1. Title Page

Title	CARE Initiative Study: Real-world emulation of the KEYNOTE-189 comparative effectiveness trial of pembrolizumab, pemetrexed, and chemotherapy vs. placebo, pemetrexed, and chemotherapy for the first-line treatment of metastatic non-small cell lung cancer
Research Question and Objectives	This study seeks to advance understanding of under what conditions real-world evidence studies can provide reliable conclusions about drug effectiveness. The objective is to emulate the KEYNOTE-189 randomized controlled trial of pembrolizumab, pemetrexed, and chemotherapy as first-line treatment in adult patients with metastatic non-small cell lung cancer without EGFR or ALK mutations using real-world electronic health record data.
Protocol version	v1.1
Last update date	August 13, 2024
Contributors	Natalie Levy, PhD Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Paige Sheridan, PhD Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Nileesa Gautam, MSc Aetion, Inc 50 Congress St. Suite 1025 Boston, MA 02109 Adina Estrin, MS

	Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Monica Iyer Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Sarah McDonald Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Ann Madsen, PhD Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001 Ulka Campbell, PhD Aetion, Inc 5 Penn Plaza 7th Floor New York, NY 10001
Sponsor	This work was collectively sponsored by Abbvie, Amgen, AstraZeneca, Bayer, Gilead, Janssen, and Pfizer.
Conflict of interest	NL, PS, NG, AE, MI, SM, AM, and UC are shareholders and employees of Aetion, Inc.

Table of contents

2. Abstract

Background

The Coalition to Advance Real-World Evidence through Randomized Controlled Trial Emulation (CARE) Initiative is a program designed to build an empirical evidence base for the use of real-world data (RWD) in clinical and regulatory decision-making.[1] Using randomized controlled trials (RCT) as a benchmark for causal effect estimates, a series of RCT emulations will be conducted across varying trials, real world data sources, and study design elements to better understand under what conditions non-interventional studies, using data generated during routine clinical care, can provide reliable conclusions about drug effectiveness.

Research Question and Objectives

In this study, real-world electronic health record (EHR) data will be used to emulate the KEYNOTE-189 efficacy trial of pembrolizumab as first-line therapy in patients with metastatic non-small cell lung cancer (NSCLC) without epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) sensitizing mutations.[2] Similarly to the KEYNOTE-189 trial, this study will compare real-world overall survival (rwOS) and real-world progression-free survival (rwPFS) between patients who initiate pemetrexed, platinum-based chemotherapy, and pembrolizumab, and patients who initiate pemetrexed and platinum-based chemotherapy alone.

Research Methods

The inclusion and exclusion criteria applied in the KEYNOTE-189 trial will be operationalized in a real-world EHR data source, as closely as is feasible, to create an observational cohort similar to the trial study population. Patients who initiate treatment with pemetrexed, platinum-based chemotherapy (cisplatin or carboplatin), and pembrolizumab within 15 days of one another in the metastatic setting will be considered exposed. Patients who initiate pemetrexed and platinum-based chemotherapy only (with no evidence of pembrolizumab) within 15 days of one another will comprise the comparator group. Inverse probability of treatment weighting (IPTW) will be used to control for measured hypothesized confounders. Kaplan-Meier methods and Cox proportional hazards models will be used to compare median rwOS and rwPFS and hazards of progression and death between the exposure and comparator groups.

3. Amendments and updates

Version date	Version number	Section of protocol	Amendment or update	Reason
August 13, 2024	1.1	Tables 4, 5, 9	Amendments to some variable assessment windows	To improve emulation of trial criteria and measures

4. Milestones

Table 1. Milestones

Milestone	Date
Initial feasibility assessment	March 21, 2024
Additional data explorations	April 8, 2024
Draft 1 of protocol complete	May 30, 2024
Final protocol shared with steering committee	June 28, 2024
Amended protocol shared with steering committee	August 13, 2024

5. Rationale and background

The potential of non-interventional studies using real-world data (RWD) — healthcare data generated during routine clinical practice — to produce evidence about the effectiveness and safety of biomedical products is increasingly recognized by clinical and regulatory decision makers.[3,4] This is reflected by the growing use of RWD to support regulatory approvals.[5] Real-world evidence (RWE) studies complement randomized controlled trials (RCTs) by generating new hypotheses, producing results more quickly and at a lower cost, including broader patient populations, reflecting clinical care patterns, and assessing longer-term outcomes.[1,6] These advantages of RWE studies are of particular value in the field of oncology due to high unmet medical need, poor patient outcomes for several cancer types, a rapidly evolving treatment landscape, and the need to generate additional confirmatory evidence following accelerated regulatory approval.[7]

At the same time, causal inference from non-randomized studies leveraging RWD may be hindered by threats to internal validity. Due to a lack of randomization, RWE studies may suffer from unmeasured or inadequately controlled confounding. Key variables may be missing or misclassified in data generated from clinical practice, which may introduce information bias and limit direct comparisons with clinical trials. Therefore, successful application of RWD to support clinical and regulatory decision-making requires a thorough understanding of the circumstances under which RWD can generate valid evidence about treatment effectiveness.

The Coalition to Advance Real-World Evidence through Randomized Controlled Trial Emulation (CARE) Initiative aims to contribute to this understanding by building an empirical evidence base for the generation of RWD-based evidence of treatment effectiveness.[1] To do this, electronic health record (EHR) data collected during routine healthcare practice will be used to emulate the primary outcomes of completed RCTs for oncology therapies. The RCT results will provide a benchmark causal effect estimate against which the findings of non-randomized emulations can be compared. No standard metric has been proposed to quantify agreement between emulation and RCT results and previous work has used a variety of measures.[8] The CARE emulations will focus on *qualitative agreement* — whether findings from a non-interventional

study and RCT are in the same direction and are of similar magnitude. The choice to use a metric that is not anchored to statistical significance reflects conclusions from the CARE pilot study about specific challenges in oncology emulations (e.g., small real-world sample sizes) and the non-inferential goals of this work.[9,10] Through this effort, the CARE Initiative seeks to identify under what conditions non-interventional studies using data generated during routine clinical care can provide reliable conclusions about drug effectiveness.

6. Research question and objectives

In this emulation of the KEYNOTE-189 trial, real-world EHR data will be used to estimate the effectiveness of initiating first-line treatment with pemetrexed and carboplatin or cisplatin plus pembrolizumab versus pemetrexed and carboplatin or cisplatin alone among patients with metastatic non-small cell lung cancer (NSCLC) without epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) sensitizing mutations.[2]

Table 2. Primary research question and objective

Objective:	The objective of this non-interventional study is to estimate the effectiveness of initiating first-line pemetrexed and cisplatin or carboplatin plus pembrolizumab versus pemetrexed and cisplatin or carboplatin alone in a real-world emulation of the KEYNOTE-189 RCT.
Hypothesis:	Patients with metastatic NSCLC without EGFR or ALK mutations treated with pemetrexed, cisplatin/carboplatin, and pembrolizumab will have improved real-world overall survival (rwOS) and real-world progression-free survival (rwPFS) compared with patients treated with pemetrexed and cisplatin/carboplatin alone.
Population:	Adult patients (≥18 years of age) with metastatic NSCLC without EGFR or ALK mutations
Exposure:	Pemetrexed, cisplatin or carboplatin, and pembrolizumab as first-line treatment for metastatic NSCLC
Comparator:	Pemetrexed and cisplatin or carboplatin as first-line treatment for metastatic NSCLC
Outcomes:	rwOS, defined as time from study treatment initiation to death rwPFS, defined as time from study treatment initiation to disease progression or death
Setting:	Clinical data sourced from oncology practices in the United States (U.S.)
Main measure of effect:	Hazard ratio for rwOS and rwPFS in the intent-to-treat population

7. Research methods

7.1 Study design

7.1.1 Overview of key design elements of the KEYNOTE-189 trial

Study design: The KEYNOTE-189 trial (NCT02578680)[11] was an international, randomized, double-blind, placebo-controlled Phase III clinical trial comparing the efficacy and safety of first-line pemetrexed and cisplatin or carboplatin plus pembrolizumab versus pemetrexed and cisplatin or carboplatin plus placebo for the treatment of metastatic NSCLC without sensitizing EGFR or ALK mutations. Figure 1 displays the study design diagram for the KEYNOTE-189 trial.

Population: The trial study population included patients who were 18 years of age or older with pathologically confirmed metastatic nonsquamous NSCLC without sensitizing EGFR or ALK mutations. Eligible patients had received no prior systemic therapy for metastatic disease; had an Eastern Cooperative Oncology Group (ECOG) performance-status score of 0 or 1; had at least one measurable lesion according to the Response Evaluation Criteria in Solid Tumors (RECIST) v1.1[12]; and provided a tumor sample for determination of programmed death-ligand 1 (PD-L1) status. Patients with symptomatic central nervous system (CNS) metastases; a history of noninfectious pneumonitis that required glucocorticoids; active autoimmune disease requiring systemic treatment; or receipt of more than 30 Gy of radiotherapy to the lung in the previous six months were excluded. Full inclusion and exclusion criteria are listed in Table 4 and Table 5.

Endpoints: The two primary trial endpoints were overall survival (OS), defined as the time from randomization to death from any cause, and progression-free survival (PFS), defined as the time from the date of randomization to the date of the first documentation of objective progression of disease, as evaluated by study investigators according to RECIST v1.1, or death due to any cause. Follow-up for OS and PFS continued until the first of: documented disease progression (PFS only); death; last disease assessment in the absence of progression (PFS only); initiation of a new anti-cancer therapy (PFS only); discontinuation from overall study participation due to withdrawal of consent, unacceptable adverse events, disease progression, or intercurrent illness; or loss to follow-up. Crossover to pembrolizumab monotherapy was permitted among patients in the comparator group who experienced disease progression.

Analysis: Efficacy was assessed in the intent-to-treat population defined by randomized treatment assignment. Median OS and PFS and corresponding 95% confidence intervals were estimated after a maximum follow-up of 20.4 months using the Kaplan-Meier method and compared by treatment status using a log-rank test. The hazard ratio for death and progression was calculated using a Cox proportional-hazards model.

7.1.2 Overview of key design elements of the real-world emulation study

Study design: This new user, non-randomized active comparator cohort study will compare rwOS and rwPFS between patients with records indicating initiation of pemetrexed and carboplatin or cisplatin plus pembrolizumab versus pemetrexed and carboplatin or cisplatin alone following qualifying metastatic NSCLC diagnosis in the EHR data source. The data originate from oncology practices in the U.S. and include both

structured and curated data elements abstracted from unstructured sources including provider notes, pathology, and imaging reports (Section 7.6.1).

Population: The study population will include adults 18 years of age or older with a diagnosis of nonsquamous metastatic NSCLC without sensitizing EGFR or ALK mutations recorded in the EHR. Patients will be selected to reflect the KEYNOTE-189 trial eligibility criteria, as feasible in the RWD source, to create a trial-similar population. Eligible patients will include those with no record of previously receiving systemic anti-cancer therapy for metastatic disease; and without evidence of an ECOG performance status score of >1. Patients with evidence of CNS metastases; history of noninfectious pneumonitis with glucocorticoid treatment; autoimmune disease with evidence of immunosuppressive treatment; or radiotherapy to the lung in the previous six months will be excluded. Real-world operationalization of all trial inclusion and exclusion criteria are listed in Table 4 and Table 5. Study exposure groups will be ascertained within a 15-day time window ('exposure ascertainment window'), beginning on the day of the first record of a study drug. Patients initiating treatment for metastatic disease with pemetrexed and carboplatin or cisplatin plus pembrolizumab within the exposure ascertainment window will be classified as exposed. Patients initiating treatment with pemetrexed and carboplatin or cisplatin and with no evidence of pembrolizumab within the exposure ascertainment window will be classified as comparator patients.

Endpoints: The two study endpoints, rwOS and rwPFS, will be based on curated progression information and date of death available in the data source. For rwOS, patients will be followed until the first of: death; the administrative end of the study period (Day 621, or 20.4 months, to conform to the maximum follow-up time at which OS and PFS were evaluated in the KEYNOTE-189 trial); end of the data; or the last date of EHR activity for patients without evidence of death. For rwPFS, patients will be followed until the first of: progression or death; initiation of a new anti-cancer therapy; the administrative end of the study period; the end of data; or loss to follow-up (the last date prior to a period of >90 days without curated EHR activity and without death).

Analysis: Inverse probability of treatment weighting (IPTW) will be used to adjust for measured baseline confounders. Median rwOS and rwPFS and corresponding 95% confidence intervals will be estimated using the Kaplan-Meier method and compared between exposure groups using a log-rank test. The hazard ratios for death and progression will be calculated using Cox proportional-hazards models.

Rationale for study design choice: The choice of study design, population, endpoint, and analysis plan are intended to emulate the KEYNOTE-189 trial design as closely as possible, including creating a trial-similar real-world population, controlling for confounding in the absence of randomization, and estimating the intent-to-treat effect as was done in the trial.

7.2 Study design diagram

The figures below display the study designs for the KEYNOTE-189 RCT (Figure 1) and this real-world emulation study (Figure 2). For simplicity, inclusion and exclusion criteria displayed correspond to those highlighted in the trial publication.[2] Full trial eligibility criteria are listed in Sections 7.3.2 and 7.3.3.

Figure 1. KEYNOTE-189 randomized controlled trial study design

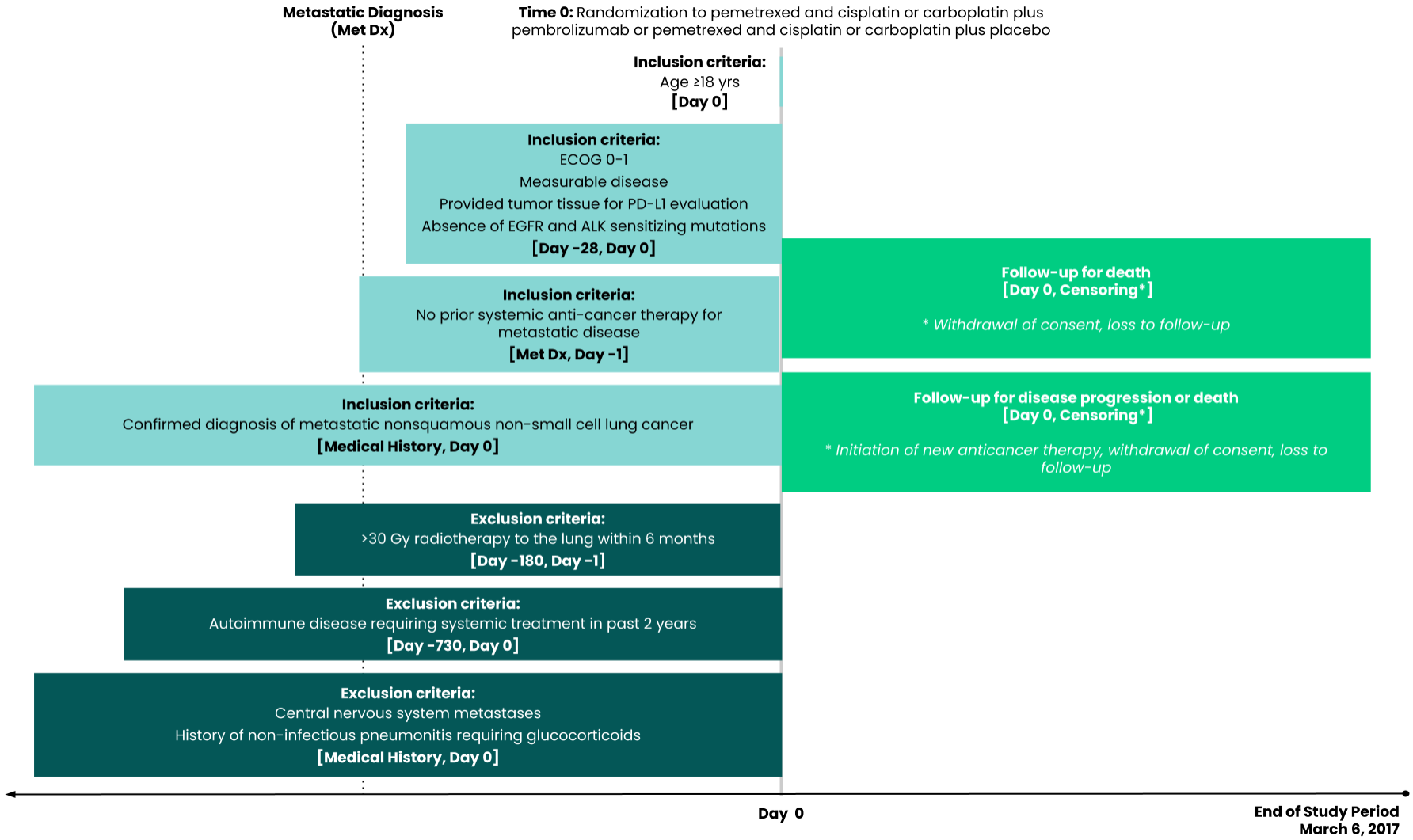
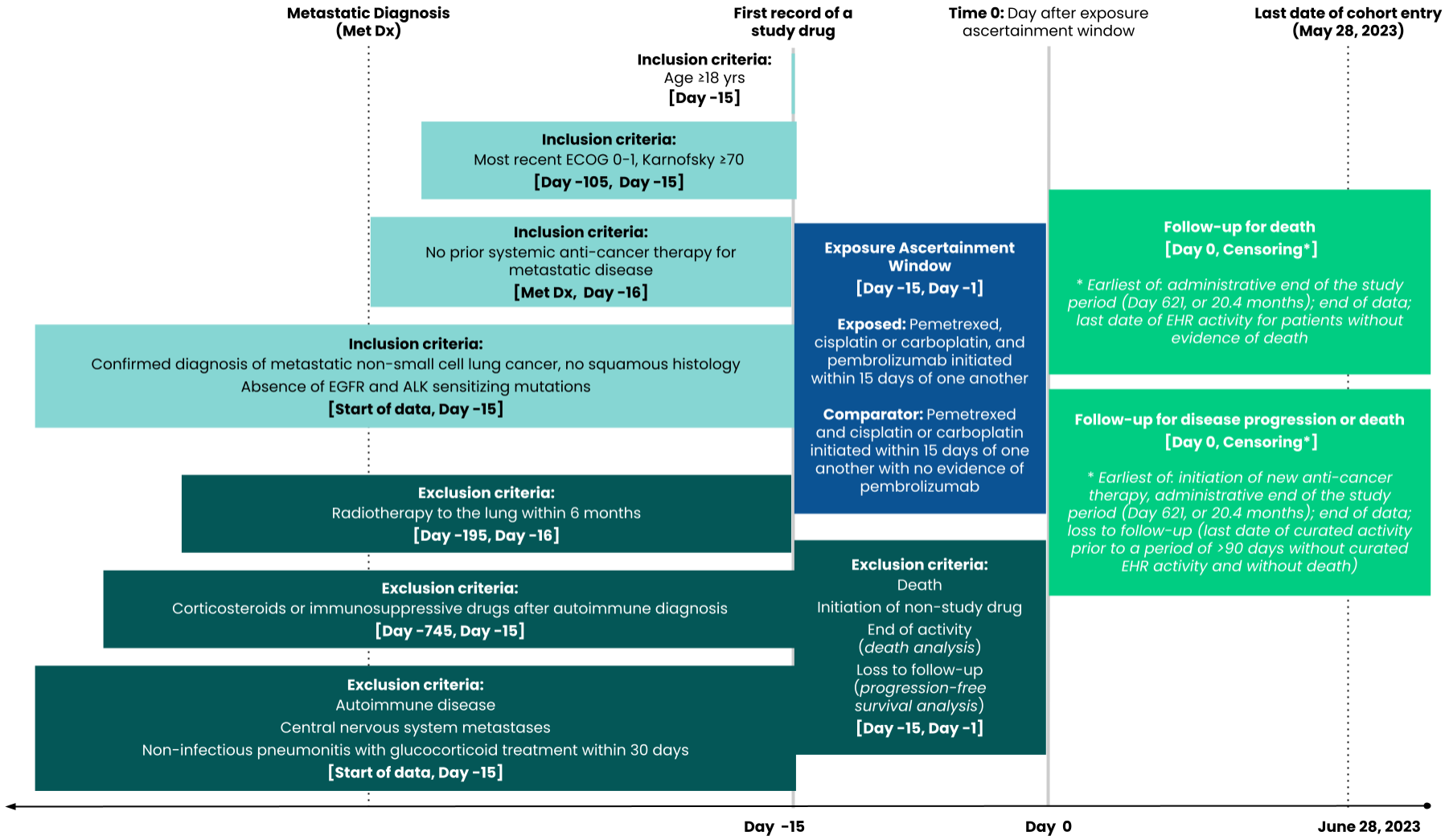


Figure 2. Real-world emulation study design



7.3 Setting

7.3.1 Context and rationale for definition of Time 0 (and other primary time anchors) for entry to the study population

Figure 2 displays the study design diagram for the primary analysis. Candidate exposed and comparator patients will be identified as those initiating treatment for metastatic disease with pemetrexed and carboplatin or cisplatin with or without pembrolizumab. Study exposure group

will be ascertained within a 15-day time window (i.e., 'exposure ascertainment window'), beginning on the day of the first record of a study drug (pemetrexed, carboplatin, cisplatin, or pembrolizumab) (Day-15), and ending 15 days later on Day -1. Patients experiencing death or a censoring event during this window will be excluded. Additional information on the exposure ascertainment window is provided in Section 7.4.1. Follow-up for progression and death will begin on Day 0 ('Time 0'). To avoid immortal time bias, Time 0 will occur after the exposure ascertainment window for all patients, irrespective of exposure status. The operational definition of Time 0 is provided in Table 3. The date of first study drug initiation will be restricted to dates from February 1, 2010 to May 28, 2023 to reflect current treatment paradigms at the time of the KEYNOTE-189 trial, and to allow a minimum available follow-up time of one month prior to the end of data for all patients, respectively. It should be noted that the first line of metastatic treatment can be initiated up to 30 days before the documented date of metastatic diagnosis per the data vendor's line of therapy definitions.

Table 3. Operational definition of Time 0

Study population name(s)	Time Anchor Description	Type of entry	Washout window	Incident with respect to…
Patients initiating first-line treatment for metastatic nonsquamous NSCLC without sensitizing EGFR or ALK mutations	Time 0 will be defined as the day after the end of the 15-day exposure ascertainment window. Exposure ascertainment is described in the variables section of this protocol.	Incident	[Metastatic diagnosis date, Day -16]	Metastatic diagnosis

7.3.2 Context and rationale for study inclusion criteria

Operational definitions for the study inclusion criteria are presented in Table 4. These inclusion criteria correspond to those applied in the KEYNOTE-189 RCT.

As indicated below, KEYNOTE-189 inclusion criteria that are not relevant in a real-world clinical setting (e.g., willingness and ability to provide tumor tissues) or are not captured in routine oncology care (e.g., measurable disease) will not be applied.

Table 4. Operational Definitions of Inclusion Criteria

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Histologically or cytologically confirmed diagnosis of stage IV nonsquamous NSCLC	NSCLC diagnosis. Histology not indicative of squamous cell carcinoma. Staging information indicative of metastatic disease.	[Start of data, Day -15]	Specific histology results cannot be identified for all patients in the dataset. As a result, patients with squamous cell histology will be excluded.	Patients with non-specific histology results who do not have nonsquamous disease and who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations. The specific histology values for all included patients will be reported descriptively.
Confirmation that EGFR or ALK-directed therapy is not indicated (documentation of absence of tumor activating EGFR mutations AND absence of ALK gene rearrangements OR presence of a Kirsten rat sarcoma viral oncogene homolog [KRAS] mutation)	EGFR and ALK biomarker tests with a result interpretation of 'Negative' or a KRAS biomarker test with a result interpretation of 'Positive.' For patients with multiple biomarker tests, the entry closest in time to study drug initiation will be used.	[Start of data, Day -15]	N/A	N/A
Measurable disease based on RECIST v.1.1	This criterion cannot be operationalized.	N/A	RECIST assessments are not performed in routine oncology care.	Patients with non-measurable disease who would have been ineligible for the trial may be included. This may result in longer estimates of rwPFS than were observed in the trial due to difficulties in objectively measuring progression for these patients.

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Had not received prior systemic treatment for their metastatic NSCLC. Subjects who received adjuvant or neoadjuvant therapy are eligible if completed ≥12 months prior to the development of metastatic disease	No first-line regimen for metastatic disease prior to study drug initiation. No record of adjuvant or neoadjuvant therapy in the 12 months before metastatic diagnosis.	No prior systemic treatment: [Metastatic diagnosis date, Day -16] No adjuvant/ neoadjuvant therapy: [Metastatic diagnosis date – 12 months, Metastatic diagnosis date – 1 day]	N/A	N/A
Provided tumor tissue from locations not radiated prior to biopsy	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
≥18 years of age on day of signing informed consent	Age at study drug initiation ≥18 years.	[Day -15]	N/A	N/A
Life expectancy of at least 3 months	This criterion cannot be operationalized.	N/A	Information on patients' life expectancy cannot be identified in the data source.	Patients with lower life expectancy who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations.

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Performance status of 0 or 1 on the ECOG performance status scale	ECOG performance status 0-1 or missing Karnofsky performance status ≥70 or missing For patients with multiple performance status records, the entry closest in time to study drug initiation will be used.	[Day -105, Day -15]	Performance status will be evaluated in the 90 days before study drug initiation to balance missingness and misclassification, given infrequent real-world assessments.	Patients with missing performance status records who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations. Sensitivity analyses restricting to patients with known values will be performed, sample size permitting.
Adequate organ function	No lab results indicating inadequate organ function as defined in the KEYNOTE-189 trial protocol.	[Day -45, Day -15]	Lab tests specified in the trial are not performed for all patients in routine oncology care. As a result, patients with lab results indicative of inadequate organ function will be excluded. Lab results will be evaluated in the 30 days before study drug initiation to balance missingness and misclassification.	Patients with inadequate organ function and therefore potentially reduced survival who would have been ineligible for the trial may be included. We do not expect this to affect a large number of patients as physicians are unlikely to start treatment for patients with poor organ function.
If female of childbearing potential, have a negative pregnancy test prior to receiving the first dose of study medication.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
If female of childbearing potential, be willing to use an adequate method of contraception.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
If male subject with a female partner(s) of child-bearing potential, must agree to use an adequate method of contraception.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Subject has voluntarily agreed to participate by giving written informed consent/assent for the trial.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A

N/A = not applicable
[a] See Appendix C for code list.

7.3.3 Context and rationale for study exclusion criteria

Operational definitions for the study exclusion criteria are presented in Table 5. These exclusion criteria correspond to those applied in the KEYNOTE-189 RCT.

As indicated below, KEYNOTE-189 exclusion criteria that are not relevant in a real-world clinical setting (e.g., known psychiatric or substance abuse disorder that would interfere with cooperation) will not be applied.

An additional exclusion criterion that was not relevant for the KEYNOTE-189 trial but is necessary for emulation using RWD is the exclusion of patients with evidence of a censoring event or death during the exposure ascertainment window (described in Section 7.4.1). Patients with evidence of a progression event during the exposure ascertainment window will not be excluded as these are likely latent progression events that preceded study drug initiation.

Table 5. Operational Definitions of Exclusion Criteria

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Predominantly squamous cell histology NSCLC. Mixed tumors will be categorized by the predominant cell type; if small cell elements are present, the subject is ineligible	Histology indicative of squamous cell carcinoma or small cell elements.	[Start of data, Day -15]	N/A	N/A
Currently participating and receiving study therapy or has participated in a study of an investigational agent or device within 4 weeks before randomization	This criterion cannot be operationalized.	N/A	Dates of clinical trial participation cannot be identified in the data source.	Patients who have recently participated in a clinical trial and who may have initiated treatment for metastatic disease in that setting who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations, and may bias comparative effect estimates if trial participation is related to the study treatment received.

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Before the first dose of trial treatment: a) Has received prior systemic cytotoxic chemotherapy for metastatic disease b) Has received antineoplastic biological therapy (e.g., erlotinib, crizotinib, cetuximab) c) Had major surgery (<3 weeks prior to first dose)	First-line regimen for metastatic disease prior to study drug initiation. Record of treatment with antineoplastic biological therapy. Record of major cancer-related surgery.	Previous treatment for metastatic disease: [Metastatic diagnosis date, Day -16] Biological therapy: [Start of data, Day -16] Major surgery: [Day -36, Day -16]	Non-cancer surgeries cannot be identified in the data source.	Patients with recent major non-cancer surgery who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations.
Received radiation therapy to the lung that is >30 Gy within 6 months of the first dose of trial treatment	Radiation therapy to the lung that is >3000 cGy	[Day -195, Day-16]	N/A	N/A
Completed palliative radiotherapy within 7 days of the first dose of trial treatment	This criterion cannot be operationalized.	N/A	Palliative radiation cannot be identified in the data source.	Patients who received palliative radiotherapy who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations.
Expected to require any other form of antineoplastic therapy while on study	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Received a live-virus vaccination within 30 days of planned treatment start. Seasonal flu vaccines that do not contain live virus are permitted.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Clinically active diverticulitis, intra-abdominal abscess, gastrointestinal obstruction, abdominal carcinomatosis	Diagnosis of clinically active diverticulitis, intra-abdominal abscess, or gastrointestinal obstruction. Abdominal carcinomatosis cannot be operationalized.	[Day -105, Day -15]	N/A	Patients with abdominal carcinomatosis who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations.
Has a known history of prior malignancy except if the subject has undergone potentially curative therapy with no evidence of that disease recurrence for 5 years since initiation of that therapy	Diagnosis type name of 'second primary malignant neoplasm.'	[Day -1840, Day -15]	N/A	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Known active CNS metastases and/or carcinomatous meningitis. Subjects with previously treated brain metastases may participate provided they are clinically stable for 2 weeks and have no evidence of new or enlarging brain metastases and are off steroids 3 days prior to dosing with study medication.	Diagnosis of brain, CNS, and/or spinal cord metastases. Active disease and clinical stability cannot be operationalized. Carcinomatous meningitis cannot be operationalized.	[Start of data, Day -15]	Active or clinically stable disease status cannot be identified in the data source. Carcinomatous meningitis cannot be identified in the data source.	Patients with stable CNS metastases who would have been eligible for the trial may be excluded, while patients with carcinomatous meningitis who would have been ineligible may be included. This may affect comparability between the study and trial populations.
Previously had a severe hypersensitivity reaction to treatment with another monoclonal antibody (mAb)	This criterion cannot be operationalized.	N/A	Hypersensitivity cannot be identified in the data source.	Patients with known hypersensitivity to treatment with another mAb who would have been ineligible for the trial may be included. We do not expect this to affect a large number of patients as physicians are unlikely to prescribe pembrolizumab to patients with a known hypersensitivity.

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Known sensitivity to any component of cisplatin, carboplatin or pemetrexed	This criterion cannot be operationalized.	N/A	Sensitivity cannot be identified in the data source.	Patients with known sensitivity to platinum therapies or pemetrexed who would have been ineligible for the trial may be included. We do not expect this to affect a large number of patients as physicians are unlikely to prescribe cisplatin, carboplatin, or pemetrexed to patients with a known sensitivity.
Active autoimmune disease that has required systemic treatment in the past 2 years (i.e., with use of disease modifying agents, corticosteroids or immunosuppressive drugs).	Diagnosis of autoimmune disease and treatment with corticosteroids or immunosuppressive drugs after diagnosis and within the past 2 years.	Autoimmune disease diagnosis: [Start of data, Day -15] Systemic treatment: [Day -745, Day -15]	Active disease cannot be identified in the data source.	Patients with inactive autoimmune disease who would have been eligible for the trial may be excluded. Patients treated with corticosteroids for reasons other than active autoimmune disease who would have been eligible for the trial may be excluded. This may affect comparability between the study and trial populations.
Receiving chronic systemic steroids	Treatment with a non-topical systemic steroid in each of the 6 months prior to study drug initiation.	[Day -195, Day -15]	Chronicity of systemic steroid treatment will be identified based on receipt of multiple or ongoing treatments in the 6 months prior to study drug initiation.	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Unable to interrupt aspirin or other nonsteroidal anti-inflammatory drugs (NSAIDs), other than an aspirin dose ≤ 1.3 g per day, for a 5-day period (8-day period for long-acting agents, such as piroxicam)	This criterion cannot be operationalized.	N/A	Inability to interrupt aspirin or NSAID treatment cannot be identified in the data source.	Patients who cannot interrupt aspirin or NSAID treatment who would have been ineligible for the trial may be included. This may affect comparability between the study and trial populations.
Unable or unwilling to take folic acid or vitamin B12 supplementation	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Prior treatment with any other anti-PD-1, or PD-L1 or programmed death ligand-2 (PD-L2) agent or an antibody targeting other immuno-regulatory receptors or mechanisms. Has participated in any other pembrolizumab (MK-3475) trial and has been treated with MK-3475	Treatment with pembrolizumab or any other anti-PD-1, PD-L1, PD-L2 agent or an antibody targeting other immuno-regulatory receptors or mechanisms.	[Start of data, Day -16]	N/A	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Active infection requiring therapy	Treatment with non-topical antibiotics, antifungals, or antivirals within two weeks after an infection diagnosis.	[Day -45, Day -15]	Active infections cannot be identified in the data source. Diagnosis of an infection followed by treatment with non-topical antibiotics, antifungals, or antivirals within two weeks will be evaluated in the 30 days before study drug initiation to reflect an active infection requiring treatment, and to balance missingness and misclassification.	N/A
Known history of Human Immunodeficiency Virus (HIV) (known HIV 1/2 antibodies positive)	Diagnosis of human immunodeficiency virus infection.	[Start of data, Day -15]	N/A	N/A
Known active Hepatitis B or C.	Diagnosis of acute hepatitis B or C.	[Day -105, Day -15]	Diagnosis of acute hepatitis B or C will be evaluated in the 90 days before study drug initiation to reflect active disease and to balance missingness and misclassification.	N/A

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Has a history or current evidence of any condition, therapy, or laboratory abnormality that might confound the results of the study, interfere with the subject's participation for the full duration of the study, or is not in the best interest of the subject to participate, in the opinion of the treating Investigator.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Known psychiatric or substance abuse disorder that would interfere with cooperation with the requirements of the trial	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Regular user of any illicit drugs or had a recent history of substance abuse (including alcohol)	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A
Symptomatic ascites or pleural effusion. A subject who is clinically stable following treatment for these conditions is eligible.	Diagnosis of ascites or pleural effusion.	[Day -105, Day -15]	Symptomatic conditions cannot be identified in the data source.	Patients with asymptomatic ascites or pleural effusion or who are clinically stable following treatment who would have been eligible for the trial may be excluded. This may affect comparability between the study and trial populations.

Trial Criterion	Real-world operationalization [a]	Assessment window	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Interstitial lung disease or a history of pneumonitis that required oral or intravenous glucocorticoids.	Diagnosis of interstitial lung disease or diagnosis of pneumonitis and record of treatment with glucocorticoids within 30 days.	[Start of data, Day -15]	Treatment with glucocorticoids will be evaluated in the 30 days after pneumonitis diagnosis to allow for delays in data entry.	Patients who receive glucocorticoids for conditions other than pneumonitis who may have been eligible for the trial may be excluded. This may affect comparability between the study and trial populations.
Is pregnant, breastfeeding, or expecting to conceive or father children within the projected duration of the study.	This criterion is not relevant in a real-world clinical setting.	N/A	N/A	N/A

N/A = not applicable
[a] See Appendix C for code list.

7.4 Variables

7.4.1 Context and rationale for exposures of interest

Operational definitions for the two treatment strategies that will be compared are presented in Table 6. Exposure will be defined based on treatment initiated during the 15-day exposure ascertainment window. While patients in the KEYNOTE-189 trial received study treatments on the same day, a 15-day window was selected to allow for potential delays in data entry and treatment administration variability that may occur in routine practice. In a feasibility assessment of 1,085 patients in the data source who initiated pemetrexed and carboplatin or cisplatin plus pembrolizumab as first-line therapy, 94.3% started all medications within 15 days. A shorter exposure window may be too restrictive and result in few eligible patients, while longer time periods may result in treatment patterns in the study population that are vastly different from the trial. As discussed above, patients who experience a censoring event or death during the exposure ascertainment window will be excluded from the study to align the start of follow-up for both exposure groups.

Table 6. Operational Definitions of Exposures

Group name(s)	Details	Washout window	Assessment Window	Incident with respect to…	Source of algorithm	Validity concerns and how they will be addressed
Exposed	Patients initiating treatment with pemetrexed and carboplatin or cisplatin plus pembrolizumab and no other treatments within 15 days of one another, in the metastatic setting.	N/A, treatment in first-line metastatic setting	[Metastatic diagnosis date, Day -1]	Metastatic Diagnosis	Curated regimen definition; clinical experts	Real-world exposure group definitions allow for more flexibility in treatment timing than in the trial. The potential impact of this on results will be explored through a sensitivity analysis shortening the duration of the exposure ascertainment window (Table 11).
Comparator	Patients initiating treatment with pemetrexed and carboplatin or cisplatin with no evidence of pembrolizumab or any other treatments within 15 days of one another, in the metastatic setting.	N/A, treatment in first-line metastatic setting	[Metastatic diagnosis date, Day -1]	Metastatic Diagnosis	Curated regimen definition; clinical experts	

N/A = not applicable

7.4.2 Context and rationale for outcome of interest

Operational definitions for the outcomes of interest, rwOS and rwPFS, are presented in Table 7. These outcomes correspond to the primary outcomes of the KEYNOTE-189 trial.

Table 7. Operational Definitions of Outcome

Outcome name	Details	Primary outcome	Type of outcome	Washout window	Source of algorithm	Validity concerns and how they will be addressed
rwOS	Time from Day 0 to death	Yes	Time-to-event	N/A	Dates of death are sourced from EMR, curation, and third-party death data.	Death data is not NDI validated. The sources of death data may have different capture rates over the study period. Deaths by year of study entry will be reported to contextualize findings.
rwPFS	Time from Day 0 to disease progression or death	Yes	Time-to-event	N/A	Progression events are curated in the data. All tumor progression events after initial cancer diagnosis are captured.	Real-world progression is not evaluated at fixed intervals as was done in the trial. The frequency and timing of real-world progression assessment results by exposure group will be reported to contextualize study findings.

7.4.3 Context and rationale for follow-up

Follow-up for rwOS will begin on Day 0, and will continue until the earliest of:

1. Date of documented death;
2. The administrative end of the study period (Day 621, or 20.4 months, of follow-up), to align with the maximum follow-up time at which OS was evaluated in the KEYNOTE-189 trial;
3. End of the data (June 28, 2023);
4. Last date of EHR activity for patients without evidence of death.

Follow-up for rwPFS will begin on Day 0, and will continue until the earliest of:

1. Date of documented progression or death (Table 7);
2. Initiation of any new anti-cancer therapy;
3. The administrative end of the study period (Day 621, or 20.4 months, of follow-up), to align with the maximum follow-up time at which PFS was evaluated in the KEYNOTE-189 trial;

4. End of the data (June 28, 2023);
5. Loss to follow-up: The last date of curated EHR activity prior to a period of >90 days without curated EHR activity and without death.

Operational definitions for the study censoring criteria are presented in Table 8. These censoring criteria correspond to those applied in the KEYNOTE-189 RCT where applicable.

Table 8. Operational Definitions of Censoring Criteria

Trial Criterion	Real-world operationalization	Rationale for real-world operationalization (where applicable)	Validity concerns and how they will be addressed (where applicable)
Initiation of a new anti-cancer therapy	Treatment with any new anti-cancer therapy.	N/A	N/A
End of study	Administrative end of study (Day 621, or 20.4 months, of follow-up) or June 28, 2023.	Align with maximum time in KEYNOTE-189 or end of available data in real-world data source.	N/A
Withdrawal of consent	This criterion is not relevant in a real-world clinical setting.	N/A	N/A
Loss to follow-up	A period of >90 days without curated EHR activity and without death.	Curated activity in the real-world data source indicates points at which progression can be recorded. Metastatic lung cancer patients likely have contact with the healthcare system at least every 90 days for lab work, prescription refills, outpatient visits, or scans. Periods greater than 90 days may indicate loss to follow-up during which censoring or outcome events cannot be captured.	Patients who use the health care system less frequently will be censored. A sensitivity analysis will be conducted expanding the period without curated activity to >180 days (Table 11).

7.4.4 Context and rationale for covariates (confounding variables and effect modifiers, e.g., risk factors, comorbidities, comedications)

Operational definitions for the study covariates are presented in Table 9. Covariates were chosen based on the primary trial publication and the research team's substantive knowledge of potential confounders. For time-varying characteristics, the value closest in time prior to study drug initiation will be used.

Table 9. Operational Definitions of Covariates

Characteristic	Details/Levels [a]	Type of variable	Assessment window
Patient Demographic Characteristics			
Age	Age at study drug initiation 18 to <65 yrs, ≥65 yrs	Continuous, Categorical	[Day -15]
Gender	Female, male, missing	Categorical	[Start of data, Day -15]
Race	American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, Black or African American, White, other or unknown race	Categorical	[Start of data, Day -15]
Ethnicity	Hispanic or Latino, not Hispanic or Latino, unknown	Categorical	[Start of data, Day -15]
Region	Northeast, south, midwest, west, missing	Categorical	[Start of data, Day -15]
Clinical Characteristics			
Oncology clinic site type	Academic, community, missing	Categorical	[Start of data, Day -15]
Performance status	ECOG performance status 0, or 1, or Karnofsky performance status 70, 80, 90, 100, or missing.	Categorical	[Day -105, Day -15]
Smoking status at initial diagnosis	Current smoker/former smoker, non-smoker, missing	Categorical	[Start of data, Initial diagnosis date]

Characteristic	Details/Levels [a]	Type of variable	Assessment window
Number of outpatient visits in past year	N/A	Continuous	[Day -380, Day -15]
Disease Characteristics and Treatment History			
Year of study treatment initiation	N/A	Categorical	[Day -15]
Previous therapy for nonmetastatic disease	Thoracic radiotherapy, neoadjuvant therapy, adjuvant therapy	Dichotomous (for each treatment type)	[Start of data, Metastatic diagnosis date - 1]
Platinum based therapy	Cisplatin, carboplatin	Dichotomous	[Day -15, Day -1]
PD-L1 tumor proportion score	Positive, negative, missing	Categorical	[Start of data, Day -15]
Histologic feature	Adenocarcinoma, other	Categorical	[Start of data, Day -15]
Disease stage at initial diagnosis	Disease stage I-IV	Categorical	[Initial diagnosis date - 30, Initial diagnosis date + 90]
Time between metastatic diagnosis and treatment initiation	N/A	Continuous	[Metastatic diagnosis date, Day -15]

Characteristic	Details/Levels [a]	Type of variable	Assessment window
Time interval between initial diagnosis and metastatic diagnosis (approximate disease-free interval)	N/A	Continuous	[Initial diagnosis date, Metastatic diagnosis date]
Recurrence type	Recurrent, de novo metastatic, missing	Categorical	[Initial diagnosis date - 30, Initial diagnosis date + 90]
Liver metastases	Yes, no	Dichotomous	[Start of data, Day -15]
Number of metastatic disease sites	1, 2, ≥3	Continuous, Categorical	[Start of data, Day -15]

N/A = not applicable

[a] See Appendix C for code list.

7.5 Data analysis

7.5.1 Context and rationale for analysis plan

Analytic Population

Primary analyses will be conducted in the real-world study population based on the first study treatment initiated. This approach is intended to emulate the intent-to-treat analysis conducted in the KEYNOTE-189 trial, where patients were analyzed based on their randomized treatment assignment. Crossover to pembrolizumab monotherapy was permitted among patients in the comparator group following disease progression. To investigate the potential for crossover between comparison groups prior to disease progression in the real-world population, descriptive analyses of all first-line treatments initiated and time between treatments will be conducted.

IPTW will be used to approximate full conditional exchangeability between the comparison groups and facilitate estimation of the population average treatment effect.[13–15] Propensity scores (PS) reflecting the conditional probability of initiating treatment with pembrolizumab combination therapy will be calculated via multivariable logistic regression. Exposure to pembrolizumab combination therapy will be regressed on *a priori* identified potential confounders (Table 9). Inverse probability of treatment weights will be calculated as the inverse of the propensity score (1/PS)

for patients in the exposed group and as the inverse of one minus the propensity score (1/1-PS) for comparator patients.[16] Patients are therefore weighted by the inverse probability of initiating the treatment they actually started, conditional on the observed covariates included as independent variables in the PS model. This approach aims to create a pseudo-population with full exchangeability on measured confounders.

Confounder balance will be assessed by comparing the absolute standardized difference (ASD) in the distribution of categorical variable levels and the mean of continuous variables between the weighted exposure groups.[13] Randomization in the KEYNOTE-189 trial was stratified by PD-L1 expression, platinum-based drug (cisplatin vs. carboplatin), and smoking history (never vs. former or current); therefore, particular attention will be paid to balance in these variables. If balance is not achieved (ASD > 0.1), alternative specifications of the PS model, e.g., including variable transformations and interaction terms or estimating propensity scores within strata, will be explored. Confounders with insufficient balance may also be included as covariates in the outcome models or stratified regression analysis may be used.

Descriptive Analyses

The distribution of baseline patient demographic and clinical characteristics will be compared between the unweighted and weighted real-world populations and the KEYNOTE-189 trial population. Characteristics to be compared are described in Table 9. Differences will be assessed using t-tests, chi squared tests, and accompanying 95% confidence intervals.

The distribution of baseline patient demographic and clinical characteristics will be compared between comparator patients who index before 2017 and on/after 2017 to understand if differences exist between those who received the comparator treatments before and after pembrolizumab approval in this indication. Characteristics to be compared are described in Table 9. Differences will be assessed using t-tests, chi squared tests, and accompanying 95% confidence intervals.

Reasons for exclusion and censoring will be presented in a Consolidated Standards of Reporting Trials (CONSORT) diagram and distributions of missingness of inclusion/exclusion criteria and potential confounders will be calculated and compared by exposure group.[17]

Treatment Effectiveness

Similar to the KEYNOTE-189 trial, comparative treatment effectiveness of pemetrexed and carboplatin or cisplatin plus pembrolizumab versus pemetrexed and carboplatin or cisplatin alone will be estimated using median rwOS and rwPFS and by comparing the hazards of death and progression in the two exposure groups. Follow-up will be administratively censored at 20.4 months to emulate maximum follow-up in the KEYNOTE-189 trial at the time that results were reported.[2] We will emulate subgroup analyses performed in the trial, except where sample size is less than 10 in each exposure group. Post-hoc analyses to explore effect modification will be conducted for characteristics where distributions differ between the trial and RW populations.

Details of the analytic approach are presented in Table 10.

Table 10. Primary and subgroup analysis specification

Hypothesis:	Median rwOS is longer and the hazard of death is lower among patients who initiated pemetrexed and carboplatin or cisplatin plus pembrolizumab versus pemetrexed and carboplatin or cisplatin alone. Median rwPFS is longer and the hazard of progression is lower among patients who initiated pemetrexed and carboplatin or cisplatin plus pembrolizumab versus pemetrexed and carboplatin or cisplatin alone.
Exposure contrast:	Patients initiating pemetrexed and carboplatin or cisplatin plus pembrolizumab compared with patients initiating pemetrexed and carboplatin or cisplatin alone in the first line metastatic setting.
Outcome(s):	rwOS and rwPFS
Analytic software:	Aetion Substantiate Version 5.01 (or latest version)
Model(s):	Median rwOS and rwPFS will be estimated using a Kaplan-Meier estimator, weighted by time-fixed inverse probability of treatment weights. The weighted survival probability $S_a(t)$ for exposure group 'A=a' at time 't' will be as follows: $$S_a(t) = \prod_t \left(1 - \frac{d_{ta}}{r_{ta}}\right)$$ where $d_{ta} = \sum_{i=1}^{N} w_{it} \cdot Y_{it} \cdot I(A_{it} = a)$ denotes the weighted number of events and $r_{ta} = \sum_{i=1}^{N} w_{it} \cdot I(A_{it} = a)$ denotes the weighted risk set.[18–20] This is equivalent to calculating the Kaplan-Meier estimator in the IPTW weighted population.[19] A non-parametric bootstrap will be used to derive 95% confidence intervals. A Cox proportional hazards model will be used to estimate the hazard ratio at 20.4 months in the IPTW weighted population as follows: $$h(t, L_0) = h_0(t) * e^{\beta_1 A_0 + \beta_2^T L_0}$$ where $h(t, L_0)$ is the hazard of progression or death at discrete time interval 't' conditional on the vector of potential confounders 'L_0' assessed at baseline; $h_0(t)$ is the baseline hazard at discrete time interval 't'; and A_0 is an indicator for treatment initiation, coded as '1' and '0' for the exposure and comparator group, respectively. Patients with an outcome or censoring event on Day 0 will be assigned a follow-up time of 0.5 days. If balance for stratification factors used in the trial is not achieved after applying IPTW, a stratified Cox proportional hazards model will be considered. The proportional hazards assumption will be checked using plots that display the scaled Schoenfeld residuals vs. time for each covariate; if violations are detected, a time-dependent or stratified Cox proportional hazards model will be considered. The Efron method of handling ties will be used.

Confounding adjustment method	Time-fixed inverse probability of treatment weights will be used to adjust for confounding. Individual-level weights will be estimated by the following formula: $$W^A = \frac{1}{f(A|L_0)}$$ where A is the first study treatment that the patient initiated and L_0 is a vector of baseline confounders. The quantity in the denominator $f(A|L_0)$—the probability of exposure to treatment A given baseline confounders L_0—will be estimated using a logistic regression model with A as the dependent variable and the vector L_0 as the independent variables. The distribution of weights will be used to identify potential extreme weights. If extreme weights are identified, weight truncation and/or stabilization will be considered All potential confounder variables will be considered for inclusion in the weight estimation (Table 9). However, as it is not possible to predict the quantity of missing values and sparseness of the data at the time of writing this protocol, the precise functional form of the final regression model will be determined at the time of analysis. Thus, variables with high missingness will be excluded from the final model. Additionally, categorical variables may be collapsed to ensure convergence of the propensity score model.
Missing data methods	Data missingness was assessed as part of an initial feasibility assessment (Appendix B); therefore, key variables are expected to have a high degree of completeness. If substantial missingness results in an insufficient sample size for the complete analytic dataset, alternative variable specifications (e.g., changing the time frames over which variables are assessed) or exclusion of variables may be considered.
Subgroup Analyses	Cox proportional hazards models will be assessed in the following subgroups to align with the trial, as feasible. International region and brain metastases were included as subgroups in the trial but cannot be emulated in the RWD source. 1. Age (18 to <65 yrs, ≥65 yrs) 2. Sex (male, female) 3. Performance status (ECOG 0 or 1) 4. Smoking status (current or former, never) 5. PD-L1 tumor proportion score (positive, negative) 6. Platinum-based drug (carboplatin, cisplatin)

7.5.2 Context and rationale for sensitivity analyses

Sensitivity analyses will be conducted to explore the potential impact of several key study design elements. Planned analyses and their respective goals are presented in Table 11.

Table 11. Sensitivity analyses – rationale, strengths and limitations

Description	Primary Analysis	Sensitivity Analysis	Rationale	Strengths of the sensitivity analysis compared to the primary	Limitations of the sensitivity analysis compared to the primary
Contemporaneous cohort	Study population includes patients with first study drug initiation from February 1, 2010 to May 28, 2023.	The study population will be restricted to patients with first study drug initiation from February 1, 2017 to May 28, 2023 to coincide with Food and Drug Administration (FDA) accelerated approval of pembrolizumab in this indication.	Removing historical comparator patients will decrease potential confounding due to changes in treatment paradigms over time.	This analysis will provide effect estimates using a control group that more closely resembles the exposed group with respect to treatment standards.	Restriction to contemporaneous comparator patients may reduce the sample size and introduce other unmeasured sources of confounding due to differences between patients who do and do not initiate newly available treatment.
Limit eligible cohort entry dates to six months prior to the end of the data	Eligible cohort entry period limited to dates at least one month prior to the end of the data.	The eligible cohort entry period will be limited to dates at least six months before the end of the data (February 1, 2010 to December 28, 2022).	Patients identified toward the end of the available study period may not have adjudicated death data, leading to underestimates of death.	Greater minimum follow-up time and opportunity for death to be identified.	The sample size will be reduced relative to the primary analysis.
Modified exposure ascertainment window (one day)	Exposure status is ascertained over a 15-day window following first study drug initiation.	The exposure ascertainment window will be shortened to occur over a single day.	Shortening the exposure ascertainment window will exclude patients who initiate pemetrexed and carboplatin or cisplatin with pembrolizumab greater than 1 day apart from the exposed group.	This analysis will more closely align the treatment strategy in the exposed group to that of the trial.	Patients who initiate pembrolizumab greater than 1 day after other study treatments will be eligible for inclusion in the comparator group, which may bias results towards the null.

Description	Primary Analysis	Sensitivity Analysis	Rationale	Strengths of the sensitivity analysis compared to the primary	Limitations of the sensitivity analysis compared to the primary
Modified exposure ascertainment window (30 days)	Exposure status is ascertained over a 15-day window following first study drug initiation.	The exposure ascertainment window will be increased to occur over 30 days.	Increasing the exposure ascertainment window will include patients who initiate pemetrexed and carboplatin or cisplatin with pembrolizumab greater than 15 days apart in the exposed group.	This analysis will increase follow-up time for patients who initiate study drugs further apart in time and will employ a more flexible definition of the exposure ascertainment period that aligns with the data vendor line of therapy definition.	Patients who initiate study treatments further apart in time will be eligible for inclusion in the exposed group, which may less accurately reflect the treatment strategy used in the trial.
Loss to follow-up censoring definition	Patients are censored on the last date prior to a period of >90 days without curated EHR activity and without death.	Censor patients on the last date prior to a period of >180 days without curated EHR activity and without death.	Lengthening the time period without curated EHR activity will allow patients who are using healthcare less frequently to have outcomes recorded after a gap in 180 days of activity.	Progression and death events that occur after a gap of 90 days without curated EHR activity will be included.	Lengthening the time period without curated EHR activity will increase the possibility of unobserved events or censoring reasons and potentially overestimate rwPFS.
ECOG performance status assessment window	ECOG or Karnofsky performance status for study inclusion will be assessed within 90 days from first study treatment initiation.	ECOG or Karnofsky performance status for study inclusion will be assessed within 30 days from first study treatment initiation.	Shortening the assessment window for ECOG performance status may more accurately reflect patients' status at the time of study treatment initiation and will provide information on the sensitivity of results to this key inclusion criteria.	Assessing ECOG status closer in time to study treatment initiation may create a study population that is more similar to the trial population.	A larger number of patients may be missing ECOG performance status.

Description	Primary Analysis	Sensitivity Analysis	Rationale	Strengths of the sensitivity analysis compared to the primary	Limitations of the sensitivity analysis compared to the primary
Complete case - ECOG performance status	Patients with missing ECOG or Karnofsky performance status in the 90 days prior to first study treatment initiation are included.	Patients with missing ECOG or Karnofsky performance status in the 90 days prior to first study treatment initiation will be excluded.	Removing patients with missing ECOG performance status may reduce misclassification of disease severity and will enable descriptive comparison of the study population with and without these patients.	Excluding patients with missing ECOG may create a study population that is more similar to the trial population.	The sample size will be reduced.

7.6 Data source(s)

7.6.1 Context and rationale for data source(s)

The data source used in this study obtains clinical data from multiple partnerships across the U.S. These partnerships include data from organizations that aggregate data from oncology practices. The data vendor receives the complete EHR from each practice, including unstructured notes and scanned documents attached to the EHR. Structured and unstructured data are semi-normalized using an Observational Medical Outcomes Partnership (OMOP)-based common data model to improve uniformity across original data sources. Data curation (i.e., abstraction) of unstructured data from provider notes, pathology and imaging reports, and scanned documents is performed by trained reviewers. Natural Language Processing (NLP) is used to identify relevant strings of data within the patient record, but all curated data points are reviewed, interpreted, and documented by a human curation team. Patients are not filtered on the basis of data completeness to allow for flexibility when selecting the analytic sample.

Reason for selection: The dataset was considered due to its focus on oncology EHR data and was further considered fit-for-purpose after a detailed feasibility assessment that considered available sample size and completeness and quality of key inclusion criteria, exclusion criteria, key confounders, and outcomes (Appendix B). Linked exposure and outcome data will not be accessed prior to conducting final analyses.

Strengths of data source(s): The dataset includes individuals from all regions of the U.S. and both community and academic providers. Due to the oncology focus of the data vendor, this dataset provides information on important diagnostic, prognostic, and clinical characteristics among NSCLC patients. The data include several important curated fields, including ECOG, line of therapy, and progression, using a broad range of clinical documentation (e.g., physician notes, pathology reports, etc.).

Limitations of data source(s): Algorithms used by the data vendor to derive treatment regimens and certain key variables such as those used to define key inclusion and exclusion criteria, potential confounders, and mortality have not been validated. The data are also limited by the accuracy of data collection in the original EHRs, the subjective nature of data abstraction, and, for some variables, the inability to determine whether missing values indicate the true absence of a condition or missing data.

Data source provenance/curation: The data vendor programmatically tracks all inbound data throughout the data pipeline for data provenance and quality control using internal identifiers. Additionally, the data curation process is operationalized to ensure consistency and high inter-rater reliability. To reduce data entry errors following abstraction, data curation software is utilized. Quality control analyses on various subsets (e.g., a random sample, a specific cohort, only curated patients) before data are finalized for external use.

Table 12. Metadata about data sources and software

Data Source(s):	[Redacted]
Study Period:	Start of data-June 28, 2023
Eligible Cohort Entry Period:	February 1, 2010-May 28, 2023
Data Version (or date of last update):	Q3 2023
Data sampling/extraction criteria:	Described above
Type(s) of data:	Clinical data sourced from oncology practices in the U.S.
Data linkage:	N/A
Conversion to Common Data Model:	Observational Medical Outcomes Partnership-based Common Data Model
Software for data management:	Aetion Substantiate Version 5.01 (or latest version)

7.7 Data management

Raw data review

At Aetion, as part of the data ingestion process, raw data review is routinely conducted to understand contents of the data table(s), establish relationships, and help inform the database connection specification. Scientific integrity checks are performed to understand if the contents of the data shipment are consistent with the expected data as laid out in the applicable data usage agreement. Some of the key characteristics explored in this process include:

- Table structure (number of rows, columns, column names etc.)
- Summary counts per table (i.e., non-missing counts, unique counts)
- Variable distribution (e.g., min, mean, median, max for numeric variables; top frequencies for categorical variables)
- Date range (min, max and distribution over a time period)
- Missingness percentage of attributes

Database connection (DBC) process

Following receipt and review of the raw data, a data connector specification is drafted by a data scientist. The specification provides a map to Engineering for transformation of raw data to the Aetion longitudinal patient timeline. It includes information such as:

- Overall schema including tables (event types), rows (events), and columns (attributes); derivation of attributes to improve data flexibility on Platform and rationale for any attributes or events that are dropped
- Event dates that define how data will be reflected on the longitudinal patient timeline, and any minimal processing rules (e.g., drop an event that does not include a start or end date)
- Skeleton structure diagram that represents the logical view of the entire database, defining how the data tables are organized and related in the longitudinal patient timeline and how the relations among them are associated
- Information for user interface and labeling
- Codes and definitions; typically used to substitute users' having to look-up multiple resources to understand/process data

Validation of the DBC is completed to ensure that the implementation of DBC logic leads to transformed data output that connects to and behaves within Aetion Substantiate exactly as intended. Raw data are never loaded as-is; rather, data are transformed (via the DBC) into a longitudinal sequence of healthcare data points for each patient. DBC validation is required to confirm that this transformation was performed correctly. This helps to ensure validity/accuracy of the connected data and its importance cannot be ignored. Validation is performed via double programming, where two different people work independently from the same DBC specification and then compare their output. The DBC is considered validated if the outputs are identical. If the outputs are not identical, then the source of the discrepancy is investigated and resolved.

Following validation, the specification files are used to create an Aetion data dictionary for the dataset. In addition, throughout the data connector spec / creation process, any issues or decisions that have to be made that are not otherwise specified in the Specification files (e.g., how missing dates are handled), are noted in the data dictionary.

7.8 Quality control

Prior to deployment on Aetion Substantiate, a manual test of certain platform features and dataset values is conducted to ensure they are visible and testable on the front-end. This test is run following any deployment activity (such as a version update and/or data/shard update). Checks include:

- Baseline values for database information (dataset name, patient counts, earliest and latest event dates)
- Database configuration (specified dataset values)
- Measure, Cohort, and Analysis Generations to confirm this functionality using the dataset
- Output from generated analysis output
- Coding Systems, if applicable

The implementation of all variables, cohorts, and analytic plans will be individually checked by two analysts. Any discrepancies will be discussed with the analysts and study lead to ensure alignment with the study design outlined in the protocol.

7.9 Study size and feasibility

The KEYNOTE-189 trial target sample size was 570 patients with 2:1 randomization. This sample size was estimated based on 90% power to detect a hazard ratio for progression or death of 0.70 at a one-sided alpha level of 0.0095 (assuming 468 events) and a hazard ratio of 0.70 for death at a one-sided alpha level of 0.0155 (assuming 416 deaths).[2] A total of 616 patients were included in the trial.

Sample size requirements to detect a range of hazard ratios relevant to the KEYNOTE-189 trial are presented in Table 13. If the unweighted study sample size falls below the lowest estimate, corresponding to the required sample size to detect the point estimate for the hazards of death observed in the trial with 80% power, implementation will pause. The study team and CARE Steering Committee will then consider the value of continuing the study with potentially insufficient power, given the lower primacy of statistical significance in an emulation setting.

Table 13. Sample size requirements

	Power	Hazard ratio for death	Ratio of exposed to unexposed	Alpha	Prevalence of death among the unexposed [a]	Total sample size required [b]
Trial sample size calculation, point estimate	90%	0.70	1:1	5%	50.6%	653

	Power	Hazard ratio for death	Ratio of exposed to unexposed	Alpha	Prevalence of death among the unexposed [a]	Total sample size required [b]
Trial sample size calculation, point estimate	80%	0.70	1:1	5%	50.6%	488
Trial result, upper confidence limit	90%	0.64	1:1	5%	50.6%	418
Trial result, upper confidence limit	80%	0.64	1:1	5%	50.6%	312
Trial result, point estimate	90%	0.49	1:1	5%	50.6%	164
Trial result, point estimate	80%	0.49	1:1	5%	50.6%	122

[a] As reported in the trial.

[b] Calculated using the *powerSurvEpi* R package[21], based on the sample size formula for proportional hazards modeling derived by Latouche et al.[22]

In feasibility analyses (Appendix B) among adults with metastatic NSCLC, negative results for EGFR/ALK mutations, no CNS metastases, and no evidence of ECOG performance status >1 in the data source, there were 1085 individuals with evidence of initiating pembrolizumab, pemetrexed, and platinum-based chemotherapy, and 564 individuals who initiated pemetrexed and chemotherapy only. This real-world study will include data for all individuals meeting study inclusion and exclusion criteria.

8. Strengths and Limitations

This emulation study is limited by several inherent differences between trial settings and non-interventional studies leveraging RWD.

In the absence of randomization, IPTW will be used to control for confounding. However, important confounders may be unavailable or imperfectly measured in the RWD source, which may result in residual confounding. There may be residual confounding by indication arising from causes of progression or death that also impact physicians' treatment decisions. Several measures of disease severity (e.g., performance status, stage at initial diagnosis, disease-free interval) will be used to generate treatment probabilities, but these may not sufficiently control for confounding by indication. In particular, real-world patients who did not initiate pembrolizumab after approval may systematically differ from those treated with pembrolizumab for reasons that are not captured in the health record. These patients may also be different from those enrolled in the KEYNOTE-189 RCT, limiting successful emulation of the trial results.

Several trial design elements cannot be perfectly emulated or operationalized due to limitations of the data source. First, the trial included international sites, while the RWD source is restricted to EHR data from U.S. oncology clinics. Second, some study inclusion and exclusion criteria and potential effect modifiers, such as measurable disease, cannot be operationalized and others, such as performance status, lab values, and patient symptoms, may be missing in the RWD due to infrequent real-world clinical assessment or inadequate capture in the EHR. Excluding patients who have missing values for these key variables may result in systematic differences between the trial and real-world populations. A preliminary data feasibility assessment was conducted prior to protocol finalization to ensure that key study eligibility criteria and potential confounders had a high degree of completeness (Appendix B). Third, while trial treatments were administered on the same day, these treatment strategies must be approximated (within 15 days) due to differences in medication dosing schedules, insurance delays, and provider decision-making in routine clinical practice. While death is a consensus variable across multiple sources and validated in this data source, it is possible that death capture varies over the study period. This will be contextualized by reporting of the death rate by year of study entry to understand potential differences in data availability. Finally, progression surveillance is conducted less frequently and regularly in real-world practice than was performed in the trial, which may affect estimates of rwPFS.

At the same time, this study proposes to use a high-quality, RWD source specifically designed for conducting RWD analyses in oncology. The data source includes curated, quality-controlled data elements (e.g., ECOG, progression) unique to oncology studies and necessary to this emulation. The preliminary feasibility assessment indicated low missingness of key variables, the ability to create a trial similar population through careful operationalization of trial characteristics, and a larger available sample size than was enrolled in the trial. Differences between the trial and real-world emulation, including rates of study treatment discontinuation and crossover, will be transparently reported and compared to contextualize final results (Appendix B).

9. Protection of Human Subjects

This study will use de-identified secondary data and therefore does not constitute research involving human subjects. Institutional review board exemption will be requested.

10. Reporting of Adverse Events

Detection and reporting of adverse events do not apply as this study involves secondary use of real-world data from an existing data collection infrastructure.

11. References

1. Merola D, Campbell U, Gautam N, et al. The Aetion Coalition to Advance Real-World Evidence through Randomized Controlled Trial Emulation Initiative: Oncology. *Clin Pharmacol Ther*. 2023;113(6):1217-1222. doi:10.1002/cpt.2800

2. Gandhi L, Rodríguez-Abreu D, Gadgeel S, et al. Pembrolizumab plus Chemotherapy in Metastatic Non–Small-Cell Lung Cancer. *N Engl J Med*. 2018;378(22):2078-2092. doi:10.1056/NEJMoa1801005

3. National Institute for Health and Care Excellence. *NICE Real-World Evidence Framework*. National Institute for Health and Care Excellence; 2022. Accessed December 21, 2022. https://www.nice.org.uk/corporate/ecd9/resources/nice-realworld-evidence-framework-pdf-1124020816837

4. Center for Drug Evaluation and Research, Center for Biologics Evaluation, Oncology Center of Excellence. *Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products*. U.S. Food and Drug Administration; 2023. Accessed December 21, 2023. https://www.fda.gov/media/171667/download

5. Purpura CA, Garry EM, Honig N, Case A, Rassen JA. The Role of Real-World Evidence in FDA-Approved New Drug and Biologics License Applications. *Clin Pharmacol Ther*. 2022;111(1):135-144. doi:10.1002/cpt.2474

6. Visvanathan K, Levit LA, Raghavan D, et al. Untapped Potential of Observational Research to Inform Clinical Decision Making: American Society of Clinical Oncology Research Statement. *J Clin Oncol*. 2017;35(16):1845-1854. doi:10.1200/JCO.2017.72.6414

7. Castellanos EH, Wittmershaus BK, Chandwani S. Raising the Bar for Real-World Data in Oncology: Approaches to Quality Across Multiple Dimensions. *JCO Clin Cancer Inform*. 2024;(8):e2300046. doi:10.1200/CCI.23.00046

8. Franklin JM, Pawar A, Martin D, et al. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clin Pharmacol Ther*. 2020;107(4):817-826. doi:10.1002/cpt.1633

9. Merola D, Campbell UB, Lenis D, Schneeweiss S, Wang SV. Calibrating real-world evidence against randomized trials: A pilot to inform the CARE Initiative. *Manuscr Submitt Publ*. Published online 2024.

10. Stang A, Poole C, Kuss O. The ongoing tyranny of statistical significance testing in biomedical research. *Eur J Epidemiol*. 2010;25(4):225-230. doi:10.1007/s10654-010-9440-x

11. Merck Sharp & Dohme LLC. *A Randomized, Double-Blind, Phase III Study of Platinum+Pemetrexed Chemotherapy With or Without Pembrolizumab (MK-3475) in First Line Metastatic Non-Squamous Non-Small Cell Lung Cancer Subjects (KEYNOTE-189)*. clinicaltrials.gov; 2023. Accessed December 31, 2023. https://clinicaltrials.gov/study/NCT02578680

12. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228-247. doi:10.1016/j.ejca.2008.10.026

13. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015;34(28):3661-3679. doi:10.1002/sim.6607

14. Hernán MA, Robins JM. *Causal Inference: What If*. Chapman & Hall/CRC; 2020.

15. Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Stat Med*. 2014;33(7):1242-1258. doi:10.1002/sim.5984

16. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivar Behav Res*. 2011;46(3):399-424. doi:10.1080/00273171.2011.568786

17. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Published online 2010.

18. Xie J, Liu C. Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Stat Med*. 2005;24(20):3089-3110. doi:10.1002/sim.2174

19. Westreich D, Cole SR, Tien PC, et al. Time Scale and Adjusted Survival Curves for Marginal Structural Cox Models. *Am J Epidemiol*. 2010;171(6):691-700. doi:10.1093/aje/kwp418

20. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed*. 2004;75(1):45-49. doi:10.1016/j.cmpb.2003.10.004

21. Qui W, Chavarro J, Lazarus R, Rosner B, Ma J. powerSurvEpi: Power and Sample Size Calculation for Survival Analysis of Epidemiological Studies. Published online 2021. https://CRAN.R-project.org/package=powerSurvEpi

22. Latouche A, Porcher R, Chevret S. Sample size formula for proportional hazards modelling of competing risks. *Stat Med*. 2004;23(21):3263-3274. doi:10.1002/sim.1915

23. Autoimmune disease. National Library of Medicine MedGen. Accessed May 30, 2024. https://www.ncbi.nlm.nih.gov/medgen/2135

24. Dalal AA, Duh MS, Gozalo L, et al. Dose-Response Relationship Between Long-Term Systemic Corticosteroid Use and Related Complications in Patients with Severe Asthma. *J Manag Care Spec Pharm*. 2016;22(7). doi:10.18553/jmcp.2016.22.7.833

25. Karime C, Travers P, Ouni A, Francis D. Association Between Coronavirus Disease 2019 and Acute Complicated Diverticulitis. *Cureus*. 2023;15(1). doi:10.7759/cureus.33252

26. Akrami M, Ghaeini Hesarooeih A, Barfei M, Zangouri V, Alborzi Z. Clinical Characteristics of Bowel Obstruction in Southern Iran; Results of a Single Center Experience. *Bull Emerg Trauma*. 2015;3(1):22-26.

27. Meehan M, Shah A, Lobo J, et al. Validation of an algorithm to identify incident interstitial lung disease in patients with rheumatoid arthritis. *Arthritis Res Ther*. 2022;24:2. doi:10.1186/s13075-021-02655-z

28. Schiefenhövel F, Poncette AS, Boyle EM, et al. Pleural effusions are associated with adverse outcomes after cardiac surgery: a propensity-matched analysis. *J Cardiothorac Surg*. 2022;17:298. doi:10.1186/s13019-022-02050-y

29. Walling AM, Sudore RL, Bell D, et al. Population-Based Pragmatic Trial of Advance Care Planning in Primary Care in the University of California Health System. *J Palliat Med*. 2019;22(Suppl 1):72-81. doi:10.1089/jpm.2019.0142

12. Appendices

Appendix A: List of abbreviations

Abbreviation	Definition
ASD	Absolute standardized difference
ALK	Anaplastic lymphoma kinase
CARE	Coalition to Advance Real-World Evidence through Randomized Controlled Trial Emulation
CNS	Central nervous system
CONSORT	Consolidated Standards of Reporting Trials
DBC	Database connection
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic health record
EGFR	Epidermal growth factor receptor
FDA	Food and Drug Administration
HIV	Human Immunodeficiency Virus
IPTW	Inverse probability of treatment weighting
KRas	Kirsten rat sarcoma viral oncogene homolog
mAb	Monoclonal antibody
MK-3475	Pembrolizumab
NLP	Natural Language Processing
NSAIDs	Nonsteroidal anti-inflammatory drugs

Abbreviation	Definition
NSCLC	Non-small cell lung cancer
OMOP	Observational Medical Outcomes Partnership
OS	Overall survival
PD-L1	Programmed death-ligand 1
PD-L2	Programmed death-ligand 2
PFS	Progression-free survival
PS	Propensity score
QC	Quality control
RCT	Randomized controlled trials
RECIST	Response Evaluation Criteria in Solid Tumours
RWD	Real-world data
RWE	Real-world evidence
rwOS	Real-world overall survival
rwPFS	Real-world progression-free survival
U.S.	United States

Appendix B: Preliminary feasibility assessment

STEP 1a: Overarching research aim					
To emulate the KEYNOTE-189 randomized controlled trial of pembrolizumab + chemotherapy for the first-line treatment of metastatic nonsquamous non-small cell lung cancer without EGFR or ALK mutations using real-world data.					
STEP 1b: Trial research question					
Among adult patients with metastatic nonsquamous non-small cell lung cancer without EGFR or ALK mutations, does initial treatment with pembrolizumab + chemotherapy compared with treatment with chemotherapy alone result in longer overall and progression-free survival?					
STEP 1c: Trial primary objective(s)					
Among adult patients with metastatic nonsquamous non-small cell lung cancer without EGFR or ALK mutations, compare overall survival and progression-free survival for patients treated with pembrolizumab + chemotherapy and patients treated with chemotherapy alone.					
DESIGN ELEMENTS	**STEP 2: DESCRIBE ORIGINAL CLINICAL TRIAL**	**STEP 3: DESCRIBE REAL-WORLD DATA STUDY EMULATION OF ORIGINAL CLINICAL TRIAL**			
		3a. Minimal criteria for valid operationalization in real-world data source	**3b. Criteria ranking with regard to uniqueness and importance**	**Data Source**	
OVERALL RATING				**5**	
GENERAL					
Sample size	**Trial sample size**	**1.5x trial sample size[2]**	**Must Have**	Sample size among adult patients with metastatic non-small cell lung cancer who received first-line pembrolizumab + pemetrexed + (carbo/cis)platin or pemetrexed + (carbo/cis)platin	
Treated	410	615		2,605	**5**
Comparator	206	309		1,925	**5**

Length and frequency of follow-up[1]	Median reported follow-up: 10.5 months (range: 0.2 to 20.4 months)	Sufficient time coverage in dataset to identify outcome after receipt of treatment	**Must Have**	Earliest metastatic diagnosis date: Q3 2004 End of datacut: Q3 2023	**Not Applicable[3]**
VARIABLE-RELATED					
Variable	**Trial criterion**	**Minimal criteria for valid operationalization in any real-world data source based on routine clinical care**	**Criteria ranking with regard to uniqueness and importance**	**Operationalization and coverage in data source**	**Rating**
Treatment	200 mg intravenous (IV) pembrolizumab + four cycles of the investigator's choice of IV cisplatin (75 mg/m^2) or IV carboplatin (area under the concentration–time curve, 5 mg per milliliter per minute) + pemetrexed (500 mg/m^2), every 3 weeks, followed by pemetrexed (500 mg/m^2) every 3 weeks	Date of pembrolizumab, pemetrexed and carboplatin or cisplatin treatment	**Must Have**	Date of pembrolizumab, pemetrexed, and carboplatin or cisplatin treatment is available. Lines of therapy are available.	**5**
Comparator	200 mg IV saline placebo + four cycles of the investigator's choice of IV cisplatin (75 mg/m^2) or IV carboplatin (area under the concentration–time curve, 5 mg per milliliter per minute) + pemetrexed (500 mg/m^2), every 3 weeks, followed by pemetrexed (500 mg/m^2) every 3 weeks	Date of pemetrexed and carboplatin or cisplatin treatment	**Must Have**	Date of pemetrexed and carboplatin or cisplatin treatment is available. Lines of therapy are available.	**5**

Variable	Trial criterion	Minimal criteria for valid operationalization in any real-world data source based on routine clinical care	Criteria ranking with regard to uniqueness and importance	Operationalization and coverage in data source	Rating
Inclusion Criterion 1	18 years of age or older	Year of birth	**Must Have**	Year of birth is available.	**5**
Inclusion Criterion 2	Pathologically-confirmed metastatic nonsquamous non-small cell lung cancer	Diagnosis of non-small cell lung cancer with histological and/or pathological confirmation of subtype; date of metastatic diagnosis	**Must Have**	Non-small cell lung cancer diagnosis and histology and/or pathology are available. Date of metastatic diagnosis is available.	**5**
Inclusion Criterion 3	No sensitizing epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) mutations	Dates and result of biomarker tests	**Must Have**	Biomarker test results and dates are available.	**5**
Inclusion Criterion 4	Received no previous systemic therapy for metastatic disease	Names/types and dates of antineoplastic treatment; date of metastatic diagnosis	**Must Have**	Lines of therapy are available.	**5**
Inclusion Criterion 5	ECOG performance status score of 0 or 1	ECOG performance status result	**Must Have**	ECOG or Karnofsky performance status information is available.	**5**

Variable	Trial criterion	Minimal criteria for valid operationalization in any real-world data source based on routine clinical care	Criteria ranking with regard to uniqueness and importance	Operationalization and coverage in data source	Rating
Inclusion Criterion 6	Has at least one measurable lesion according to RECIST v1.1	RECIST is not used to assess progression or response in a real-world setting	**Not Applicable**	RECIST is not used to assess progression or response in a real-world setting. Progression will be assessed with available real-world information (see below).	**Not Applicable**
Inclusion Criterion 7	Provided a tumor sample for determination of programmed death-ligand 1 (PD-L1) status	Patient agreement to provide a tumor tissue sample is not captured outside of a clinical trial setting and therefore is not relevant to a real-world emulation	**Not Applicable**	This criterion will not be operationalized.	**Not Applicable**
Exclusion Criterion 1	Evidence of symptomatic central nervous system metastases	Dates and locations of distant metastases	**Nice to Have**	Date and site of metastases are available.	**5**
Exclusion Criterion 2	History of noninfectious pneumonitis that required the use of glucocorticoids	Date of noninfectious pneumonitis diagnosis; date of glucocorticoid treatment	**Nice to Have**	Date of noninfectious pneumonitis diagnosis is available, but completeness could not be ascertained. Date of glucocorticoids treatment is available.	**4**
Exclusion Criterion 3	Active autoimmune disease or systemic immunosuppressive treatment	Date of autoimmune disease diagnosis; dates of immunosuppressant treatment	**Nice to Have**	Date of autoimmune disease diagnosis is available, but completeness could not be ascertained. Date of immunosuppressive treatment is available.	**4**

Variable	Trial criterion	Minimal criteria for valid operationalization in any real-world data source based on routine clinical care	Criteria ranking with regard to uniqueness and importance	Operationalization and coverage in data source	Rating
Exclusion Criterion 4	Received >30 Gray of radiation therapy to the lung in the 6 months prior to the first dose of study medication	Date, location, and dose of radiation therapy	**Nice to Have**	Dates, doses, and sites of radiation therapy are available, but completeness could not be ascertained.	**4**
Primary Outcome 1 (Definition & Ascertainment)	Overall survival	Date of death; dates of healthcare interactions	**Must Have**	Date of death is available. Date of last activity is available.	**4**
Primary Outcome 2 (Definition & Ascertainment)	Progression-free survival	Date of death; curated progression variable; imaging results; dates of healthcare interactions	**Must Have**	Curated progression information is available. Date of death is available. Date of last activity is available.	**4**
Confounding Variable 1	Not applicable in a randomized setting	Age	**Must Have**	Year of birth is available.	**5**
Confounding Variable 2	Not applicable in a randomized setting	Sex	**Must Have**	Sex is available.	**5**
Confounding Variable 3	Not applicable in a randomized setting	Race/ethnicity	**Must Have**	Race/ethnicity is available.	**5**
Confounding Variable 4	Not applicable in a randomized setting	Performance status	**Must Have**	ECOG or Karnofsky performance status information is available.	**5**
Confounding Variable 5	Not applicable in a randomized setting	Smoking status	**Must Have**	Smoking status is available.	**5**
Confounding Variable 6	Not applicable in a randomized setting	Progression/disease free-interval (Time from initial diagnosis to metastatic diagnosis)	**Must Have**	Date of initial diagnosis is available. Date of metastatic diagnosis is available.	**5**

Variable	Trial criterion	Minimal criteria for valid operationalization in any real-world data source based on routine clinical care	Criteria ranking with regard to uniqueness and importance	Operationalization and coverage in data source	Rating
Confounding Variable 7	Not applicable in a randomized setting	Year of study treatment initiation	**Must Have**	Year of study treatment initiation is available.	**5**
Confounding Variable 8	Not applicable in a randomized setting	Number and/or location(s) of metastatic sites	**Must Have**	Metastatic site location is available. Number of metastatic sites can be determined.	**5**
Confounding Variable 9	Not applicable in a randomized setting	PD-L1 tumor proportion score status	**Must Have**	Biomarker test results and dates are available.	**5**

Abbreviations: ALK = anaplastic lymphoma kinase; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; IV = intravenous; PD-L1 = programmed death-ligand 1; Q1/Q2/Q3/Q4 = quarter of year (Q1: January - March, Q2: April - June, Q3: July - September, Q4: October - December); RECIST = Response Evaluation Criteria in Solid Tumors; TNM = Tumor, Node, Metastasis.

Footnotes:

1. Follow-up time is stated as reported in the trial publication. Maximum available observation time is reported for the real-world data source. These are not directly comparable.
2. The minimum sample size for feasibility analyses was selected to account for expected attrition when all eligibility criteria are applied.
3. The final study period would be defined in the study protocol based on the date of treatment approval and relevant updates to treatment guidelines.

Keys for Ranking	
Scoring for Data Sources by Data Elements	
Scoring	**Description**
1	Data Requirements are not met
2	
3	Some data requirements are met
4	

5	All or nearly all data requirements are met

Appendix C: Code lists

Design Element	Variable	Code Type	Code
Inclusion	Absolute neutrophil count (ANC)	Lab test name	Neutrophils [#/volume] in Blood Neutrophils [#/volume] in Blood by Manual count Neutrophils [#/volume] in Blood by Automated count
		Units	x10(3)/mcL
Exclusion	Active infection requiring therapy	ICD-10	A00 A01 A02 A03 A04 A06 A07 A08 A09 A15 A18 A20 A21 A22 A23 A24 A25 A26 A27 A28 A30 A31 A32 A33 A34 A35 A36 A37 A38

Design Element	Variable	Code Type	Code
			A39 A40 A41 A42 A43 A44 A46 A48 A49 A50 A51 A52 A53 A54 A55 A56 A57 A58 A59 A60 A63 A64 A65 A66 A67 A68 A69 A70 A71 A74 A75 A77 A78 A79 A80 A81

Design Element	Variable	Code Type	Code
			A82 A83 A84 A85 A86 A87 A88 A89 A90 A91 A92 A93 A94 A95 A96 A98 A99 B00 B01 B02 B03 B04 B05 B06 B07 B08 B09 B10 B15 B16 B17 B18 B19 B25 B26 B27

Design Element	Variable	Code Type	Code
			B30 B33 B34 B35 B36 B37 B38 B39 B40 B41 B42 B43 B44 B45 B46 B47 B48 B49 B50 B51 B52 B53 B54 B55 B56 B57 B58 B59 B60 B64 B65 B66 B67 B68 B69 B70

Design Element	Variable	Code Type	Code
			B71 B72 B73 B74 B75 B76 B77 B78 B79 B80 B81 B82 B83 B85 B86 B87 B88 B89 B95 B96 B97 B99 D59 G00 G01 G02 G04 G14 H05 I00 I02 J00 J02 J03 J04 J05

Design Element	Variable	Code Type	Code
			J06 J12 J13 J14 J15 J16 J20 J21 J39 K65 L00 L01 L02 L03 L04 L05 L08 L30 M00 N10 N11 N12 N13 N15 N30 N34 N39 O23 O26 O41 O86 O98 P00 P23 P35 P36

Design Element	Variable	Code Type	Code
			P37 P38 P39
		Generic name	Abacavir Acyclovir Amantadine Atazanavir Baloxavir Marboxil Bictegravir Brivudin Cobicistat Daclatasvir Darunavir Delavirdine Didanosine Dolutegravir Doravirine Efavirenz Elvitegravir Emtricitabine Enfuvirtide Entecavir Etravirine Famciclovir Fomivirsen Fosamprenavir Foscarnet Fostemsavir Ganciclovir Ibalizumab Indinavir Lamivudine Ledipasvir Lenacapavir Letermovir

Design Element	Variable	Code Type	Code
			Lopinavir Maraviroc Maribavir Molnupiravir Nelfinavir Nevirapine Nevirapine Nirmatrelvir/Ritonavir Oseltamivir Pemivibart Penciclovir Peramivir Raltegravir Remdesivir Rilpivirine Rimantadine Ritonavir Saquinavir Simeprevir Sofosbuvir Stavudine Tecovirimat Tenofovir Tenofovir Tipranavir Valaciclovir Valacyclovir Valganciclovir Vilobelimab Zanamivir Zidovudine Penicillin Tobramycin Cycloserine Vancomycin Capreomycin

Design Element	Variable	Code Type	Code
			Cefazolin Streptomycin Neomycin Bacitracin Sulfadiazine Isoniazid Cefaclor Loracarbef Cefuroxime Cephalothin Cefamandole Ceftazidime Drotrecogin Alfa Amoxicillin Cephradine Ampicillin Sulfamethoxazole Erythromycin Dicloxacillin Cloxacillin Metronidazole Tetracycline Doxycycline Cephalexin Nitrofurantoin Aztreonam Oxacillin Nafcillin Cefepime Sulfisoxazole Trimethoprim Sulfisoxazole Acetyl Ceftriaxone Erythromycin Ethylsuccinate Cefixime Sulfasalazine

Design Element	Variable	Code Type	Code
			Ethambutol Pyrazinamide Minocycline Demeclocycline Norfloxacin Cefoxitin Imipenem Ertapenem Ceftizoxime Cefonicid Ethionamide Tigecycline Clindamycin Lincomycin Spectinomycin Gentamicin Cefpodoxime Linezolid Rifabutin Gatifloxacin Amikacin Kanamycin Cefadroxil Cephapirin Sulfanilamide Sulfacetamide Chloramphenicol Dapsone Polymyxin Nalidixic Acid Mafenide Lomefloxacin Mezlocillin Ciprofloxacin Ciprofloxacin Lactate Moxifloxacin

Design Element	Variable	Code Type	Code
			Clofazimine Mupirocin Ticarcillin Cefotetan Cefotaxime Silver Sulfadiazine Levofloxacin Oxytetracycline Cefoperazone Trovafloxacin Mesylate Clarithromycin Tinidazole Ofloxacin Doripenem Sulfathiazole Rifampin Chloramphenicol Sod Succinate Meropenem Azithromycin Cefdinir Paromomycin Colistin Enoxacin Sparfloxacin Quinupristin Netilmicin Ceftibuten Cefprozil Rifapentine Telithromycin Daptomycin Furazolidone Grepafloxacin Piperacillin Cefditoren Ceftaroline Fosamil

Design Element	Variable	Code Type	Code
			Fosfomycin Tromethamine Telavancin Polymyxin B Sulfate Gemifloxacin Rifaximin Sarecycline Bacitracin Chloramphenicol Palmitate Chlortetracycline Trimethoprim Aminosalicylic Pretomanid Bacitracin Zinc Fidaxomicin Rifamycin Dalbavancin Cefiderocol Bedaquiline Oritavancin Ceftolozane Tedizolid Plazomicin Ozenoxacin Delafloxacin Meglumine Triamcinolone Omadacycline Eravacycline Lefamulin Ivermectin
Inclusion	ALT (SGPT)	Lab test name	Alanine aminotransferase [Enzymatic activity/volume] in Blood Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma by No addition of P-5'-P

Design Element	Variable	Code Type	Code
			Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma by With P-5'-P
		Units	U/L IU/L
Inclusion	Anti-PD-1, PD-L1, PD-L2 agent or an antibody targeting other immuno-regulatory receptors or mechanisms	Generic name	Pembrolizumab Nivolumab Atezolizumab Avelumab Durvalumab Ipilimumab Relatlimab
Exclusion	Antineoplastic biological therapy	Generic Name	Pembrolizumab Ipilimumab Avelumab Durvalumab Bevacizumab Nivolumab Atezolizumab Ramucirumab Amivantamab Cetuximab Relatlimab Panitumumab Enfortumab Crizotinib Erlotinib Dacomitinib Lorlatinib Binimetinib Sotorasib Trametinib Cobimetinib Abemaciclib Trastuzumab

Design Element	Variable	Code Type	Code
			Axitinib Gefitinib Alectinib Mobocertinib Dabrafenib Osimertinib Imatinib Ado-Trastuzumab Emtansine Olaparib Pertuzumab Sorafenib Cabozantinib Sunitinib Afatinib Entrectinib Capmatinib Lapatinib Lenvatinib Palbociclib Adagrasib Brigatinib Selpercatinib Vemurafenib Neratinib Ceritinib Fam-Trastuzumab Deruxtecan-Nxki Regorafenib Pralsetinib Rucaparib Niraparib Tepotinib Ivosidenib Encorafenib Vismodegib Pazopanib Sacituzumab

Design Element	Variable	Code Type	Code
			Thalidomide
Inclusion	AST (SGOT)	Lab test name	Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma by With P-5'-P
		Units	U/L IU/L
Exclusion	Autoimmune disease[23]	SNOMED	85828009
Exclusion	Chronic systemic steroids (Corticosteroids/Glucocorticoids) [24]	Generic name	Methylprednisolone Triamcinolone Prednisone Dexamethasone Budesonide Cortisone acetate Hydrocortisone Prednisolone Betamethasone
Exclusion	Clinically active diverticulitis[25]	ICD-10	K57
Inclusion	Creatinine clearance	Lab test name	Creatinine [Mass/volume] in Blood Creatinine [Mass/volume] in Serum or Plasma Creatinine [Mass/volume] in Urine
		Units	mg/dL
		Lab test name	Creatinine renal clearance in Urine and Serum or Plasma collected for unspecified duration Creatinine renal clearance predicted by Cockcroft-Gault formula
		Units	mL/min mL/min/SA
Inclusion	Serum Bilirubin	Lab test name	Bilirubin.total [Mass/volume] in Serum or Plasma

Design Element	Variable	Code Type	Code
			Bilirubin.direct [Mass/volume] in Serum or Plasma Bilirubin.conjugated [Mass/volume] in Serum or Plasma Bilirubin.total [Mass/volume] in Blood
		Units	mg/dL
Exclusion	GI obstruction[26]	ICD-10	K56.6
Inclusion	Hemoglobin	Lab test name	Hemoglobin [Mass/volume] in Blood Hemoglobin [Mass/volume] in Arterial blood Hemoglobin [Mass/volume] in Venous blood Hemoglobin [Mass/volume] in Blood by calculation
		Units	g/dL mg/dL
Exclusion	Hepatitis B	ICD-10	B16 B19.1
Exclusion	Hepatitis C	ICD-10	B17.1 B19.2
Exclusion	HIV	SNOMED	445945000 91947003
Exclusion	Immunosuppressant or DMARD	Generic name	Hyaluronidase / rituximab Mycophenolate mofetil Rituximab Sirolimus Prednisone Cyclosporine Tacrolimus Azathioprine Everolimus Alemtuzumab Methotrexate Leflunomide Sulfasalazine

Design Element	Variable	Code Type	Code
			Hydroxychloroquine Etanercept Infliximab Adalimumab Certolizumab pegol Golimumab Tocilizumab Abatacept Tofacitinib Baricitinib Upadacitinib
Inclusion	International normalized ratio (INR)	Lab test name	International normalized ratio INR in Platelet poor plasma by Coagulation assay INR in Platelet poor plasma or blood by Coagulation assay INR in Blood by Coagulation assay
		Units	Ratio International normalized ratio
Exclusion	Interstitial lung disease[27]	ICD-10	J84.10 J84.112 J84.9 M05.10
Exclusion	Intra-abdominal abscess	ICD-10	L02.211 K65.1
Inclusion	Major surgery	Surgery Name	Lobectomy of lung Wedge resection Thoracoscopic wedge resection of lung Thoracoscopic lobectomy of lung Excision Total pneumonectomy Thoracoscopic pneumonectomy Bilobectomy of lung Thoracoscopy, surgical; with diagnostic wedge resection followed by

Design Element	Variable	Code Type	Code
			anatomic lung resection (List separately in addition to code for primary procedure) Laparoscopy, surgical; with removal of adnexal structures (partial or total oophorectomy and/or salpingectomy) Excision, tumor, soft tissue of abdominal wall, subcutaneous; 3 cm or greater Salpingo-oophorectomy, complete or partial, unilateral or bilateral (separate procedure) Excision of segment of lung Thoracotomy; with diagnostic wedge resection followed by anatomic lung resection (List separately in addition to code for primary procedure) Laparoscopy, surgical; with lysis of adhesions (salpingolysis, ovariolysis) (separate procedure) Thoracoscopy, surgical; with removal of a single lung segment (segmentectomy) Thoracoscopy; with biopsy(ies) of pleura Completion pneumonectomy
Inclusion	Partial Thromboplastin Time (includes activated)	Lab test name	aPTT actual/normal in Platelet poor plasma by Coagulation assay aPTT in Blood by Coagulation 1:1 saline aPTT in Blood by Coagulation assay aPTT in Control Platelet poor plasma by Coagulation assay aPTT in Platelet poor plasma by Coagulation 1:1 saline aPTT in Platelet poor plasma by Coagulation assay aPTT normal/actual in Platelet poor plasma by Coagulation assay aPTT panel - Platelet poor plasma
		Units	Seconds
Inclusion	Platelets	Lab test name	Platelets [#/volume] in Blood Platelets [#/volume] in Blood by Automated count Platelets [#/volume] in Blood by Estimate Platelets [#/volume] in Blood by Manual count
		Units	10*9 cells/L x10(3)/mcL

Design Element	Variable	Code Type	Code
Exclusion	Pleural effusion[28]	ICD-10	J90 J91
Inclusion	Prothrombin Time (PT)	Lab test name	Prothrombin time Prothrombin time (PT) Prothrombin time (PT) in Control Platelet poor plasma by Coagulation assay Prothrombin time (PT) PIVKA sensitive in Platelet poor plasma Prothrombin time (PT) in Blood by Coagulation assay
		Units	Seconds
Inclusion	Radiation to the lung	Treatment Site	Chest and Lung Chest wall structure Left lung structure Lung structure Mediastinal structure Right lung structure Structure of hilum of right lung Structure of lower lobe of left lung Structure of lower lobe of lung Structure of lower lobe of right lung Structure of lung and/or mediastinum Structure of right half of chest wall Structure of upper lobe of left lung Structure of upper lobe of lung Thoracic structure Pleura Both lungs Lung and pleura, CS Middle lung lobe bronchus Structure of apex of right lung Structure of hilum of left lung Structure of hilum of lung Structure of middle lobe of right lung Structure of upper lobe of right lung Main bronchus structure

Design Element	Variable	Code Type	Code
			Mass of thoracic structure Left thorax structure
		Units	cGY
Inclusion / Exclusion	Small cell elements	Histology Description	Small cell carcinoma, metastatic site Small cell carcinoma, NOS Combined small cell carcinoma Malignant tumor, small cell type Squamous cell carcinoma, small cell, nonkeratinizing
Inclusion / Exclusion	Squamous histology	Histology Description	Squamous cell carcinoma, NOS Squamous cell carcinoma in situ, NOS Squamous cell carcinoma, sarcomatoid Adenosquamous carcinoma Basaloid squamous cell carcinoma Squamous cell carcinoma, large cell, keratinizing Mixed adenocarcinoma and squamous cell carcinoma Squamous cell carcinoma, nonkeratinizing, NOS Squamous cell carcinoma, large cell, nonkeratinizing, NOS Papillary squamous cell carcinoma Squamous cell carcinoma, metastatic, NOS Squamous cell carcinoma, keratinizing, NOS Squamous cell carcinoma, spindle cell Squamous cell papilloma, NOS Squamous carcinoma Adenocarcinoma with squamous metaplasia
Exclusion	Ascites[29]	ICD-10	R18.8
Inclusion	Thyroid stimulating hormone	Lab test name	Thyrotropin [Units/volume] in Serum or Plasma Thyrotropin [Units/volume] in Blood
		Units	uIU/mL mIU/L mIU/mL mU/L

Design Element	Variable	Code Type	Code
			uU/mL IU/mL
		Lab test name	Thyroxine (T4) free [Mass/volume] in Serum or Plasma Thyroxine (T4) [Mass/volume] in Serum or Plasma Thyroxine (T4) free index in Serum or Plasma by calculation
		Units	pg/mL ng/dL ng/mL pg/dL
		Lab test name	Triiodothyronine (T3) Free [Mass/volume] in Serum or Plasma Triiodothyronine (T3) [Mass/volume] in Serum or Plasma Triiodothyronine (T3) Free [Moles/volume] in Serum or Plasma
		Units	ng/dL ug/dL ug/mL ng/mL