and unspecified sites	ICD-9	209.2
Malignant poorly differentiated neuroendocrine tumors	ICD-9	209.3
Malignant neoplasms of lip, oral cavity and pharynx	ICD-10	C00-C14
Malignant neoplasms of digestive organs	ICD-10	C15-C26

Malignant neoplasms of respiratory and intrathoracic organs	ICD-10	C30-C39
Malignant neoplasms of bone and articular cartilage	ICD-10	C40-C41
Malignant melanoma of skin	ICD-10	C43
Malignant neoplasms of mesothelial and soft tissue	ICD-10	C45-C49
Malignant neoplasm of breast	ICD-10	C50
Malignant neoplasms of female genital organs	ICD-10	C51-C58
Malignant neoplasms of male genital organs	ICD-10	C60-C63
Malignant neoplasms of urinary tract	ICD-10	C64-C68
Malignant neoplasms of eye, brain and other parts of central nervous system	ICD-10	C69-C72
Malignant neoplasms of thyroid and other endocrine glands	ICD-10	C73-C75
Malignant neoplasm of other and ill-defined sites	ICD-10	C76
Malignant (primary) neoplasm, unspecified	ICD-10	C80.1
Malignant neoplasm associated with transplanted organ	ICD-10	C80.2
Malignant neoplasms of lymphoid, hematopoietic and related tissue	ICD-10	C81-C96
Malignant neuroendocrine tumors	ICD-10	C7A

Appendix 3. ICD Diagnosis Codes Used to Identify Comorbidities

Condition	ICD-9-CM	ICD-10-CM
COPD	490-492, 496	J41-J44
Chronic pulmonary disease	416.8, 416.9, 490-505, 506.4, 508.1, 508.8	I27.8-I27.9, J40-J47, J60-J67, J68.4, J70.1, J70.3
Diabetes	250	E10, E11, E12, E13, E14
Peripheral vascular disease	093.0, 437.3, 440, 441, 443.1-443.9, 447.1, 557.1, 557.9, V43.4	I70-I71, I73.1, I73.8-I73.9, I77.1, I79.0, I79.2, K55.1, K55.8-K55.9, Z95.8-Z95.9
Congestive heart failure	428	I09.9, I11.0, I13.0, I13.2, I25.5, I42.0, I42.5-I42.9, I43, I50, P29.0
Hypertension	362.11, 401-405, 437.2	I10, I11, I13, I15, I16
Myocardial infarction	410, 412	I21, I22, I25.2
Atrial fibrillation	427.31	I48.0, I48.1, I48.2, I48.91
Renal disease	403.01, 403.11, 403.91, 404.02, 404.03, 404.12, 404.13, 404.92, 404.93, 582, 583.0-583.7, 585, 586, 588.0, V42.0, V45.1, V56	I12.0, I13.1, N03.2-N03.7, N05.2-N05.7, N18-N19, N25.0, Z49.0-Z49.2, Z94.0, Z99.2
Liver disease	070.22, 070.23, 070.32, 070.33, 070.44, 070.54, 070.6, 070.9, 570, 571, 573.3, 573.4, 573.8, 573.9, V42.7, 456.0-456.2, 572.2-572.8	B18, K70.0-K70.3, K70.9, K71.3-K71.5, K71.7, K73-K74, K760, K762-K764, K768-K769, Z94.4, I85.0, I85.9, I86.4, I98.2, K70.4, K71.1, K72.1, K72.9, K76.5-K76.7

Document Approval Record

Document Name:	A8081077_Non Interventional Study Protocol_FINAL-27Nov2024
Document Title:	A8081077_Non Interventional Study Protocol_FINAL-27Nov2024

Signed By:	Date(GMT)	Signing Capacity
Redacted	06-Dec-2024 22:06:14	Redacted
Redacted	07-Dec-2024 13:21:41	Redacted