PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

POST-AUTHORIZATION EFFECTIVENESS STUDY (PAES) INFORMATION

Title	Real-world post-authorization effectiveness study (PAES) of pembrolizumab for the treatment of NSCLC across races, ethnicities, and age groups
Protocol Version identifier	3475-G18-00-v3
Date of last version of protocol	28-MAR-2025
EU PAS Register No:	Study not registered
Active substance	L01XC18, monoclonal antibody, pembrolizumab
Medicinal product(s):	Keytruda®, pembrolizumab
Joint PASS	No
Research question and objectives	Primary Study Objectives: 1. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **early-stage NSCLC** by estimating real-world overall survival (rwOS) among the following subgroups: • Race (e.g., Black or African American, Asian, other [American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and other race] and White patients) • Hispanic/Latino ethnicity (Hispanic/Latino patients and not Hispanic/Latino patients) • Age (elderly patients ≥75 years and younger patients <75 years) 2. To evaluate the real-world effectiveness of **pembrolizumab in combination with platinum-based chemotherapy (PBC)** in US patients with **early-stage NSCLC** by estimating rwOS among the above subgroups 3. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups 4. To evaluate the real-world effectiveness of **pembrolizumab in combination with PBC** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

	Exploratory study objectives: 1. To evaluate the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 1) and pembrolizumab in combination with PBC (see Primary Objective 2) in US patients with **early-stage NSCLC** by estimating **real-world disease-free survival (rwDFS) and real-world event-free survival (rwEFS), respectively,** among the above subgroups 2. To apply **propensity score methods** to the evaluation of the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 3) and pembrolizumab in combination with PBC (see Primary Objective 4) in US patients with **advanced NSCLC** among the above subgroups
Country of study	United States of America (US)
Author	PPD Principal Scientist, Epidemiology Biostatistics and Research Decision Sciences (BARDS)
Marketing authorisation holder including MAH Contact Person	Merck Sharp & Dohme, LLC A subsidiary of Merck & Co., Inc. 126 East Lincoln Ave. P.O. Box 2000 Rahway, NJ 07065 PPD Principal Scientist, Epidemiology Biostatistics and Research Decision Sciences (BARDS) PPD PPD
Merck Final Repository (REDS) Date	03-DEC-2025
Date of Health Authority Approval of Protocol	08-DEC-2025

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

TABLE OF CONTENTS

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

LIST OF TABLES

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

LIST OF FIGURES

N/A

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

LIST OF ABBREVIATIONS

AE	Adverse event
ALK	Anaplastic lymphoma kinase
aNSCLC	Advanced non-small cell lung cancer
BRAF	B-Raf Proto-Oncogene Serine/Threonine Kinase
CI	Confidence interval
DSUR	Development Safety Update Reports
HER2	Human Epidermal Growth Factor Receptor 2
EC	Ethics Committee
ECOG	Eastern Cooperative Oncology Group
EHR	Electronic health record
EGFR	Epidermal growth factor receptor
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
eNSCLC	Early non-small cell lung cancer
FDA	Food and Drug Administration
FDAAA	Food and Drug Administration Amendments Act
FDAMA	Food and Drug Administration Modernization Act
FHRD	Flatiron Health Research Database
FHRN	Flatiron Health Research Network
GPP	Good Pharmacoepidemiology Practice
GQPPV	Global Qualified Person for Pharmacovigilance
GVP	Good Pharmacovigilance Practice
HR	Hazard ratio
HMA-EMA	Head of Medicines Agencies – European Medicines Agency
ICD	International Classification of Disease
IO	Immuno-oncology
IPW	Inverse Probability Weighting
IRB	Institutional Review Board
KRAS	Kirsten Rat Sarcoma Virus
MET	Mesenchymal-Epithelial Transition
MSD	Merck Sharp & Dohme LLC
NCCN	National Comprehensive Cancer Network
NPCR	National Program of Cancer Registries
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic Tyrosine Receptor Kinase
OS	Overall survival

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

PAES	Post-authorization effectiveness study
PASS	Post-authorization safety study
PBRER	Periodic Benefit Risk Evaluation Report
PBC	Platinum-based chemotherapy
PD-L1	Programmed death-ligand 1
PMC	Post-marketing commitment
PQC	Product Quality Complaint
Protocol CI	Protocol Coordinating Investigator
PRS	Program Requirements and Specification
PSUR	Periodic Safety Update Report
QA	Quality assurance
QC	Quality control
RET	Rearranged During Transfection
RCT	Randomized Clinical Trial
ROS1	ROS Proto-Oncogene 1, Receptor Tyrosine Kinase
RWD	Real World Data
SEER	Surveillance Epidemiology and End Results
SMD	Standardized Mean Differences
SQI	Significant quality issue
SSDI	Social Security Death Index
TPS	Tumor Proportion Score
US	United States of America

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

1 RESPONSIBLE PARTIES

Principal investigator	PPD Principal Scientist, Epidemiology Biostatistics and Research Decision Sciences (BARDS) PPD PPD
Coordinating investigator for each country in which the study is to be performed	Not applicable
Sponsor contacts	Merck Sharp & Dohme, LLC A subsidiary of Merck & Co., Inc. 126 East Lincoln Ave. P.O. Box 2000 Rahway, NJ 07065 Contact: PPD Principal Scientist, Epidemiology Biostatistics and Research Decision Sciences (BARDS) PPD PPD
Other contacts	PPD Senior Director, Epidemiology Biostatistics and Research Decision Sciences (BARDS) PPD PPD
Supplier/Collaborator	Flatiron Health, Inc. 233 Spring St, New York, NY 10013
Investigators	PPD Senior Research Scientist, RWE PPD

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

2 ABSTRACT

Title	Real-world post-authorization effectiveness study (PAES) of pembrolizumab for the treatment of NSCLC across races, ethnicities, and age groups
Protocol Number / Version	3475-G18-00-v3
Date	24-NOV-2025
Author	PPD Principal Scientist, Epidemiology Biostatistics and Research Decision Sciences (BARDS)
Rationale & Background	Limited data exist on the efficacy of immune checkpoint inhibitors across diverse patient populations, largely due to the underrepresentation of racial/ethnic minorities and elderly patients in oncology clinical trials. The Sponsor is proposing to conduct a real-world post-authorization effectiveness study (PAES) to characterize the real-world effectiveness (overall survival) of pembrolizumab for the treatment of NSCLC among US patients across races, ethnicities, and age groups to address part of an FDA post-marketing commitment (PMC) for KEYNOTE-671 (4531-2).

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Research Question(s) & Objective(s)	Primary Study Objectives: 1. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **early-stage NSCLC** by estimating real-world overall survival (rwOS) among the following subgroups: • Race (e.g., Black or African American, Asian, other [American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and other race] and White patients) • Hispanic/Latino ethnicity (Hispanic/Latino patients and not Hispanic/Latino patients) • Age (elderly patients ≥75 years and younger patients <75 years) 2. To evaluate the real-world effectiveness of **pembrolizumab in combination with platinum-based chemotherapy (PBC)** in US patients with **early-stage NSCLC** by estimating rwOS among the above subgroups 3. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups 4. To evaluate the real-world effectiveness of **pembrolizumab in combination with PBC** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups Exploratory study objectives: 1. To evaluate the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 1) and pembrolizumab in combination with PBC (see Primary Objective 2) in US patients with **early-stage NSCLC** by estimating **real-world disease-free survival (rwDFS) and real-world event-free survival (rwEFS), respectively,** among the above subgroups

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

	2. To apply **propensity score methods** to the evaluation of the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 3) and pembrolizumab in combination with PBC (see Primary Objective 4) in US patients with **advanced NSCLC** among the above subgroups
Study Design	This retrospective cohort study will be conducted using Flatiron Health US-based deidentified electronic health record (EHR)-derived databases.
Population	US adults (≥18 years of age) diagnosed with NSCLC and treated with ≥1 cycle of pembrolizumab (monotherapy or in combination with platinum-based chemotherapy)
Variables	Exposure: number of pembrolizumab cycles, pembrolizumab treatment type (i.e., monotherapy or combined with platinum-based chemotherapy), pembrolizumab duration, year of first pembrolizumab administration, chemotherapy name, and chemotherapy duration, date of surgery, and radiation therapy type, if applicable and as available Outcome: vital status, date of death, date of last follow-up, recurrence, date of first local recurrence, date of first distant recurrence, progression, date of progression Covariates: race, ethnicity, and age Other patient characteristics for descriptive purposes: sex, ECOG performance status, smoking status, disease stage, histology, PD-L1 status, genomic alteration status as available, treatment setting (e.g., adjuvant), and chemotherapy agents, as available
Data Sources	Flatiron Health: early non-small cell lung cancer (eNSCLC) Flatiron Health Research Database (FHRD) and advanced non-small cell lung cancer (aNSCLC) FHRD
Study Size	The study will include all patients in the research database who meet study inclusion and exclusion criteria within the data cutoff period.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Data Analysis	Four study cohorts will be evaluated based on disease stage and treatment type: 1) early-stage patients treated with pembrolizumab monotherapy, 2) early-stage patients treated with pembrolizumab in combination with PBC, 3) advanced-stage patients treated with pembrolizumab monotherapy, and 4) advanced-stage patients treated with pembrolizumab in combination with PBC. Within each study population, real-world overall survival will be estimated using Kaplan-Meier methods for each of the following subgroups: race (e.g., Black or African American, Asian, other race, White), ethnicity (Hispanic/Latino, not Hispanic/Latino), and age group (elderly, younger). Median survival times and the probabilities of survival at key milestones of interest (e.g., 1-year, 5-year) and corresponding 95% confidence intervals will be calculated. Exploratory analyses of rwDFS and rwEFS will be conducted in the early-stage population (study cohorts 1 and 2, respectively). rwDFS and rwEFS will each be estimated using Kaplan-Meier methods for race, ethnicity, and age groups. An exploratory analysis of rwOS using propensity score weighting methods will be conducted in the advanced stage population to balance baseline characteristics that may confound the effect of each exposure subgroup (race, ethnicity, or age group) on survival following pembrolizumab treatment.
Milestones	
Start of data collection:	1Q2026
End of data collection:	3Q2027
Interim report(s) of study results:	Not applicable
Study progress report(s):	Not applicable
Final report of study results:	4Q2028

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

3 AMENDMENTS AND UPDATES

Amendment or Update no	Date	Section of Study Protocol	Amendment or Update	Reason	NIR DRC Approval Date	NIR DRC Version No
<1>	< 24-OCT-2024>	N/A	Original	N/A	14-OCT-2024	1
<2>	<14-APR-2025>	List of abbreviations	Update 1	Added additional abbreviations	28-MAR-2025	2
		1 Responsible Parties		Updated Flatiron Health as Supplier		
		2 Abstract		Updated protocol version, study design and data sources to reference Flatiron		
		5 Rationale and Background		Added FDA approval dates for each indication		
		6 Research Question and Objectives		Updated Table 1 to refine exclusion criteria		

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Amendment or Update no	Date	Section of Study Protocol	Amendment or Update	Reason	NIR DRC Approval Date	NIR DRC Version No
		7 Research Methods		7.1 Updated study design details regarding feasibility assessment 7.2 Added detail to inclusion and exclusion criteria 7.3.2 Added detail to outcomes and added references 7.3.3 Added detail to covariates and added references 7.4 Updated data source and status of feasibility assessment Phases 1 and 2 7.5 Updated Table 2 to reflect subpopulation >=75 years 7.9 Added quality control details to be conducted by supplier; Edits approved by Merck Legal/OGC 7.10 Added detail to limitations and added references		
		8 Protection of Human Subjects		Updated to include Flatiron involvement in IRB		

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Amendment or Update no	Date	Section of Study Protocol	Amendment or Update	Reason	NIR DRC Approval Date	NIR DRC Version No
		Annex 2		Updated to reflect Flatiron involvement; Edits approved by Merck Legal/OGC		
<3>	<03-DEC-2025>	List of abbreviations	Revision	Added additional abbreviations	24-NOV-2025	3
		1 Responsible Parties		Updated to include new Flatiron Investigator		
		2 Abstract		Revised to reflect protocol changes, including updates to research objectives, variables, data analysis, and milestones.		
		6 Research Question and Objectives, 7.3.3 Covariates, 7.8 Data Analysis		Racial groups updated to include four separate subgroups (e.g., Black or African American, Asian, other [American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and other race] and White) when sample size allows.		

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Amendment or Update no	Date	Section of Study Protocol	Amendment or Update	Reason	NIR DRC Approval Date	NIR DRC Version No
		6 Research Question and Objectives, 7.2 Setting, 7.3.2 Outcomes, 7.8 Data Analysis		Per FDA recommendation separated out early-stage NSCLC and advanced NSCLC patient cohorts in analysis and endpoints clarified		
		7.3.3 Covariates		Added actionable biomarkers to the description of patient characteristics.		
		7.5 Study size		Added Table 2, preliminary estimates of study sample size by stage group. Updated Table 3 to include precision estimates for a sample size of n=1,000.		
		6 Research Question and Objectives, 7.8 Data Analysis		Incorporated propensity score methods as an exploratory objective for comparable analyses among advanced NSCLC		
		6 Research Question and Objectives, 7.3.2 Outcomes, 7.8 Data Analysis		Added an exploratory objective in early-stage NSCLC estimating real-world event-free survival (rwEFS) and real-world disease-free survival (rwDFS)		

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Amendment or Update no	Date	Section of Study Protocol	Amendment or Update	Reason	NIR DRC Approval Date	NIR DRC Version No
		7.10 Limitations of the Research Methods		Added limitation for propensity score weighting and noted that rwEFS and rwDFS will be evaluated in early-stage patients to address data maturity limitations. Added limitation related to eligibility criteria and comparability to clinical trial data.		

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

4 MILESTONES

Milestone	Planned Date
Registration in the HMA-EMA Catalogue of RWD Studies	4Q2025
Start of data collection	1Q2026
End of data collection	3Q2027
Final report of study results	4Q2028

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

5 RATIONALE AND BACKGROUND

In the US, lung cancer is the second most frequently diagnosed cancer (11.7% of new cancer cases in 2024) [National Cancer Institute 2024] and the leading cause of cancer death in both men and women, causing nearly as many deaths as breast, prostate, and colorectal cancers combined [Siegel, R. L., et al 2024]. Lung cancer is most frequently diagnosed among individuals over 65 years of age [National Cancer Institute 2023]. Although rates of incidence and mortality are higher among men than women [Sung, H., et al 2021], the incidence of lung cancer in women has more than doubled since the mid-1970s [Barta, J. A., et al 2019]. Tobacco exposure is the primary risk factor for developing lung cancer [Barta, J. A., et al 2019], with 81% of lung cancer deaths attributable to cigarette smoking [Siegel, R. L., et al 2024]. Non-small cell lung cancer (NSCLC) represents the majority (~89%) of all lung cancer cases [Zhang, Y., et al 2023]. Among US adults aged 18 years and older, the overall age-adjusted incidence of NSCLC was 49.7 per 100,000 in 2021 [National Cancer Institute (NCI) 2023], corresponding to an estimated 169,834 new cases of NSCLC in 2024 (US Census Bureau, International Programs, International Data Base Version: 12.03.21: International Database [census.gov]). The overall 5-year relative survival rate for patients with NSCLC is 28% [American Cancer Society 2023]; however, survival varies by stage at diagnosis, with early-stage lung cancer having a higher survival rate compared to late-stage diagnosis. Approximately 44% of patients with NSCLC are diagnosed with metastatic disease [Ganti, A. K., et al 2021] and the 5-year relative survival rate for those with distant metastases is only 9% [American Cancer Society 2023]. Lung cancer exhibits notable disparities across various sociodemographic groups, highlighting the need for a multifaceted approach to lung cancer diagnosis and treatment.

Age-related Disparities in Lung Cancer

Older and elderly adults (individuals over the ages of 65 and 75, respectively) represent a substantial proportion of lung cancer cases in the US and present unique challenges. Age-related disparities in lung cancer are a significant concern, with older individuals being at higher risk of lung cancer diagnosis and mortality [National Cancer Institute 2023] due to longer exposure to risk factors such as smoking and environmental carcinogens, as well as natural age-related changes in DNA damage repair. Older individuals are more likely to experience delays in diagnosis [Mills, S., et al 2023] and other disparities related to healthcare access, including financial barriers. Elderly patients with lung cancer in the US are less likely to receive guideline-concordant treatment and more likely receive less intensive treatment than recommended [Blom, E. F., et al 2020], although it is unclear whether this undertreatment reflects a greater toxicity risk, patient preferences, or poor adherence. Underlying cardiovascular and pulmonary comorbidities which increase in prevalence with advancing age are associated with poor lung cancer prognosis [Medrek, S. and Szmit, S. 2022] [Sogaard, M., et al 2013].

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Racial and Ethnic Disparities in Lung Cancer

Notable racial and ethnic disparities in NSCLC persist within the US and are driven by complex interactions between structural inequities, social determinants of health, as well as biological factors attributed to ancestry [American Association for Cancer Research 2024].

Black or African American: Accumulating evidence has shown that Black men are 15% more likely to develop lung cancer compared with White men [American Cancer Society 2022] and are at a higher risk at a younger age, even if they smoke less [Aldrich, M. C., et al 2019]. Mortality rates also remain higher in Black men compared to White men, although the absolute disparity in mortality has narrowed substantially between 2000 and 2020 [American Association for Cancer Research 2024].

American Indian or Alaskan Native: AI/AN people living in counties which include all or part of a reservation or share a common boundary with a reservation (ie, Purchased/Referred Care Delivery Areas), have higher lung cancer incidence rates than White Americans [Kratzer, T. B., et al 2022]. However, these patterns vary considerably by geographic region with incidence rates 80% higher among AI/AN people compared with White people in the Northern Plains and 64% lower than White people in the Southwest [Siegel, R. L., et al 2022].

Asian or Pacific Islander: Despite Asian/Pacific Islander people having overall lower incidence rates of NSCLC (any stage) compared with White people in the US, risk varies by Asian subgroup and nativity [Raz, D. J., et al 2008]. For example, never-smoking Japanese American women have a lower risk of NSCLC compared with never-smoking non-Hispanic White women; while the risk of NSCLC among never-smoking Native Hawaiian, Chinese American, and Filipinx American women is 1.7 to 2.3-fold greater than that in never-smoking non-Hispanic White women [DeRouen, M. C., et al 2022]. However, these patterns vary considerably by geographic region with rates 80% higher among AI/AN people compared with White people in the Northern Plains and 64% lower than White people in the Southwest [Siegel, R. L., et al 2022].

Hispanic or Latino: Compared with the non-Hispanic White population in the US, both lung cancer incidence and mortality rates are 53% lower among the Hispanic population (rate ratio=0.47) [American Association for Cancer Research 2022], demonstrating a known survival benefit among Hispanic adults [Miao, E., et al 2023].

Underrepresentation of Racial/Ethnic Minority and Elderly Patients in Clinical Trials

Racial/ethnic minority and elderly patients have historically been underrepresented in oncology clinical trials relative to the incidence of lung cancer [Aldrighetti, C. M., et al 2021] [Kilic, S., et al 2024]. The lack of representation of diverse patients in clinical trials hinders the generalizability of trial results, leading to limited understanding of the treatment efficacy and safety in racial/ethnic minority and older populations. Several factors contribute to the low participation of racial/ethnic minorities in clinical trials, including lack of

awareness of trial opportunities, cultural and language differences, medical mistrust, socioeconomic disparities, and other systemic barriers [Hamel, L. M., et al 2016]. Elderly patients tend to be under-recruited due to concerns about their ability to tolerate and respond to treatments, particularly in those with low performance status and those with comorbidities that could complicate trial outcomes [Sedrak, M. S., et al 2021]. Additionally, strict eligibility criteria, such as the exclusion of individuals with coexisting health conditions or those taking multiple medications, further limit the participation of older individuals who are more likely to have complex medical profiles.

KEYNOTE-671 Post-marketing Commitment

Multiple clinical trials investigating the safety and efficacy of pembrolizumab (MK-3475) in NSCLC have been conducted globally by the Sponsor, and resulted in the following current indications for NSCLC:

- Pembrolizumab, in combination with pemetrexed and platinum chemotherapy, is indicated for the first-line treatment of patients with metastatic nonsquamous non–small cell lung cancer (NSCLC), with no epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) genomic tumor aberrations [08/2018].

- Pembrolizumab, in combination with carboplatin and either paclitaxel or paclitaxel protein-bound, is indicated for the first-line treatment of patients with metastatic squamous non–small cell lung cancer (NSCLC) [10/2018].

- Pembrolizumab, as a single agent, is indicated for the first-line treatment of patients with non–small cell lung cancer (NSCLC) expressing programmed death ligand 1 (PD-L1) [tumor proportion score (TPS) ≥1%] as determined by an FDA-approved test, with no epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) genomic tumor aberrations, and is stage III where patients are not candidates for surgical resection or definitive chemoradiation, or metastatic [04/2019].

- Pembrolizumab, as a single agent, is indicated for the treatment of patients with metastatic non–small cell lung cancer (NSCLC) whose tumors express programmed death ligand 1 (PD-L1) [tumor proportion score (TPS) ≥1%] as determined by an FDA-approved test, with disease progression on or after platinum-containing chemotherapy. Patients with epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) genomic tumor aberrations should have disease progression on FDA-approved therapy for these aberrations prior to receiving pembrolizumab [10/2015].

- Pembrolizumab, as a single agent, is indicated for adjuvant treatment following resection and platinum-based chemotherapy (PBC) for adult patients with stage IB (T2a ≥4 cm), II, or IIIA non–small cell lung cancer (NSCLC) [01/2023].

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

- Pembrolizumab is indicated for the treatment of patients with resectable (tumors ≥4 cm or node positive) non–small cell lung cancer (NSCLC) in combination with platinum-containing chemotherapy as neoadjuvant treatment, and then continued as a single agent as adjuvant treatment after surgery [10/2023].

The most recent FDA approval was for the neoadjuvant and adjuvant treatment of resectable NSCLC based on results from KEYNOTE-671. With this approval, FDA issued a post-marketing commitment (PMC) (4531-2) to "Conduct an integrated analysis from ongoing, completed, or planned clinical trials and other potential data sources as appropriate enrolling a sufficient representation of older adults ages 75 years and older, and United States (US) racial and ethnic minority patients that is reflective of the US population of patients with NSCLC, to further characterize the efficacy and safety of pembrolizumab in combination with platinum-containing chemotherapy and pembrolizumab as a single agent in these patients."

To fulfill a portion of the FDA PMC, Sponsor is proposing to conduct a real-world post-authorization effectiveness study (PAES) to characterize the real-world effectiveness (overall survival) of pembrolizumab for the treatment of NSCLC among US patients across races, ethnicities, and age groups. This real-world study will complement an integrated analysis of clinical trial data.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

6 RESEARCH QUESTION AND OBJECTIVES

This study will characterize the real-world effectiveness of pembrolizumab (monotherapy or in combination with platinum-based chemotherapy [PBC], separately) for the treatment of NSCLC across races, ethnicities, and age groups. Population, Intervention, Comparison, Outcomes, Time, and Study design (PICOTS) criteria are provided in Table 1 and detailed inclusion and exclusion criteria are listed in Section 7.2.

Primary study objectives:

1. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **early-stage NSCLC** by estimating real-world overall survival (rwOS) among the following subgroups:
    - Race (e.g., Black or African American, Asian, other [American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and other race] and White patients)
    - Hispanic/Latino ethnicity (Hispanic/Latino patients and not Hispanic/Latino patients)
    - Age (elderly patients ≥75 years and younger patients <75 years)
2. To evaluate the real-world effectiveness of **pembrolizumab in combination with PBC** in US patients with **early-stage NSCLC** by estimating rwOS among the above subgroups
3. To evaluate the real-world effectiveness of **pembrolizumab monotherapy** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups
4. To evaluate the real-world effectiveness of **pembrolizumab in combination with PBC** in US patients with **advanced NSCLC** by estimating rwOS among the above subgroups

Exploratory study objectives:

1. To evaluate the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 1) and pembrolizumab in combination with PBC (see Primary Objective 2) in US patients with **early-stage NSCLC** by estimating **real-world disease-free survival (rwDFS) and real-world event-free survival (rwEFS), respectively,** among the above subgroups
2. To apply **propensity score methods** to the evaluation of the real-world effectiveness of pembrolizumab monotherapy (see Primary Objective 3) and pembrolizumab in combination with PBC (see Primary Objective 4) in US patients with **advanced NSCLC** among the above subgroups

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Table 1 PICOTS criteria

PICOTS	Inclusion	Exclusion
Population	Adults (≥18 years of age) diagnosed with early-stage or advanced NSCLC	Prior treatment with an IO agent or an investigational drug Treatment with pembrolizumab in combination with any of the following: concurrent chemoradiation therapy; radiation therapy, or an investigational drug Participation in clinical trials for NSCLC before pembrolizumab
Intervention	Received pembrolizumab (monotherapy or in combination with platinum-based chemotherapy [PBC]) for the treatment of NSCLC	Treatment with pembrolizumab in the second-line setting or later
Comparison	No formal statistical comparisons	
Outcomes	Real-world overall survival (rwOS) among the following subgroups: • Race (e.g., Black or African American, Asian, other [American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and other race] and White patients) • Hispanic/Latino ethnicity (Hispanic/Latino patients and not Hispanic/Latino patients) • Age (elderly patients ≥75 years and younger patients <75 years) Real-world disease-free survival (rwDFS) in early-stage NSCLC patients (exploratory) Real-world event-free survival (rwEFS) in early-stage NSCLC patients (exploratory)	
Time	Patients initiating pembrolizumab between October 2, 2015 (first FDA approval of pembrolizumab for NSCLC) and June 30, 2025 (allows patients to accumulate ≥2 years of follow-up time to adequately assess rwOS); Data cutoff planned June 30, 2027	
Study design	Retrospective cohort study using structured EHR-derived data	

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

7 RESEARCH METHODS

7.1 Study Design

This retrospective cohort study will utilize electronic health record (EHR)-derived databases that are largely representative of the US NSCLC population and include sizable racial and ethnic minorities and elderly patients. A two-phased feasibility assessment is ongoing, with Phase 1 having been completed January 2025. Phase 1 of the feasibility assessment addressed data availability and determined whether a data source appropriately addressed the study objectives, and led to the selection of Flatiron Health. Phase 2 includes a comprehensive assessment to evaluate the reliability and relevancy of Flatiron data, with assessment concluding by April 2026.

7.2 Setting

Adult patients diagnosed with NSCLC and treated with pembrolizumab (monotherapy or in combination with PBC) between October 2, 2015 (first FDA approval of pembrolizumab for NSCLC) and June 30, 2025 in the US will be eligible for inclusion in this study. Eligible patients will be identified from two Flatiron databases: the early non-small cell lung cancer (eNSCLC) Flatiron Health Research Database (FHRD) and advanced non-small cell lung cancer (aNSCLC) FHRD. Early-stage and advanced NSCLC patients will be analyzed separately. Details of each database are described in Section 7.4.

The follow-up time for each patient in this study spans from the start of patient activity in the database through the date of death (if any) or last confirmed activity date through June 30, 2027 (data cutoff). In the proposed study design, patients will have opportunity to accumulate ≥2 years of follow-up time for outcome ascertainment through June 30, 2027.

Inclusion criteria

- Diagnosis of NSCLC (any stage)
    - Patients initially diagnosed with early-stage NSCLC (Stage I-IIIA or IIIB with known surgical resection).
    - Patients initially diagnosed with advanced NSCLC (Stage IIIB- IV), or who were previously diagnosed with eNSCLC and subsequently developed recurrent or progressive disease.
- Age 18 years or older at the time of NSCLC diagnosis
- Received ≥1 cycle of pembrolizumab (monotherapy or in combination with PBC) for the treatment of NSCLC
    - Patients with evidence of treatment with pembrolizumab as monotherapy or in combination with PBC as neoadjuvant and/or adjuvant therapy for eNSCLC will be selected from the eNSCLC FHRD.

  - Patients with evidence of treatment of pembrolizumab as monotherapy or in combination with PBC as first-line therapy in the advanced setting will be selected from the aNSCLC FHRD.

- Pembrolizumab initiation date between October 2, 2015 and June 30, 2025 (to allow ≥2 years of follow-up time)
- Known (non-missing) race, ethnicity, or age (depending on the subgroup of interest)

Exclusion criteria

- Prior treatment with an IO agent or an investigational drug on or before initiation of pembrolizumab
  - Patients selected from the aNSCLC FHRD who received treatment with an IO agent (including pembrolizumab) or an investigational drug before the initiation of pembrolizumab *in the advanced setting* will be excluded.
- Treatment with pembrolizumab in combination with any of the following: concurrent chemoradiation therapy (CCRT), radiation therapy (RT), or an investigational drug
  - Patients treated with pembrolizumab as monotherapy or in combination with PBC in the neoadjuvant or adjuvant setting with evidence of CCRT or RT will be excluded based on abstracted RT data.
  - Patients treated with pembrolizumab as monotherapy or in combination with PBC in the advanced setting will be assumed to not concurrently be receiving CCRT or RT.
- Treatment with pembrolizumab in the second-line setting or later
  - Patients selected from the aNSCLC FHRD who were treated with pembrolizumab in the second-line setting (or later) will be excluded.

Eligible patients will meet the inclusion and exclusion criteria listed above. Patients will not be excluded if they have a prior cancer diagnosis (including lung cancer), receive pembrolizumab outside of NCCN guidelines or approved indications (i.e., off-label or compassionate use), or participate in clinical trials that include treatment with an investigational drug for NSCLC after treatment with pembrolizumab.

7.3 Variables

7.3.1 Exposure

The primary exposure is treatment with ≥1 cycle of pembrolizumab. Treatment details will be extracted from structured data to derive relevant variables of interest, including but not limited to number of pembrolizumab cycles, treatment type (i.e., monotherapy or combined with PBC), pembrolizumab duration, year of first pembrolizumab administration,

chemotherapy name, chemotherapy duration, date of surgery, and radiation therapy type, if applicable and as available.

7.3.2 Outcomes

Health outcomes: Clinical events or outcomes which may be represented as diagnoses, treatment or procedures (examples include syncope, disease progression or hypoglycemia collected as study endpoints, or death). Any herein described health outcomes, collected per the protocol, will be summarized as part of any interim analysis (if required) and in the final study report.

The primary outcome of interest for the current study will be death, as measured by rwOS, defined as the time from pembrolizumab treatment start to date of death. rwOS is a validated endpoint to measure the effectiveness of advanced cancers, including NSCLC [Lasiter, L., et al 2022] [Curtis, M. D., et al 2018], and some early-stage cancers, such as breast cancer [Zhang, Q., et al 2021]; however, rwOS has not yet been validated for early-stage NSCLC. Patients without a date of death will be censored on the date of the last follow-up or data cutoff date (June 30, 2027), whichever occurs first. A composite mortality variable that already exists in the FHRDs will be used. This composite mortality variable represents a patient's vital status and date of death, combining publicly available Social Security Death Index (SSDI) data and obituary data (processed from obituaries, funeral homes and other sources) with the EHR-derived mortality data (i.e., structured date of death and abstracted information from available EHR documentation) [Curtis, M. D., et al 2018] [Zhang, Q., et al 2021]. Separate estimates of rwOS will be provided for each of the following: race, ethnicity, and age group.

In the eNSCLC FHRD, analyses of rwDFS and rwEFS as exploratory endpoints will be conducted for each of the following subgroups: race, ethnicity, and age group. The rwDFS endpoint assesses time from the date of a patient's surgery with evidence of complete resection to first instance of local recurrence, distant recurrence, progression, or death. rwEFS assesses the time from index date (i.e., pembrolizumab therapy start date) to first instance of local recurrence, distant recurrence, progression, or death. For both endpoints, patients without an event will be censored on the date of the last follow-up or data cutoff date (June 30, 2027), whichever occurs first. Both analyses will leverage composite variables derived from EHR data elements, although neither outcome has been validated in the eNSCLC FHRD.

7.3.3 Covariates

Covariates

Race and ethnicity are primary variables of interest. Racial groups will include White, Black or African American, Asian, and other (American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, multiple races, etc.). Individual racial groups will remain separate to the extent possible given sample size limitations. Any subgroup with <75 patients

will be combined into 'other'. Final subgroup categories will be based on actual patient numbers at the time of data cut off. Ethnicity will be categorized as Not Hispanic/Latino or Hispanic/Latino.

Race and ethnicity information is collected as structured data in the EHR based on information collected from patients via intake interviews and forms, but exact processes may vary between oncology practices in the Flatiron Health Research Network (FHRN). Race categories with low representation in the US population (American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander), and source data that contains a race description which falls in multiple race categories are included in 'other race'. Source data whose race description states that patient declines to answer, or where it is otherwise unknown or uncoded is represented as NULL (i.e., missing) values. Patients missing race or ethnicity will be excluded from race- or ethnicity-specific analyses, respectively, but included in the age-specific analyses. Race data collected from the EHR in this manner has been used in peer-reviewed research on racial disparities in oncology [Maignan, K., et al 2022] [Tan, R., et al 2022] [Guadamuz, J. S., et al 2024].

Age is also a primary variable of interest. Patients will be categorized as elderly (≥75 years of age) or younger (<75 years of age) based on age at initiation of pembrolizumab treatment.

Other Patient Characteristics for Descriptive Purposes

Other patient characteristics may be used to describe patients included in each analysis, including sex, ECOG performance status, smoking status, disease stage, histology, PD-L1 status, actionable genomic alteration status, treatment setting (e.g., adjuvant), and chemotherapy agents. The status of actionable genomic alterations sourced from the EHR is available in the aNSCLC FHRD (EGFR, ALK, KRAS, ROS1, BRAF, MET, RET, NTRK, HER2) and the eNSCLC FHRD (EGFR, ALK, HER2).

7.4 Data Sources

Two databases will be used to identify eligible patients: the eNSCLC FHRD and aNSCLC FHRD. Both FHRDs are longitudinal databases, comprising deidentified patient-level data, curated via technology-enabled abstraction. During the study period, the deidentified data are expected to originate from approximately 280 US cancer clinics (~800 sites of care). The majority of patients in the FHRDs originate from community oncology settings, with the remainder from academic medical centers.

The eNSCLC FHRD includes patients with an international classification of disease (ICD) code for lung cancer (ICD-9 162.x or ICD-10 C34x or C39.9) who visited a practice in the FHRN at least twice on or after January 1, 2019, and were included in a probabilistic sample of patients for abstraction to confirm a diagnosis of early (Stage I-IIIA) NSCLC with diagnosis date on or after January 1, 2019.

The aNSCLC FHRD includes patients with an ICD code for lung cancer (ICD-9 162.x or ICD-10 C34x or C39.9) who visited a practice in the FHRN at least twice on or after January

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

1, 2011, and were included in a probabilistic sample of patients for abstraction to confirm a diagnosis of advanced (Stage IIIB or higher) or recurrent NSCLC with diagnosis date on or after January 1, 2011.

Study data will be based on structured data and previously extracted data fields which will be sourced from EHR data. The exposures of interest, along with outcomes and other relevant variables will be assessed for reliability and validity.

A final dataset will be determined that meets the minimum data standards as set forth in Section 7.4.2 Feasibility Assessment and the following:

- Identification of the study population as set forth in Section 7.4.2 Feasibility Assessment
- Evaluation of the study exposures, outcomes, and covariates as set forth in Section 7.3 Variables

7.4.1 Study Procedures

The proposed study will use commercially or publicly available de-identified secondary data sources. Research using de-identified data from commercially or publicly available secondary data sources represents the lowest risk to potential subjects because it involves the collection of anonymous or publicly available data. Large electronic healthcare databases purchased through suppliers like Flatiron Health contain a significant level of protection against the release of personal information to outside entities.

7.4.2 Feasibility Assessment

A two-phased feasibility assessment is ongoing. Phase 1 (completed) consisted of an initial assessment to determine whether the necessary data were available and to select a data source that appropriately addressed the study objectives. The results of Phase 1 led to the selection of the aforementioned Flatiron databases. Phase 2 (ongoing) consists of a more thorough evaluation of the Flatiron databases, including an examination of data reliability and relevancy, and will conclude by April 2026. The following standards are being used in the Phase 2 feasibility assessment.

- Data are fit-for-purpose: Relevancy and reliability will be assessed by checking availability, accuracy, and completeness of data needed to identify the study population based on eligibility criteria (i.e., exposure, outcome, and subgroup of interest), as well as the representativeness of the data source in identifying a diverse sample of NSCLC patients that is representative of the general NSCLC population eligible for the exposure of interest (i.e., treatment with pembrolizumab). In addition to race, ethnicity, and age distribution, representativeness may include broad geographical coverage and a mix of academic and community practice settings.

- Data capture and validity of critical information: The data source must contain adequate detail and completeness to assess diagnosis of NSCLC, age and date at pembrolizumab initiation, treatment type (i.e., monotherapy or combined with PBC), and race/ethnicity. The minimum data needed for assessment of rwOS include date of death or date of last follow-up, although validated composite mortality endpoints are preferred. If a critical variable must be ascertained and is not directly available within the source data, this assignment must be clearly defined and assessed for validity.

- Data curation and missingness: The dataset owners/users must be able to provide adequate documentation/understanding of data entry practices, source documentation, transformation (if applicable), and production of a study-specific dataset.

If Phase 2 feasibility assessment results suggest that Flatiron databases are unable to meet all minimum requirements listed above, the Sponsor will propose changes to the study design/approach that minimize any impact on evaluating the objectives.

7.5 Study Size

The study will include all patients in the research databases who meet study inclusion and exclusion criteria within the data cutoff period. Early-stage NSCLC analyses will include fewer patients and events than advanced NSCLC analyses given the more recent pembrolizumab approvals in early-stage populations. Table 2 includes preliminary sample size estimates of patient subgroups, stratified by stage group (i.e., FHRD database), as of April 30, 2025. Sample sizes do not account for stratification by treatment type (pembrolizumab monotherapy vs. pembrolizumab in combination with PBC), therefore the number of patients included in each subgroup will be further reduced in the final analysis.

Table 2 Preliminary estimates of study sample size, by stage group

Eligible patient population (as of April 30, 2025)	**Early-stage NSCLC Stage I-IIIA + Resected IIIB [eNSCLC FHRD]**	**Advanced NSCLC Stage IIIB-IV [aNSCLC FHRD]**
Adult NSCLC patients	10,787	109,074
Received pembrolizumab	1,244	25,623
Race		
Asian	24	531
Black or African American	103	2,737
White	893	17,298
Other or Unknown	222	5,017
Ethnicity		
Hispanic or Latino	41	918
Not Hispanic or Latino	914	18,802
Unknown	289	5,903
Age		
<75 years	627	12,922
≥75 years	617	12,701

As this study will be descriptive in nature, no hypothesis tests or formal power calculations will be performed. Assuming an expected median OS of 1.67 years (20 months) that is consistent across all subgroups, precision estimates (95% CI) are provided for each subpopulation across a range of overall sample sizes.

Table 3 Estimates of precision for expected median OS (1.67 years), by overall sample size and expected proportion of subpopulations

Subpopulation proportion	N=1,000 95% CI	N=2,500 95% CI	N=5,000 95% CI	N=7,500 95% CI	N=10,000 95% CI
2% (Asian)	(0.98-3.48)	(1.17-2.58)	(1.29-2.25)	(1.35-2.12)	(1.39-2.05)
5% (Hispanic or Latino)	(1.17-2.58)	(1.32-2.17)	(1.41-2.01)	(1.46-1.94)	(1.48-1.9)
13% (Black or African American)	(1.33-2.16)	(1.44-1.96)	(1.50-1.87)	(1.53-1.83)	(1.55-1.81)
85% (White)	(1.46-1.92)	(1.57-1.78)	(1.60-1.74)	(1.61-1.73)	(1.62-1.72)
41% (≥75 years)	(1.52-1.84)	(1.53-1.82)	(1.57-1.78)	(1.59-1.76)	(1.6-1.74)

7.6 Data Management

All data collected for the study should be recorded accurately, promptly, and legibly. For primary data collection, the investigator or qualified designee is responsible for recording and verifying the accuracy of subject data. For data not obtained from a primary source (i.e., secondary data, such as claims and electronic health records), the investigator is responsible for reviewing data quality and relevance to the best of the investigator's knowledge. By signing this protocol either electronically or written, the investigator confirms that the quality and relevance of data has been assessed to meet the minimum requirements for all study objectives. If this study has been outsourced, the institutional policies of the supplier should be followed for development of data management plans. However, the supplier should ensure compliance with Good Pharmacoepidemiology Practice, and all applicable federal, state, and local laws, rules and regulations relating to the conduct of the study.

Data Management Software and Hardware:

Statistical analyses will be implemented using relevant software, including SAS or R.

Description of Data Preparation and Methods for Data Retrieval and Collection:

The investigator will review all documentation related to data preparation, retrieval, and cleaning. Detailed descriptions of data processing, including abstraction, curation, quality assurance, and quality checks will be provided once a supplier has been selected.

7.7 Programming Quality

This study will incorporate the following quality checks for data analysis and reporting programming:

- Creating a program requirements and specification (PRS) document

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

- Developing and testing of statistical programs which includes ensuring the programs run successfully and all output are reviewed to ensure they meet the criteria included in the PRS. This includes validating that all inputs (metadata or parameter values) are correctly specified in the programs and are consistent with the PRS.

- Independent Review and Testing, conducted by a second programmer to ensure that the input and outputs of the programs created by the first programmer meet the documented PRS. This includes the activities listed below. If discrepancies are found, the programmers will discuss and repeat steps until consistency is achieved.
    - Review of code to ensure the program aligns with the PRS
    - Execution of code and review of results for some or all scenarios

      And may include the following activity:
    - Parallel programming of a small piece of critical code

- Review of outputs/results to ensure accuracy and format of each deliverable.

7.8 Data Analysis

Data analyses will be conducted through 4Q2027. Operational details further defining variables, analytic methodology, and table shells will be defined in a separate detailed statistical analysis plan.

Analysis Plan

The primary study objectives will require four study cohorts to evaluate the real-world effectiveness of pembrolizumab in US NSCLC patients across racial, ethnic, and age groups:

1. Early-stage patients treated with pembrolizumab monotherapy,

2. Early-stage patients treated with pembrolizumab in combination with PBC,

3. Advanced-stage patients treated with pembrolizumab monotherapy, and

4. Advanced-stage patients treated with pembrolizumab in combination with PBC

For each of the four study cohorts, sociodemographic and clinical characteristics among NSCLC patients receiving pembrolizumab in a real-world setting will be described among racial, ethnic, and age groups. For categorical variables, frequencies, and percentages will be generated. For continuous variables, descriptive statistics will include means, standard deviations, medians, interquartile ranges, and/or minimum and maximum values, as applicable. The number of patients with unknown values for both categorical and continuous variables will be provided. Given the descriptive nature of the study, no statistical testing will be performed.

Real-world OS as defined in Section 7.3.2 will be estimated among patients using Kaplan-Meier methods. Median survival times and the probabilities of survival at key milestones of interest (e.g., 1-year, 5-year, etc.) and corresponding 95% confidence intervals (CI) will be calculated. Separate analyses will be conducted for each the following subgroups: Black or African American, Asian, other race, White, Hispanic/Latino, not Hispanic/Latino, elderly (≥75 years), and younger (<75 years). Racial subgroups with small sample sizes (<75 patients) may be combined with 'other race'. A descriptive summary of rwOS estimates will be provided for all races, ethnicities, and age groups. Given the descriptive nature of the study, no imputation for missing data or statistical testing will be performed. Sensitivity analyses may be conducted to evaluate the effectiveness of pembrolizumab among the subset of patients receiving pembrolizumab in line with currently approved indications (i.e., on-label use).

Exploratory analyses of rwDFS and rwEFS as defined in Section 7.3.2 will be conducted in the early-stage NSCLC population. The rwDFS endpoint will be evaluated among patients analyzed in Primary Objective 1 (i.e., surgically-resected patients who did not receive neoadjuvant therapy prior to surgery). The rwEFS endpoint will be evaluated among patients analyzed in Primary Objective 2 (i.e., resectable patients who received neoadjuvant therapy). For both outcomes Kaplan-Meier methods will be used to estimate median survival times and the probabilities of survival at key milestones of interest and corresponding 95% CI will be calculated. Separate analyses will be conducted for each race, ethnicity, and age group. Sensitivity analyses may be conducted to evaluate the effectiveness of pembrolizumab among the subset of patients receiving pembrolizumab in line with currently approved indications.

An exploratory analysis of rwOS using propensity score weighting methods will be conducted in the advanced NSCLC population to balance baseline characteristics that may confound the effect of each exposure group (i.e., race, ethnicity, or age group) on survival following pembrolizumab treatment. Due to sample size limitations, propensity score methods will not be applied to the early-stage population. Propensity scores will be estimated using multinomial logistic regression for multiple group comparisons (i.e., race) and logistic regression for binary group comparisons (i.e., ethnicity or age group), incorporating covariates associated with group membership and/or the outcome of interest. The selection of covariates to be used for propensity scoring will be determined based on subject-matter expertise, a review of published literature, and results of the completed feasibility assessment. Covariates must have minimal missingness and may include variables such as sex, ECOG performance status, smoking status, disease stage at initial diagnosis, histology, PD-L1 status, brain metastasis, age (for the analyses of race and ethnicity), and race (for the analysis of age). If a covariate could also act as a mediator (i.e., lies along the causal pathway by which race, ethnicity, or age influence survival), it will be excluded from propensity scoring. Missingness may be considered as a separate category for each covariate. The final propensity score model will be selected based on model stability, parameter estimation accuracy, and the distribution of predicted propensity scores. Models showing extreme propensity score distributions (values approaching 0 or 1) or poor parameter estimation will

be considered inadequate, and alternative model specifications will be evaluated. Once estimated, the propensity scores will be used to create balanced comparisons between subgroups using an appropriate weighting method, such as inverse probability weighting (IPW). Weights will be derived from the marginal distribution of covariates in the entire cohort, so that each group more closely resembles the overall study population. Stabilized weights will be applied to reduce the influence of extreme weights and improve precision of estimates. Covariate balance will be assessed using standardized mean differences (SMD) before and after propensity score adjustment. An absolute SMD < 0.1will be considered indicative of adequate balance. Additional balance diagnostics may include overlap assessment through propensity score distribution plots, variance ratios for continuous variables, and Kolmogorov-Smirnov tests for distributional balance. If adequate balance is not achieved through the initial approach, alternative weighting strategies or adjustments to covariate inclusion will be explored. Following propensity score adjustment, rwOS outcome analyses will be conducted on the balanced pseudo-population. Weighted Kaplan-Meier estimators will be used to generate median survival times and probabilities of survival at key milestones along with 95% CI, separately for each treatment type.

7.9 Quality Control

By signing this protocol, all parties agree to following applicable standard operating procedures (SOPs). All parties also agree to ensuring all existing and new study personnel are appropriately trained to ensure the study is conducted and data are generated, documented, and reported in compliance with the protocol, Good Pharmacoepidemiology Practice (GPP), Good Pharmacovigilance Practices (GVP), and all applicable federal, state, and local laws, rules and regulations. All parties should maintain transparency and open communication in order to effectively manage the study and proactively mitigate any risks.

The Sponsor may conduct routine or for-cause audits to ensure oversight and conduct of the study are completed in accordance with the protocol, quality standards (e.g., GPP and GVP), and applicable laws and regulations. If a significant quality issue (SQI) is identified at any time during the conduct of the study, it must be escalated to the Sponsor immediately. A SQI is any issue with the potential to negatively impact, either directly or indirectly, the rights, safety and well-being of patients or study participants and/or the integrity of the data. In the event an audit or SQI results in corrective or preventive actions, all parties are expected to appropriately implement the action plan in a timely manner.

Flatiron Health will conduct all activities related to developing and ensuring the quality of the real-world dataset for this research study. This includes the selection of patients for the study cohort and the inclusion of relevant data variables. Flatiron will also manage all activities involved in ensuring the quality and delivery of the data, including database creation, approval by central IRB, quality assurance of data, and building and securely transmitting the data files to Sponsor.

Flatiron Health will comply with the audits described under this section in accordance with Flatiron Health’s contract with Sponsor for this study, and Flatiron Health will support

quality assurance verification activities as reasonably requested by Sponsor, or a third party on behalf of Sponsor. Additionally, auditors/inspectors will be limited to reviewing documentation specific to this study in accordance with Flatiron Health's policy regarding confidential information.

7.10 Limitations of the Research Methods

Validity and interpretability of measurement:

- EHR-derived data are frequently collected for the purpose of the clinical management of oncology patients and not primarily for research purposes. The proposed study is subject to data that are available in the selected database and the level of granularity provided for available data.

- OS estimates may be biased due to possible misclassification of events resulting from incomplete or low sensitivity mortality data. However, the mortality data in the secondary data sources proposed for this study are enhanced by linking multiple data sources including data abstraction from EHR, SSDI, and obituaries, and the mortality variable used in the FHRDs has been shown to have high sensitivity (>90%) when compared to the National Death Index [Curtis, M. D., et al 2018] [Zhang, Q., et al 2021]. Validation of the mortality endpoint is one of the required criteria needed to select the data source used for the study.

- OS estimates may be biased due to uncontrolled confounding factors, including those that are available in the dataset (e.g., stage, treatment setting, year of diagnosis) and those that are not available in the dataset (e.g., age-related comorbidities). An exploratory analysis using propensity score weighting methods will attempt to balance confounders across subgroups using available covariates. It should be noted that the use of propensity score weighting may result in less precise effect estimates.

- Combining racial/ethnic minorities does not capture the heterogeneity that may exist within individual races or ethnicities. Individual racial groups will remain separate to the extent possible given sample size limitations.

Data completeness:

- As is common with observational studies utilizing real-world data, the main limitations for this study include lack of availability of data on certain variables and missingness in data obtained from the EHR. Certain variables may not be available with full completeness, such as race/ethnicity and treatment details.

- Additionally, patients may seek care outside of the supplier's network. This may contribute to missing, incomplete data and patient loss to follow-up.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Study design:

- Although the same patient may be present in both eNSCLC and aNSCLC FHRD databases, the exclusion criteria pertaining to prior treatment ensure that each patient will only be included in the study sample once. Patients included in the eNSCLC FHRD who received pembrolizumab in both the early- and advanced-stage settings will only be included if they meet all eligibility criteria at earliest initiation of pembrolizumab (i.e., in the early-stage setting), while patients included in the aNSCLC FHRD who have evidence of treatment with pembrolizumab in both the early- and advanced-stage settings will not be included.
- Study inclusion/exclusion criteria are intended to reflect a real-world patient population. As such, study populations may not mirror the corresponding clinical trial populations and, as a result, observed results may not be comparable to those reported using clinical trial data. Sensitivity analyses may be conducted to evaluate the effectiveness of pembrolizumab among subsets of patients receiving pembrolizumab in line with currently approved indications (i.e., on-label use).
- Mortality data may not have sufficient maturity for all patients. For example, OS among patients with early-stage NSCLC recently treated with pembrolizumab may be particularly likely to be impacted by limited data maturity, affecting interpretability. Exploratory outcomes rwEFS and rwDFS will be evaluated for early-stage patients to mitigate the limitation in data maturity. Since neither outcome has been validated for use in the eNSCLC FHRD, results should be interpreted with caution.

Generalizability of study results:

- The data source may not be representative of the overall US NSCLC population. Of note, the eNSCLC FHRD and aNSCLC FHRD are similar databases that select probabilistic samples of patients diagnosed with eNSCLC or aNSCLC from practices in the FHRN. The eNSCLC FHRD samples approximately 17% of patients diagnosed with eNSCLC from the FHRN, while the aNSCLC FHRD samples approximately 52% of patients with aNSCLC from the FHRN. Consequently, a pooled sample combining patients from both databases will include a higher proportion of patients treated in the advanced setting than would a representative sample. There is evidence that the aNSCLC FHRD share a similar demographic and geographic distribution with the US Surveillance, Epidemiology, and End Results program (SEER) program and the National Program of Cancer Registries (NPCR) [Ma, X., et al 2020], but study results may not be generalizable to NSCLC patient populations globally due to variation in patient demographic and clinical characteristics and standard of care.
- OS estimates may be biased due to confounding factors that may not be available in the dataset (e.g., socioeconomic status, age-related comorbidities, etc.).

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

8 PROTECTION OF HUMAN SUBJECTS

8.1 Informed Consent

This study will not require participant informed consent. No patient contact will be sought for this project and all data will be retrospectively accessed. As such, no patient consent is sought or needed as part of this study.

This study may require participant IRB/EC review.

Data used in this study will be anonymized and no personal identifiers will be available to maintain patient confidentiality. Studies of this type are typically determined to be exempt from IRB review. Flatiron will also manage all activities involved in ensuring the quality and delivery of the data, including database creation and approval by central IRB.

Investigators shall ensure that personal identifiers will be removed from any study files that are accessible to non-study personnel in accordance with applicable laws and regulations. Whenever feasible, study files should be coded and stripped of personal identifiers, and code keys should be stored separately from study files.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

9 MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse Event (AE) and Product Quality Complaint (PQC) Reporting Language for Non-Interventional Study Protocols

Adverse Event and Product Quality Complaint Reporting

This is a non-interventional database study based on secondary use of data collected for other purposes. No administration of any therapeutic or prophylactic agent is required in this protocol. No reporting of individual adverse events or product quality complaints to regulatory agencies is planned for this database study because there is no access to individual patient/subject records and it is not possible to assess the causality of individual cases. The investigator should refer to their institution's policy or local laws and regulations regarding reporting of any suspected adverse reactions and product quality complaints.

Any health outcomes (if collected per Section 7.3), including any that qualify as adverse events, will be summarized as part of any interim analysis (including safety analysis, if required) and in the final study report, which will be provided to regulatory agencies by the Sponsor as required. Any relevant safety information will be summarized, and the Sponsor will include in the appropriate Periodic Safety Update Report (PSUR)/Periodic Benefit Risk Evaluation Report (PBRER) and/or Development Safety Update Reports (DSUR), if required.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

10 PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

In addition to submitting the primary results of this study to the FDA in the form of a final study report, results may be externally disseminated in a manuscript submitted to a peer-reviewed, scientific journal, abstract/presentation at a scientific conference or symposium, or results posted on the HMA-EMA Catalogue of RWD Studies. Any publication related to the study will need to be reviewed/approved by the Sponsor prior to submitting results externally. Any publication resulting from this work will adhere to the procedures and pre-specified analysis plans within this protocol.

11 REFERENCES

[Aldrich, M. C., et al 2019]	Aldrich MC, Mercaldo SF, Sandler KL, Blot WJ, Grogan EL, Blume JD. Evaluation of USPSTF lung cancer screening guidelines among African American Adult Smokers. JAMA Oncol. 2019 Sep;5(9):1318-24.	[08C2G0]
[Aldrighetti, C. M., et al 2021]	Aldrighetti CM, Niemierko A, Van Allen E, Willers H, Kamran SC. Racial and ethnic disparities among participants in precision oncology clinical studies. JAMA Netw Open. 2021 Nov 8;4(11):e2133205.	[08P7F0]
[American Association for Cancer Research 2022]	American Association for Cancer Research. AACR cancer disparities progress report 2022. Philadelphia (PA): American Association for Cancer Research (AACR); 2022. 200 p.	[08H0Z2]
[American Association for Cancer Research 2024]	American Association for Cancer Research. AACR cancer disparities progress report 2024. Philadelphia (PA): American Association for Cancer Research (AACR); 2024. 216 p.	[08N0T6]
[American Cancer Society 2022]	American Cancer Society. Cancer facts and figures 2022. Atlanta (GA): American Cancer Society (ACS); 2022. 80 p.	[07ZJ90]
[American Cancer Society 2023]	American Cancer Society [Internet]. Atlanta (GA): American Cancer Society (ACS); c2023. Lung cancer survival rates; [cited 2023 Dec 14]; [about 5 screens]. Available from: https://www.cancer.org/cancer/types/lung-cancer/detection-diagnosis-staging/survival-rates.html#5-year-relative-survival-rates-for-non-small-cell-lung-cancer%201/2.	[08GXGR]
[Barta, J. A., et al 2019]	Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. Ann Glob Health. 2019 Jan 22;85(1):8.	[05DQP9]
[Blom, E. F., et al 2020]	Blom EF, Ten Haaf K, Arenberg DA, de Koning HJ. Disparities in receiving guideline-concordant treatment for lung cancer in the United States. Ann Am Thorac Soc. 2020 Feb;17(2):186-94.	[08P7F2]

[Curtis, M. D., et al 2018]	Curtis MD, Griffith SD, Tucker M, Taylor MD, Capra WB, Carrigan G, et al. Development and validation of a high-quality composite real-world mortality endpoint. Health Serv Res. 2018 Dec;53(6):4460-76.	[08SNMQ]
[DeRouen, M. C., et al 2022]	DeRouen MC, Canchola AJ, Thompson CA, Jin A, Nie S, Wong C, et al. Incidence of lung cancer among never-smoking Asian American, Native Hawaiian, and Pacific Islander females. J Natl Cancer Inst. 2022;114(1):78-86.	[08C2G4]
[Ganti, A. K., et al 2021]	Ganti AK, Klein AB, Cotarla I, Seal B, Chou E. Update of incidence, prevalence, survival, and initial treatment in patients with non-small cell lung cancer in the US. JAMA Oncol. 2021 Dec;7(12):1824-32.	[07Y2MQ]
[Guadamuz, J. S., et al 2024]	Guadamuz JS, Wang X, Altomare I, Camelo Castillo W, Sarkar S, Mamtani R, et al. Mediators of racial and ethnic inequities in clinical trial participation among patients with cancer, 2011-2023. JNCI Cancer Spectr. 2024;8(5):pkae085.	[08SNMP]
[Hamel, L. M., et al 2016]	Hamel LM, Penner LA, Albrecht TL, Heath E, Gwede CK, Eggly S. Barriers to clinical trial enrollment in racial and ethnic minority patients with cancer. Cancer Control. 2016 Oct;23(4):327-37.	[08P7F3]
[Kilic, S., et al 2024]	Kilic S, Zhao J, Okut H, Jani CT, Radwan A, Dudipala H, et al. Disparities in US lung cancer clinical trial enrollment. J Racial Ethn Health Disparities. 2024;11:3201-9.	[090299]
[Kratzer, T. B., et al 2022]	Kratzer TB, Jemal A, Miller KD, Nash S, Wiggins C, Redwood D, et al. Cancer statistics for American Indian and Alaska Native individuals, 2022: including increasing disparities in early onset colorectal cancer. CA Cancer J Clin. 2022:1-27.	[08FND0]

[Lasiter, L., et al 2022]	Lasiter L, Tymejczyk O, Garrett-Mayer E, Baxi S, Belli AJ, Boyd M, et al. Real-world overall survival using oncology electronic health record data: Friends of Cancer Research Pilot. Clin Pharmacol Ther. 2022 Feb;111(2):444-54.	[08PF6Y]
[Ma, X., et al 2020]	Ma X, Long L, Moon S, Adamson BJS, Baxi SS. Comparison of population characteristics in real-world clinical oncology databases in the US: flatiron health, SEER, and NPCR [manuscript]. 2020. 88 p.	[09029D]
[Maignan, K., et al 2022]	Maignan K, Fashoyin-Aje LA, Torres AZ, Fernandes LL, Gwise T, Baxi SB, et al. Exploring racial disparities in treatment patterns and outcomes for patients with multiple myeloma using real world data. Blood Cancer J. 2022;12:65.	[08SNMW]
[Medrek, S. and Szmit, S. 2022]	Medrek S and Szmit S. Are cardiovascular comorbidities always associated with a worse prognosis in patients with lung cancer? Front Cardiovasc Med. 2022 Sep 23;9:984951.	[08P7F5]
[Miao, E., et al 2023]	Miao E, Klugman M, Rohan T, Dean Hosgood H. Hypothesized explanations for the observed lung cancer survival benefit among Hispanics/Latinos in the United States. J Racial Ethn Health Disparities. 2023;10:1339-48.	[08C3RF]
[Mills, S., et al 2023]	Mills S, Donnan P, Buchanan D, Smith BH. Age and cancer type: associations with increased odds of receiving a late diagnosis in people with advanced cancer. BMC Cancer. 2023;23:1174.	[08P7F6]
[National Cancer Institute (NCI) 2023]	SEER*Stat [Internet]. Bethesda (MD): National Cancer Institute (NCI). 2023. Age-specific incidence of stage NSCLC by age and race in 2018-2021, SEER 17; 20 p. Available from: https://seer.cancer.gov/seerstat/analysis.html.	[08M3K0]

[National Cancer Institute 2023]	Surveillance, Epidemiology, and End Results Program [Internet]. Bethesda (MD): National Cancer Institute (NCI). Cancer stat facts: lung and bronchus cancer; [cited 2023 Mar 31]; [about 18 screens]. Available from: https://seer.cancer.gov/statfacts/html/colorect.html.	[088Y52]
[National Cancer Institute 2024]	Surveillance, Epidemiology, and End Results Program [Internet]. Bethesda (MD): National Cancer Institute (NCI). Cancer stat facts: lung and bronchus cancer; [cited 2024 Apr 24]; [about 26 screens]. Available from: https://seer.cancer.gov/statfacts/html/lungb.html.	[08KPSP]
[Raz, D. J., et al 2008]	Raz DJ, Gomez SL, Chang ET, Kim JY, Keegan THM, Pham J, et al. Epidemiology of non-small cell lung cancer in Asian Americans: incidence patterns among six subgroups by nativity. J Thorac Oncol. 2008 Dec;3(12):1391-7.	[08B5D6]
[Sedrak, M. S., et al 2021]	Sedrak MS, Freedman RA, Cohen HJ, Muss HB, Jatoi A, Klepin HD, et al. Older adult participation in cancer clinical trials: a systematic review of barriers and interventions. CA Cancer J Clin. 2021;71:78-92.	[08P7F7]
[Siegel, R. L., et al 2022]	Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022 Jan/Feb;72(1):7-33.	[07YSZM]
[Siegel, R. L., et al 2024]	Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74:12-49.	[08HZXM]
[Sogaard, M., et al 2013]	Sogaard M, Thomsen RW, Bossen KS, Sorensen HT, Norgaard M. The impact of comorbidity on cancer survival: a review. Clin Epidemiol. 2013 Nov 1;5(Suppl 1):3-29.	[08P7F8]

[Sung, H., et al 2021]	Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May-Jun;71(3):209–49.	[0892DT]
[Tan, R., et al 2022]	Tan R, Cassoli L, Yan Y, Shen V, Day BM, Mitchell EP. Assessing real-world racial differences among patients with metastatic triple-negative breast cancer in US community practices. Front Public Health. 2022 May 24;10:859113.	[08SNMX]
[Zhang, Q., et al 2021]	Zhang Q, Gossai A, Monroe S, Nussbaum NC, Parrinello CM. Validation analysis of a composite real-world mortality endpoint for patients with cancer in the United States. Health Serv Res. 2021;56(6):1281-7.	[08SNMR]
[Zhang, Y., et al 2023]	Zhang Y, Vaccarella S, Morgan E, Li M, Etxeberria J, Chokunonga E, et al. Global variations in lung cancer incidence by histological subtype in 2020: a population-based study. Lancet Oncol. 2023 Nov;24:1206-18.	[08JD46]

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

12 ANNEXES

Annex 1 ENCePP Checklist for Study Protocols (Revision 4)

Doc.Ref. EMA/540136/2009

ENCePP Checklist for Study Protocols (Revision 4)

Adopted by the ENCePP Steering Group on 15/10/2018

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) welcomes innovative designs and new methods of research. This Checklist has been developed by ENCePP to stimulate consideration of important principles when designing and writing a pharmacoepidemiological or pharmacovigilance study protocol. The Checklist is intended to promote the quality of such studies, not their uniformity. The user is also referred to the ENCePP Guide on Methodological Standards in Pharmacoepidemiology, which reviews and gives direct electronic access to guidance for research in pharmacoepidemiology and pharmacovigilance.

For each question of the Checklist, the investigator should indicate whether or not it has been addressed in the study protocol. If the answer is "Yes", the section number of the protocol where this issue has been discussed should be specified. It is possible that some questions do not apply to a particular study (for example, in the case of an innovative study design). In this case, the answer 'N/A' (Not Applicable) can be checked and the "Comments" field included for each section should be used to explain why. The "Comments" field can also be used to elaborate on a "No" answer.

This Checklist should be included as an Annex by marketing authorisation holders when submitting the protocol of a non-interventional post-authorisation safety study (PASS) to a regulatory authority (see the Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies). The Checklist is a supporting document and does not replace the format of the protocol for PASS presented in the Guidance and Module VIII of the Good pharmacovigilance practices (GVP).

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Study title: Real-world post-authorization effectiveness study (PAES) of pembrolizumab for the treatment of NSCLC across races, ethnicities, and age groups

EU PAS Register® number: Not yet registered
Study reference number (if applicable): Not applicable

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	☒	☐	☐	4
1.1.2 End of data collection[2]	☒	☐	☐	4
1.1.3 Progress report(s)	☐	☐	☒	
1.1.4 Interim report(s)	☐	☐	☒	
1.1.5 Registration in the EU PAS Register®	☒	☐	☐	4
1.1.6 Final report of study results.	☒	☐	☐	4

Comments:

Study will be registered on the HMA-EMA Catalogue of RWD Studies after HA approval of protocol.

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain:	☒	☐	☐	6
2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue)	☒	☐	☐	5
2.1.2 The objective(s) of the study?	☒	☐	☐	6
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	☒	☐	☐	6, 7.2
2.1.4 Which hypothesis(-es) is (are) to be tested?	☐	☐	☒	
2.1.5 If applicable, that there is no *a priori* hypothesis?	☐	☐	☒	

Comments:

[1] Date from which information on the first study is first recorded in the study dataset or, in the case of secondary use of data, the date from which data extraction starts.

[2] Date from which the analytical dataset is completely available.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 3: Study design		**Yes**	**No**	**N/A**	**Section Number**
3.1	Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	☒	☐	☐	7.1
3.2	Does the protocol specify whether the study is based on primary, secondary or combined data collection?	☒	☐	☐	7.1, 7.2, 7.4
3.3	Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	☒	☐	☐	7.8
3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	☐	☒	☐	
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)	☒	☐	☐	9

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	☒	☐	☐	7.2, 7.4
4.2	Is the planned study population defined in terms of:				
	4.2.1 Study time period	☒	☐	☐	
	4.2.2 Age and sex	☒	☐	☐	7.2
	4.2.3 Country of origin	☒	☐	☐	
	4.2.4 Disease/indication	☒	☐	☐	
	4.2.5 Duration of follow-up	☒	☐	☐	
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	☒	☐	☐	7.4

Comments:

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	☒	☐	☐	7.3.1
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)	☒	☐	☐	7.3.1
5.3 Is exposure categorised according to time windows?	☐	☒	☐	
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	☒	☐	☐	7.3.1
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☒	☐	
5.6 Is (are) (an) appropriate comparator(s) identified?	☐	☐	☒	

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?	☒	☐	☐	7.3.2
6.2 Does the protocol describe how the outcomes are defined and measured?	☒	☐	☐	7.3.2
6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)	☒	☐	☐	7.3.2
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)	☐	☐	☒	

Comments:

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 7: Bias	**Yes**	**No**	**N/A**	**Section Number**
7.1 Does the protocol address ways to measure confounding? (e.g. confounding by indication)	☒	☐	☐	7.10
7.2 Does the protocol address selection bias? (e.g. healthy user/adherer bias)	☒	☐	☐	7.10
7.3 Does the protocol address information bias? (e.g. misclassification of exposure and outcomes, time-related bias)	☒	☐	☐	7.10

Comments:

Section 8: Effect measure modification	Yes	No	N/A	Section Number
8.1 Does the protocol address effect modifiers? (e.g. collection of data on known effect modifiers, sub-group analyses, anticipated direction of effect)	☒	☐	☐	7.3.3

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	☒	☐	☐	7.4, 7.3.1
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)	☒	☐	☐	7.4, 7.3.2
9.1.3 Covariates and other characteristics?	☒	☐	☐	7.4, 7.3.3
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	☒	☐	☐	7.4, 7.3.1
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)	☒	☐	☐	7.4, 7.3.2
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	☒	☐	☐	7.4, 7.3.3

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	☒	☐	☐	7.4, 7.3.1
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))	☒	☐	☐	7.4, 7.3.2
9.3.3 Covariates and other characteristics?	☒	☐	☐	7.4, 7.3.3
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)	☒	☐	☐	7.3.2, 7.4.2

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	☒	☐	☐	7.8
10.2 Is study size and/or statistical precision estimated?	☒	☐	☐	7.5
10.3 Are descriptive analyses included?	☒	☐	☐	7.8
10.4 Are stratified analyses included?	☐	☒	☐	7.8
10.5 Does the plan describe methods for analytic control of confounding?	☒	☐	☐	7.8
10.6 Does the plan describe methods for analytic control of outcome misclassification?	☐	☒	☐	
10.7 Does the plan describe methods for handling missing data?	☒	☐	☐	7.8
10.8 Are relevant sensitivity analyses described?	☐	☒	☐	

Comments:

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	☒	☐	☐	7.4, 7.6, 7.9
11.2 Are methods of quality assurance described?	☒	☐	☐	7.9
11.3 Is there a system in place for independent review of study results?	☐	☒	☐	

Comments:

Within the study conduct framework, a system of independent review of study results are not planned. However, the results within a final study report would be shared with the FDA.

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of:				
12.1.1 Selection bias?	☒	☐	☐	7.10
12.1.2 Information bias?	☒	☐	☐	7.10
12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	☒	☐	☐	7.10
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	☒	☐	☐	7.4

Comments:

Section 13: Ethical/data protection issues	**Yes**	**No**	**N/A**	**Section Number**
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	☒	☐	☐	8
13.2 Has any outcome of an ethical review procedure been addressed?	☒	☐	☐	8
13.3 Have data protection requirements been described?	☒	☐	☐	8

Comments:

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Section Number**
14.1 Does the protocol include a section to document amendments and deviations?	☒	☐	☐	3

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Section Number**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	☒	☐	☐	10
15.2 Are plans described for disseminating study results externally, including publication?	☒	☐	☐	10

Comments:

Name of the main author of the protocol: PPD PPD

Date: 11/March/2025

Signature: PPD

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

Annex 2 Administrative and Regulatory Details

Confidentiality:

Confidentiality of Data

By signing this protocol, the investigator affirms to the Sponsor that information furnished to the investigator by the Sponsor will be maintained in confidence. If applicable such information will be divulged to Institutional Review Board, Ethics Review Committee or similar or expert committee; affiliated institution and employees, only under an appropriate understanding of confidentiality with such board or committee, affiliated institution and employees. Data generated by this study will be considered confidential by the investigator, except to the extent that it is included in a publication as provided in the Publications section of this protocol.

Confidentiality of Subject Records

By signing this protocol, the investigator agrees that the Sponsor (or Sponsor representative), Institutional Review Board/Independent Ethics Committee (IRB/IEC), or Regulatory Agency representatives may consult and/or copy study documents in order to verify worksheet/case report form data. If study documents will be photocopied during the process of verifying worksheet/case report form information, the subject will be identified by unique code only; full names/initials will be masked prior to transmission to the Sponsor.

By signing this protocol, the investigator agrees to treat all subject data used and disclosed in connection with this study in accordance with all applicable privacy laws, rules, and regulations.

Confidentiality of Investigator Information

By signing this protocol, the investigator recognizes that certain personal identifying information with respect to the investigator, and all sub-investigators and study site personnel (if applicable), may be used and disclosed for study management purposes, as part of a regulatory submissions, and as required by law. This information may include:

- name, address, telephone number and e-mail address;
- hospital or clinic address and telephone number;
- curriculum vitae or other summary of qualifications and credentials; and
- other professional documentation.

Consistent with the purposes described above, this information may be transmitted to the Sponsor, and subsidiaries, affiliates and agents of the Sponsor, in your country and other countries, including countries that do not have laws protecting such information. By signing

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

this protocol, the investigator expressly consents to these uses and disclosures. Additionally, the investigator's name and business contact information may be included when reporting certain serious adverse events to regulatory agencies or to other investigators. The investigator is hereby notified that the collection, processing and sharing of their personal data with respect to adverse event reports to the Sponsor and regulatory agencies occurs on the basis of performance of a legal obligation, and the investigator expressly consents to these uses and disclosures when reporting such events to other investigators.

If this is a multicenter study, in order to facilitate contact between investigators, the Sponsor may share an investigator's name and contact information with other participating investigators upon request.

Administrative:

Compliance with Financial Disclosure Requirements

Financial Disclosure requirements are outlined in the US Food and Drug Administration Regulations, Financial Disclosure by Clinical Investigators (21 CFR Part 54). It is the Sponsor's responsibility to determine, based on these regulations, whether a request for Financial Disclosure information is required. It is the investigator's/subinvestigator's responsibility to comply with any such request.

The investigator/subinvestigator(s) agree, if requested by the Sponsor in accordance with 21 CFR Part 54, to provide his/her financial interests in and/or arrangements with the Sponsor to allow for the submission of complete and accurate certification and disclosure statements. The investigator/subinvestigator(s) further agree to provide this information on a Certification/Disclosure Form, commonly known as a financial disclosure form, provided by the Sponsor. The investigator/subinvestigator(s) also consent to the transmission of this information to the Sponsor in the United States for these purposes. This may involve the transmission of information to countries that do not have laws protecting personal data.

Compliance with Law, Audit and Debarment

By signing this protocol, the investigator agrees to conduct the study in an efficient and diligent manner and in conformance with this protocol; generally accepted standards of Good Pharmacoepidemiology Practice and all applicable federal, state and local laws, rules and regulations relating to the conduct of the study.

The investigator will comply with quality assurance audits/inspections as governed by Flatiron Health's contract with Sponsor for this study and as set forth in Section 7.9 of the protocol.

The investigator agrees not to seek reimbursement from subjects, their insurance providers or from government programs for procedures included as part of the study reimbursed to the investigator by the Sponsor.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

The Investigator shall prepare and maintain complete and accurate study documentation in compliance with Good Pharmacoepidemiology Practice, standards and applicable federal, state and local laws, rules and regulations; and, for each subject participating in the study, provide all data, and, upon completion or termination of the study, submit any other reports to the Sponsor as required by this protocol or as otherwise required pursuant to any agreement with the Sponsor.

The investigator agrees to promptly take any reasonable steps that are requested by the Sponsor as a result of any audit to cure deficiencies in the study documentation and worksheets/case report forms.

The investigator must maintain copies of all documentation and records relating to the conduct of the study in accordance with their institution's records retention schedule which is compliant with all applicable regional and national laws and regulatory requirements. If an institution does not have a records retention schedule to manage its records long-term, the investigator must maintain all documentation and records relating to the conduct of the study for 5 years after final report or first publication of study results, whichever comes later, per GPP guidelines. This documentation includes, but is not limited to, the protocol, worksheets/case report forms, advertising for subject participation, adverse event reports, subject source data, correspondence with regulatory authorities and IRBs/ERCs, consent forms, investigator's curricula vitae, monitor visit logs, laboratory reference ranges, laboratory certification or quality control procedures and laboratory director curriculum vitae. All study documents shall be made available if required by relevant regulatory authorities. The investigator must consult with the Sponsor prior to discarding study and/or subject files.

The investigator will promptly inform the Sponsor of any regulatory agency inspection conducted for this study.

Persons debarred from conducting or working on studies by any court or regulatory agency will not be allowed to conduct or work on this Sponsor's studies. The investigator will immediately disclose in writing to the Sponsor if any person who is involved in conducting the study is debarred or if any proceeding for debarment is pending or, to the best of the investigator's knowledge, threatened.

In the event the Sponsor prematurely terminates a particular study site, the Sponsor will promptly notify that site's IRB/IEC.

According to European legislation, a Sponsor must designate an overall coordinating investigator for a multi-center study (including multinational). When more than one study site is open in an EU country, Merck, as the Sponsor, will designate, per country, a national principal coordinator (Protocol CI), responsible for coordinating the work of the principal investigators at the different sites in that Member State, according to national regulations. For a single-center study, the Protocol CI is the principal investigator. In addition, the Sponsor must designate a principal or coordinating investigator to review the study report

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

that summarizes the study results and confirm that, to the best of his/her knowledge, the report accurately describes the conduct and results of the study in the study's final report. The Sponsor may consider one or more factors in the selection of the individual to serve as the Protocol CI and or CSR CI (e.g., availability of the CI during the anticipated review process, thorough understanding of study methods, appropriate enrollment of subject cohort, timely achievement of study milestones). The Protocol CI must be a participating study investigator.

Compliance with Study Registration and Results Posting Requirements

Under the terms of the Food and Drug Administration Modernization Act (FDAMA) and the Food and Drug Administration Amendments Act (FDAAA), as well as the European Medicines Agency GVP Module VIII, the Sponsor of the study is solely responsible for determining whether the study and its results are subject to the requirements for submission to one or more study registries such as the HMA EMA Catalogue of RWD Studies. Merck, as Sponsor of this study, will review this protocol and submit the information necessary to fulfill these requirements for all post-marketing safety and efficacy studies. Information posted will allow subjects to identify potentially appropriate primary data collection studies for their disease conditions and pursue participation by calling a central contact number for further information on appropriate study locations and site contact information.

By signing this protocol, the investigator acknowledges that the statutory obligations under FDAMA/FDAAA and EMA GVP Module VIII are that of the Sponsor and agrees not to submit any information about this study or its results to a study registry without consulting with the Sponsor.

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

13 SIGNATURES

13.1 Sponsor's Representative

PRINTED NAME	PPD
TITLE	Principal Scientist, Epidemiology
SIGNATURE	
DATE SIGNED	

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

13.2 Investigator

I agree to conduct this study in accordance with the design outlined in this protocol and to abide by all provisions of this protocol (including other project plans and documents referenced from this protocol); changes from the protocol are acceptable only with a mutually agreed upon protocol amendment. I agree to conduct the study in accordance with generally accepted standards of Good Pharmacoepidemiology Practice. I also agree to report all information or data in accordance with the protocol and, in particular, I agree to report any adverse events and product quality complaints as defined in the Safety and Product Quality Complaint Reporting and Related Procedures section. I understand that information that identifies me will be used and disclosed as described in the protocol and the Use and Disclosure of Personal Data notice provided to me, and that such information may be transferred to countries that do not have laws protecting such information. Since the information in this protocol is confidential, I understand that its disclosure to any third parties, other than those involved in approval, supervision, or conduct of the study is prohibited. I will ensure that the necessary precautions are taken to protect such information from loss, inadvertent disclosure, or access by third parties.

PRINTED NAME	PPD
TITLE	Senior Research Scientist, RWE
SIGNATURE	
DATE SIGNED	

PRODUCT: MK-3475	PROTOCOL/AMENDMENT VERSION NO.: 00-v03
REVOPS ID NO: NIS104408 EPIDEMIOLOGY NO.: EP05026.164	NIR DRC APPROVAL DATE: 11/24/2025

13.3 Supplier

I agree to conduct this study in accordance with the design outlined in this protocol and to abide by all provisions of this protocol (including other manuals and documents referenced from this protocol); changes from the protocol are acceptable only with a mutually agreed upon protocol amendment. I agree to conduct the study in accordance with generally accepted standards of Good Pharmacoepidemiology Practice. I also agree to report all information or data in accordance with the protocol and, in particular, I agree to report any adverse events and product quality complaints as defined in the Safety and Product Quality Complaint Reporting and Related Procedures section. I understand that information that identifies me will be used and disclosed as described in the protocol and in order to perform any agreement between myself and the Sponsor, and that such information may be transferred to countries that do not have laws protecting such information. Since the information in this protocol is confidential, I understand that its disclosure to any third parties, other than those involved in approval, supervision, or conduct of the study is prohibited. I will ensure that the necessary precautions are taken to protect such information from loss, inadvertent disclosure, or access by third parties.

PRINTED NAME	PPD on behalf of Flatiron Health, Inc.
TITLE	Senior Research Scientist, RWE
SIGNATURE	
DATE SIGNED	