NON-INTERVENTIONAL (NI) STUDY PROTOCOL

Study information

Title	**D**rug treatment patterns and **E**ffects for metastatic non-small cell **L**ung cancer patients **I**n **NOR**way (DELINOR)
Protocol number	B7461041
Protocol version identifier	Version 1.0
Date	18 March 2022
Research question and objectives	**The primary objective** of the study is to describe the cancer drug treatment sequence and its effects, measured as time to treatment discontinuation and overall survival, for different types of metastatic NSCLC. **The secondary objectives** include the following: • To describe the baseline characteristics of patients diagnosed with metastatic NSCLC, including cancer stage at time of diagnosis (IIIb-IVb), histologic type, comedication and patients' demographics (e.g.: age, gender, place of residence) • To describe the prescription length for selected patient administered anti-cancer drugs and evaluate regional differences in prescription practice
Author	PPD , Oslo Economics): PPD

1. TABLE OF CONTENTS

2. LIST OF ABBREVIATIONS

Abbreviation	Definition
AE	Adverse Event
ALK	Anaplastic Lymphoma Kinase
ATC	Anatomical Therapeutic Chemical
CRN	Cancer Registry of Norway
DRG	Diagnoses Related Groups
EGFR	Epidermal Growth Factor Receptor
GDPR	General Data Protection Regulation
ICD	International Classification of Diseases
ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision
ICMJE	International Committee of Medical Journal Editors
ICPC	International Classification of Primary Care
ID	Identification
Incl.	Including
INSPIRE	Increase pharmaceutical reporting
IT	Information Technology
MED	Minimum Effective Dose
N/A	Not applicable
NDR	Norwegian Drug Registry
NI	Non-interventional
NoMA	Norwegian Medicines Agency
NorPD	Norwegian Prescription Database (Reseptregisteret)
NPR	Norwegian Patient Registry
NSCLC	Non-small-cell lung cancer

OS	Overall survival
PD-L1	Programmed death-ligand 1
PID	Personal identification number
RD	Regression Discontinuity
REC	Regional Committees for Medical and Health Research Ethics
ROS1	ROS proto-oncogene 1
SCLC	Small cell lung cancer
SMS	Short text message
SSN	Social security number
TNM, cTNM	Clinical classification of primary tumors, regional lymph nodes and metastases
QS	Quality system

3. RESPONSIBLE PARTIES

Principal Investigator(s) of the Protocol

Name, degree(s)	Job Title	Affiliation	Address
PPD		Pfizer	Drammensveien 288, 0283 Oslo, Norway
		Oslo Economics	Kronprinsesse Märthas plass 1, 0160 Oslo
		Oslo Economics	Kronprinsesse Märthas plass 1, 0160 Oslo
		Oslo Economics	Kronprinsesse Märthas plass 1, 0160 Oslo
		Oslo Economics	Kronprinsesse Märthas plass 1, 0160 Oslo
		Drammen Hospital	Dronninggata 28, 3004 Drammen
		Norwegian Radium Hospital	Ullernchausseen 70, 0379 Oslo
		St Olavs Hospital	Prinsesse Kristinas gate 3, 7030 Trondheim
		Rikshospitalet	Sognsvannsveien 20, 0372 Oslo

4. ABSTRACT

Title

Drug treatment patterns and effects for metastatic non-small cell lung cancer patients in Norway (DELINOR)

Version 1.0, 18 March 2022, PPD , Oslo Economics).

Rationale and Background

Diagnostics, radiation therapy, and drug therapy for patients with metastatic non-small cell lung cancer (NSCLC) have changed significantly over the last 2 decades. Until recently, reporting of anti-cancer drug treatment from hospitals to the Cancer registry of Norway (CRN) was not done routinely. However, the INSPIRE (INcreaSe PharmaceutIcal REporting) project initiated automatic collection of data on anti-cancer drugs from the hospitals to the CRN. This systematic collection of data presents new opportunities for cancer pharmacoepidemiology research in Norway. It is now possible to study individual patients` complete anti-cancer drug treatment sequence from start to finish.

Research question and objectives

The primary objective of the study is to describe the anti-cancer drug treatment patterns, its impact in terms of overall survival, and time from treatment initiation to discontinuation for subgroups of NSCLC. The secondary objectives include a description of baseline patient characteristics, prescription length, as well as regional differences in prescription practice for patient administered anti-cancer drugs.

Study design

The study is a retrospective cohort study based on secondary data collection extracted from 3 national health registries.

Setting and population

The study will be conducted in Norway and include all patients ≥18 years with histologically confirmed stage IIIb, IIIc, IVa, and IVb NSCLC diagnosed between 01 January 2009, and 31 December 2022 (or latest available year). The study encompasses only patients who were in stage IIIb, IIIc (only non-curatively treated) or IV at the time of the NSCLC diagnosis.

Data sources and variables

The study will be based on data extracted from 3 national health registries: the Cancer Registry of Norway (CRN), the Norwegian Patient Registry (NPR), and the Norwegian Drug Registry (NDR). Data from these registries will be linked at an individual patient level to create a single, unified dataset.

Key variables include diagnosis, cancer stage at diagnosis, date of diagnosis, birth year, type of NSCLC, type of drug treatment, date of treatment initiation and discontinuation, treating hospital and patient characteristics (e.g., age and gender).

Data analysis

The data analyses include descriptive analyses (counts and proportions) and survival analyses. All analyses will be documented in STATA do-files or R-scripts to ensure reproducibility of the results.

Milestones

The end of data collection (data delivery) is expected by 01 June 2023. The final statistical report (study report) is expected to be completed by 31 December 2023. If any milestones in the process are reached earlier or later than expected, the milestone plan will be revised accordingly.

5. AMENDMENTS AND UPDATES

None

6. MILESTONES

Milestone	Planned date
First draft study protocol	20 January 2022
Approved study protocol	18 March 2022
Application to Ethical Committee (Regional Committees for Medical and Health Research Ethics (REC))	22 March 2022
Start of data collection (data application)	01 June 2022
End of data collection (data delivery)	01 March 2023
Draft statistical report (study report)	01 October 2023
Final statistical report (study report)	31 December 2023
Draft manuscript for publication (first manuscript)	01 March 2024
Submission of manuscript to scientific journal (first manuscript)	01 May 2024
Extraction of second dataset (if applicable)	01 June 2023
Data delivery for the second dataset (if applicable)	31 December 2023
Updated draft study report (including (incl.) second dataset) (if applicable)	01 February 2024
Final complete study report (incl. second dataset) (if applicable)	01 March 2024
Draft manuscript for publication (second manuscript) (if applicable)	01 September 2024
Submission of manuscript to scientific journal (second manuscript) (if applicable)	31 December 2024

If any milestone in the process is reached/finalized earlier or later than expected, the milestone plan will be revised accordingly.

7. RATIONALE AND BACKGROUND

Lung cancer (ICD-10 C34) is the disease that claims the most lives of any illness in Norway [1][1]. In 2020, 3,331 patients were diagnosed with lung cancer and 2,168 patients died from it[2]. 9,936 patients were alive by 31 December 2020 after previously being diagnosed with lung cancer (point prevalence). Of the patients diagnosed with lung cancer, approximately 85% are diagnosed with non-small cell lung cancer (NSCLC) and 15% are diagnosed with small cell lung cancer (SCLC) [3, 4]. The main types of NSCLC are adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Most adenocarcinoma patients are diagnosed at a metastatic stage where curative treatment is not possible and 5-year survival is low.

In the last 2 decades, there have been significant changes in diagnostics, radiotherapy, and drug therapy for patients with metastatic NSCLC. All patients with NSCLC in Norway are supposed to have their tumor tested for PD-L1 (Programmed death-ligand 1), and all patients, except those with squamous cell carcinoma, are reflex tested (i.e., immediately tested after diagnosis of NSCLC) for EGFR- (Epidermal Growth Factor Receptor), ALK- (Anaplastic Lymphoma Kinase), and ROS1- (ROS proto-oncogene 1) alterations. EGFR testing of all NSCLC in conjunction with EGFR-directed therapy was implemented nationwide in Norway in 2010, while ALK reflex testing and accompanying therapy was implemented in 2013. Norwegian metastatic NSCLC patients did not receive immunotherapy outside of clinical trials until 2016. As of today (March 2022), multiple anti-cancer drugs targeting ALK+, EGFR+, ROS1+, or PD-L1+ NSCLC are available in Norway. The choice of drug is therefore a less straightforward process than previously. In addition, several new treatment options for NSCLC patients targeting other genetic markers than those in this protocol are currently under evaluation by the Norwegian Medicines Agency (NoMA).

Until recently, there was a lack of routine reporting of lung cancer drug treatment to the Cancer Registry of Norway (CRN). The INSPIRE (INcreaSe PharmaceutIcal Reporting) project, financed by 12 pharmaceutical companies (including Pfizer) and the Norwegian Cancer Society, was launched in order to automatically collect data on cancer drugs from the hospitals systems and report these to the CRN. The improved collection of data presents new opportunities for cancer pharmacoepidemiologic research in Norway.

INSPIRE allows for the study of individual patients` complete treatment sequence from start to finish. It enables evaluation of how recent changes in the drug treatment of metastatic NSCLC, including the introduction of the first targeted therapies, have changed treatment practice in a real world setting and whether these changes have improved patient survival rates. A paper published in 2018 concluded that Norwegian patients with metastatic lung cancer showed little improvement in relative survival during the 2000-2016 period [5]. The study included all metastatic lung cancer patients, regardless of their eligibility for targeted therapy or immunotherapy. Literature searches have indicated that no RWE studies have investigated treatment outcome in this patient group by type of anti-cancer drug treatment in

[1] Trachea (ICD-10 C33) is included in the public available numbers for incidence, prevalence, and mortality reported in the Cancer Registry of Norway.

Norway. Therefore, subgroup analyses of the metastatic NSCLC patients treated with targeted anti-cancer therapies or immunotherapies will be performed in this study to better understand the impact of changes in the diagnostics and anti-cancer drug treatment of this patient group.

An improved understanding of National treatment patterns could potentially improve quality and equality of lung cancer treatment in Norway. It would also help promote evaluations of possible measures for better patient outcomes. Aggregate analyses of treatment patterns for NSCLC are presented in the annual report from Norwegian Lung Cancer Registry [6]. However, this study will include more detailed analyses of the anti-cancer drug treatment of subgroups, as well as changes over time to explore potential regional differences. The current real-world evidence of the effectiveness of the new targeted therapies and immunotherapy in Norway is also lacking. It is unclear how long duration of treatment is for anti-cancer drugs in a real-world setting in Norway and whether the new therapies (both targeted and immunotherapy) have improved patient survival in patients with metastatic NSCLC.

This study aims to use Norwegian registry data to analyze anti-cancer drug treatment patterns (drug sequence), time on treatment, and survival for NSCLC patients in the metastatic setting to better understand the impact of the introduction of targeted therapies and immunotherapy.

8. RESEARCH QUESTION AND OBJECTIVES

Primary objectives
The primary objective of the study is to describe the cancer drug treatment sequence and its effects, measured as time to treatment discontinuation and overall survival, for different types of metastatic NSCLC.

The primary objective includes the following:

- **A1**: To describe anti-cancer drug treatment patterns for metastatic NSCLC patients in Norway during the period 2009 to 2022
- **A2**: To describe the time from treatment initiation to discontinuation (time on treatment) for anti-cancer drug treatment for metastatic NSCLC patients in Norway
- **A3**: To evaluate if the introduction of targeted therapies and immunotherapies has influenced overall survival for metastatic NSCLC patients in Norway

Secondary objectives
The secondary objectives include the following:

- **B1**: To describe the baseline characteristics of patients diagnosed with metastatic NSCLC, including the stage at time of diagnosis (IIIb, IIIc, IVa, and IVb), histologic type, comedication, and other patient demographics (e.g.: age, gender, place of residence)
- **B2**: To describe the prescription length for selected patient administered anti-cancer drugs and evaluate regional differences in prescription practice

9. RESEARCH METHODS

9.1. Study design

The study will be a retrospective cohort study based on secondary data extracted from 3 National Health registries:

- the Cancer Registry of Norway (CRN)
- the Norwegian Drug Registry (NDR)
- the Norwegian Patient Registry (NPR)

Data from the 3 registries will be linked at the individual patient level to create a single, unified dataset. In addition to their diagnosis, the dataset will include information on patient characteristics and all available information regarding treatment (including drugs, radiotherapy, and surgery) received by these patients during the period.

9.2. Setting

The study will be conducted in Norway and include all patients ≥18 years with histologically confirmed metastatic NSCLC who received their first diagnosis between 01 January 2009 and the year of which the most recent data are available (hereby referred to as latest available year).

At the time of developing this research protocol the study is expected to include data from 01 January 2009 to 31 December 2022. If more recent data are available at the time of data extraction, the inclusion period will be extended to include more updated data. To extend the study period, there will be two data extractions. The second data extraction will be an updated version of the dataset from the first extraction but also include the most recent data at the time of data extraction (estimated 2023). The study aims to include NSCLC patients with stage IIIb, IIIc, IVa, or IVb at the time of diagnosis[2]. NSCLC include patients with confirmed adenocarcinoma, squamous cell carcinoma, and non-classified NSCLC. Patients who were diagnosed with stage I, II or IIIa, but later progressed into a more advanced stage, will be excluded from the study. Some patients diagnosed with stage IIIb and IIIc are potentially curable and receive a different treatment regime than the non-curable metastatic patients. Thus, these will be excluded from the analyses by excluding those who have received ``curative`` radiation, ie >50Gy. The inclusion and exclusion criteria for the data extraction are presented in the following.

9.2.1. Inclusion criteria

Patients must meet all of the following inclusion criteria to be eligible for inclusion in the study:

[2] Patients diagnosed with stage NSCLC stage III are often referred to as advanced rather than metastatic. In this protocol we use the term metastatic when referring to patients with NSCLC stage IIIb, IIIc, IVa, or IVb.

1. Patients ≥ 18 years old with histologically confirmed stage IIIb, IIIc, IVa, or IVb NSCLC at the time of diagnosis

2. Received their first NSCLC diagnosis (stage IIIb, IIIc, IVa, or IVb) between 01 January 2009, and latest available year

9.2.2. Exclusion criteria

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

1. Patients ≥ 18 years old with histologically confirmed stage IIIb, IIIc, IVa, or IVb NSCLC at the time of diagnosis who received radiation therapy with curative intent, defined as a radiation dose larger than 50gy

2. If the data include patients diagnosed after 2021, IIIb patients will be excluded due to changes in the guideline. The updated guidelines for 2022 recommends that patients diagnosed with stage IIIb who have surgery with the intent to cure should not have ``curative`` radiation (i.e.: it is not possible to exclude these from the population from 2022).

9.3. Study populations and subgroups

The analyses will be performed on the whole population and four subpopulations. Patients with histologically confirmed metastatic NSCLC will be grouped based on their type of NSCLC into the following groups (classification will be based on the first recorded mutation observed in the data):

- ALK+
- EGFR+
- PD-L1+
- ROS1+

Data collection will capture all patients irrespective of the results of the tested biomarkers. Patients who were not tested or tested negative for all biomarkers will be assessed together. The main study groups are not mutually exclusive, and patients may be both ALK/EGFR+ and PD-L1+. Cases where patients tested positive for multiple biomarkers will be assessed individually and assigned to one of the main study groups. ROS1+ patients will be defined as those with a recorded ROS1+ mutation or patients who received Xalkori who were not confirmed as ALK+. The result of ROS1 tests were not implemented until 2021, thus we will use Xalkori treatment in patients that are not ALK+ as proxy.

Estimated sample sizes for the 3 subgroups ALK+, EGFR+ and PD-L1+ are presented in Table 1 below. Estimated sample size is based on available data from INSPIRE [7]. Further, it is assumed that 50% of all incident lung cancer patients are metastatic non-small cell [6]. For the ROS1+ group, there is no previous data available because results from testing are not

registered in CRN. ROS1+ patients will instead be identified based on medication use observed in data. Therefore, sample size for this subgroup cannot be estimated.

Table 1: Estimated Number of Patients Included in Each Study Group

NSCLC subtype:	Number of patients:
ALK+	240
EGFR+	465
PD-L1+	4140

Calculations based on available data from INSPIRE [7].

Study periods
For research questions A1 (treatment patterns) and A3 (changes in overall survival (OS) over time), the analyses will be divided into different study periods. These study periods will be determined based on the time of introduction (i.e.: public reimbursement) of the different targeted therapies and immunotherapies in Norway, as well as the completeness of the recorded targets (ALK+, EGFR+, ROS1+, PD-L1+) in the data.

Subgroups
Additional relevant subgroups for exploratory analyses include:

- Age
- Gender
- Region (regional health authority and health authority)
- Stage at diagnosis (stage IIIb, IIIc, IVa, and IVb)
- Histologic type (adenocarcinoma, squamous cell carcinoma, and non-classified NSCLC)
- Groups based on median survival (e.g.: quartiles)
- Year of diagnosis

9.4. Potential further limitation of study population (to be decided)

The aim of the study is to perform the analyses on patients with NSCLC stage IIIb, IIIc, IVa, or IVb at the time of diagnosis. However, there are some challenges related to identifying patients diagnosed with these stages in the CRN. These challenges and potential measures to handle them are described as follows:

The CRN classifies stage[8] as:
- **Localized stage**: All cases where the tumor is confined to the primary organ
- **Regional stage**: All cases where the tumor has invaded neighboring tissue outside of the primary organ or metastasized to regional lymph nodes

- **Distant stage**: All cases where the tumor has metastasized to other organs or distant lymph nodes
- **Unknown**: All cases where the primary origin of the tumor is not known and cases with insufficient information to set stage

For NSCLC, the localized stage is equivalent to stage I, regional stage to stage II and III and distant stage to stage IV in the clinical classification of primary tumors, regional lymph nodes and metastases (cTNM). To separate stage II from III, and stage IIIa from IIIb/c it will be necessary to use information from the investigation report (in Norwegian: "utredningsmelding") for each patient. The CRN has information on cancer stages, which is based on cTNM for NSCLC from 2014. However, the quality of the data between 2014 and 2017 is uncertain and needs to be assessed when the data is analyzed. The coverage is assumed to be high from 2017 and onwards (80-90%).

As the completeness of cTNM registration, and thus the possibility to identify patients with stage IIIb/c, is uncertain, 2 different approaches are planned to define the study period and study population (Table 2). The final choice of study period and study population (alternative 1 and 2 described in Table 2) will be made after assessing the completeness of the data and will be done as a part of the project. "Alternative 1" will be the preferred option. However, this alternative is conditional on a high coverage rate of cTNM in CRN in the period 2014-2017. If preliminary analyses of coverage indicate that the registration is poor, the study will be limited to include patients only with stage IV at the time of diagnosis ("Alternative 2").

Table 2: Alternatives to Handle Missing Data on Stage Classification in CRN

	Alternative 1	**Alternative 2**
Study period	2014 – 2021*	2009 – 2021*
Study population	NSCLC stage IIIb, IIIc and IV	NSCLC stage IV(distant metastasis)
Classification system used to identify patients	cTNM	CRN classification

*Or latest available year

9.5. Outcomes and Endpoints

A description of outcomes and endpoints are provided below.

Primary objectives

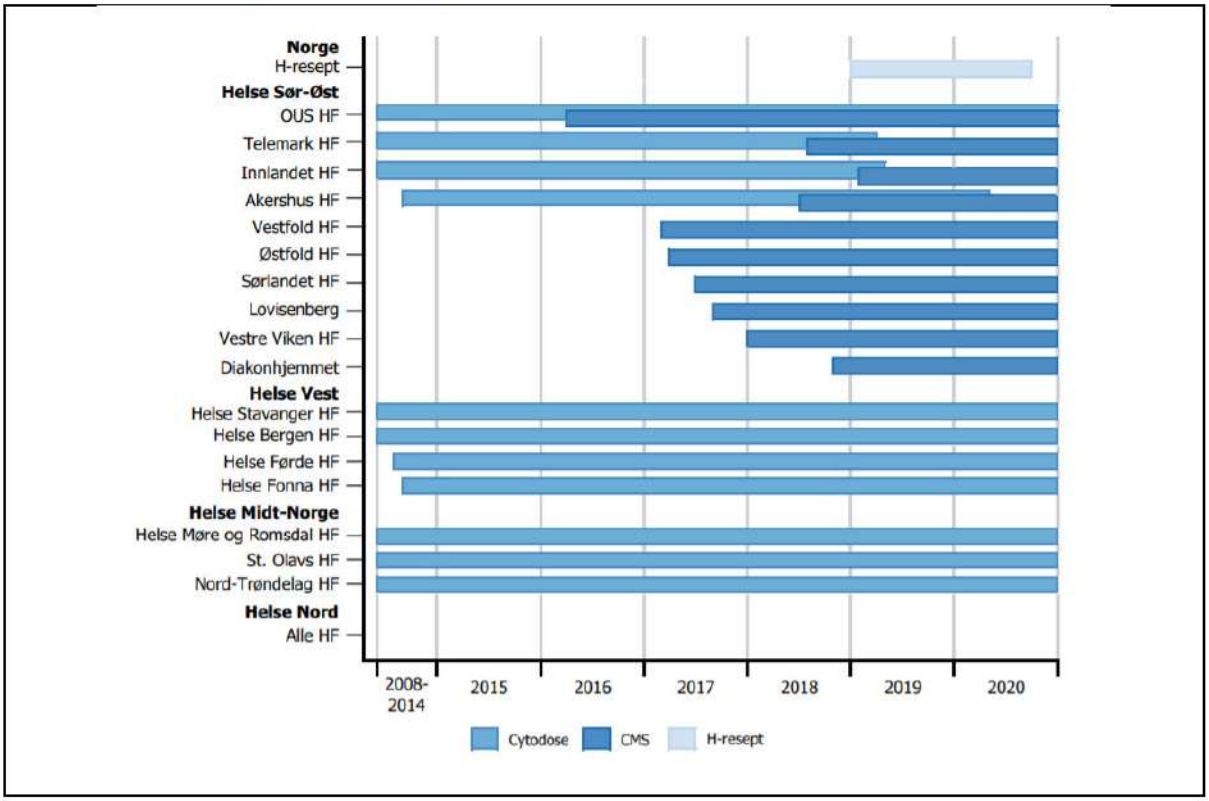
Objective A1: Anti-cancer drug treatment patterns for metastatic NSCLC patients

Treatment patterns will be described as the number and proportion of patients receiving different drug treatments (i.e., treatment lines). Additionally, the number and proportion of patients who do not receive any anti-cancer drug treatment will be described. Treatments will include all anti-cancer drug treatment received by the patients. Results will be presented using Sankey diagrams and tables (see table shells in section 9.14).

Analysis of treatment patterns will be performed for the 4 main study populations (PD-L1+, ALK+, EGFR+, and ROS1+) and by separate study periods (defined for each study group as a part of the project). Additional exploratory subgroup analysis may also be relevant and will be considered as a part of the study (for examples of relevant subgroups, see subsection 9.3).

Drug treatment will be identified using the recorded Anatomical Therapeutic Chemical (ATC) -codes (5-level digits) in the 3 registries. With the INSPIRE-project, the CRN now holds data on cancer drug treatment for 3 of the 4 regional health authorities in Norway (excluding Helse-Nord). Except for patient-administered drugs financed by the hospitals (in Norwegian: "H-resept"), the information on drug treatment is in large parts complete from 2008 and onwards (see Figure 1). NDR holds data on all patient administered drug treatments (tablets and subcutaneous treatments), including drugs on "H-resept", from 2008. Data from CRN and NDR will be supplemented with data from NPR to capture in-hospital drug use, including chemotherapy and immunotherapy. From NPR, cancer drug treatment will by identified using reported ATC-codes and registered outpatient medical treatment of lung cancer (DRG 856D).

Figure 1: Data Capture for Drug Cancer Treatment



Source: Cancer Registry of Norway (2021) [7]

Objective A2: time to treatment discontinuation

The time to treatment discontinuation (i.e., treatment length) of the different anti-cancer drugs will be estimated by calculating their average time on treatment using survival analyses

(drug survival). Time to treatment discontinuation will be evaluated for targeted therapies and immunotherapies (Table 3). The analyses will be performed separately for each of the 4 main study populations (PD-L1+, ALK+, EGFR+, and ROS1+).

The first step in the analyses will be to construct the observed treatment trials in the collected data. Each patient that receives a specific treatment during the study period will be considered in this analysis. All treatment trials are observed until the end of data collection (31 December 2021 or latest available year). Based on the first observed drug treatment (prescription or infusion) for a drug for each patient, a treatment trial of uninterrupted treatment will be constructed. This will be done based on the dose retrieved by these patients (either from the pharmacy or in hospitals) and the minimum effective dose (MED) for each drug in a dataset. The MEDs for the different drugs will be defined in collaboration with the clinicians involved in the study. Treatment discontinuation will be defined at the point where a treatment break of more than 30 days is observed (30 days grace period), the patient is recorded as deceased, or if a new anti-cancer drug treatment is initiated.

Time to treatment discontinuation will be measured as median treatment length and the proportion of patients still on treatment after 1, 3, 6, 9, 12, …, and 36 months (i.e.: after 1 month and every 3 months up to 3 years). Patients will be censored when lost to follow-up. This may occur either when they are recorded as dead, move abroad or are still on treatment at the end of the data period. Immunotherapies will only be evaluated up to 2 years as it is believed that the responders to immunotherapies have long lasting response and that it is unnecessary to continue drug treatment after 2 years. Results will be presented both as tables and as Kaplan-Meier curves (see table shells in chapter section 9.14).

The base case analysis will be restricted to the first drug received by each patient (1st line treatment). Additional analyses evaluating time to treatment discontinuation for drugs in later treatment lines, or in selected subgroups, may also be investigated in exploratory analyses.

Table 3: Overview of Cancer Drugs for Metastatic NSCLC

Study population	*Cancer drugs*
ALK+	• crizotinib • ceritinib • alectinib • brigatinib • lorlatinib • atezolizumab • bevacizumab • paklitaksel • karboplatin • docetaksel
EGFR+	• erlotinib • gefitinib • afatinib

	• osimertinib • dacomitinib • atezolizumab • bevacizumab • paklitaksel • karboplatin • docetaksel
PD-L1+	• pembrolizumab • nivolumab • atezolizumab • paklitaksel • karboplatin • docetaksel
ROS1+	• crizotinib • entrektinib • lorlatinib • atezolizumab • bevacizumab • paklitaksel • karboplatin • docetaksel

Objective A3: changes in overall survival

The aim of this objective is to evaluate if the introduction of targeted therapies and immunotherapy has influenced overall survival for metastatic NSCLC patients in Norway.

The overall survival will be estimated by calculating the average time from diagnosis to death using survival analysis (Kaplan-Meier estimator). Overall survival will be evaluated for patient cohorts diagnosed within different study periods and changes over time will be assessed. The analyses will be performed separately for each of the 4 main study populations (PD-L1+, ALK+, EGFR+, ROS1+).

As with the analysis of time to treatment discontinuation, the first step in the analyses will be to construct each patient survival time in the collected data (i.e.: time from diagnosis to death/end of data collection). Each patient that is diagnosed with NSCLC during the study period will be considered in this analysis. All patients are observed until death or the end of data collection (31 December 2021 or latest available year). Patients that are alive by the end of the observation period will be censored.

Overall survival will be measured as median survival and the proportion of patients alive after 3, 6, 9, 12, …., and 60 months (i.e.: every 3 months up to 5 years). Results will be presented both as tables and as Kaplan-Meier curves (see table shells in chapter section 9.14). Log-rank tests will be conducted to test for differences in survival by time periods.

The intention with this research objective will be to describe changes in survival over time for different subgroups based on descriptive statistical methods. However, changes in survival over time may be caused by several factors other than the introduction of new treatments, including changes in diagnostics. Several different statistical methods will be applied to better understand the relationship between the introduction of targeted therapies/immunotherapies and changes in overall survival. Exploratory analyses using more complex statistical methods will be performed as described in the next paragraph. The final choice of subgroups to be analyzed, assumptions and time periods will be defined as a part of the study depending on suitability and data quality.

A Cox regression model can be fitted to the data to assess whether there is a causal relationship between the introduction of targeted therapies/immunotherapy and changes in overall survival. The model will include variables for different treatments as regressors. The regression model can be used to quantity the effect of different treatments on a survival outcome (e.g., hazard ratio) if key model assumptions, such as proportional hazards, are met. A Cox regression model will be able to control for different illness- and patient-related characteristics that are observable in the data.

As in many regression models, there are some variables that cannot be control for due to unobservability or lack of data. This may lead to biased estimates of treatment effects. This challenge can be mitigated by the use of a regression discontinuity (RD) design [9]. Since different treatments are introduced at a specific point in time (time of public reimbursement), type of treatment is assigned deterministically (all patients receive the new treatment after reimbursement) or probabilistically (a proportion of the patients receive the new treatment after reimbursement) by the continuous variable time. Assuming that there are no other significant changes influencing survival at the time of reimbursement, the survival of the patients before and after the introduction of a new treatment can be compared and a causal inference can be drawn on the effect of new treatments on overall survival. This method requires that the group of patients diagnosed before the introduction of a specific treatment are comparable (i.e., on average similar in other covariates) to the group of patients who received the new treatment. This is essentially an untestable assumption, but specification checks can increase confidence that the assumption holds. One important specification check is to test for discontinuities in other average covariates in the data. For metastatic NSCLC patients, several treatments/patient cohorts may be applicable for this method, for example the introduction of pembrolizumab and nivolumab for the PD-L1+ group in 2016/2017.

Secondary objectives

Objective B1: Baseline characteristics

Baseline characteristics will be described as number, and proportion, of patients according to the following definitions:

Age
Patient age will be defined as the age at diagnosis and divided into the following groups:

- All ages
- 75+ years
- 65 – 74 years
- 55 – 64 years
- 45 – 54 years
- <45 years

Gender
Patients will be classified by their recorded gender at the time of diagnosis and divided into the following groups:

- Male
- Female
- Other/unknown

Region
Patients will be classified by geographic region based on their place of residence at the time of diagnosis based on health region:

- South Eastern (Helse Sør-Øst)
- Western (Helse Vest)
- Central (Helse Midt)
- Northern (Helse Nord)

Additionally, analyses will be performed on hospital level (i.e.: health authority level).

Stage at diagnosis
Patients will be classified based on cancer stage at diagnosis:

- IIIb
- IIIc
- IVa
- IVb

Type of NSCLC
Patients will be classified by the type of NSCLC:

- ALK+
- EGFR+
- PD-L1+
- ROS1+
- Wild type/no mutation

Histologic type
Patients will be classified by histologic type based on their confirmed type at diagnosis:

- Adenocarcinoma

- Squamous cell carcinoma
- Non-classified NSCLC

Comedication
Comedication will be described as the number, and proportion, of patients receiving selected patient administered drugs based on data from the NDR. Examples of drugs of special interest are:

- Statins (C10AA)
- Anticoagulants (B01)
- Proton pump inhibitors (A02BC)

Objective B2: Prescription length and regional differences in prescription practice

Patient administered drugs can be prescribed for different lengths depending on the number of pills or syringes on the prescription. The length of a prescription will influence potential drug waste, as many patients stop their treatment before it is completed, either due to lack of effect, toxicity, or death. If a patient is prescribed drugs for a long period of time, there is an increased probability that a proportion of the drugs redeemed are never used. At the same time, a longer prescription may reduce a patient's time- and travel costs related to filling the prescription.

The aim of this research question is to:

- describe the prescription length for selected patient administered anti-cancer drugs for patients diagnosed with metastatic NSCLC in Norway
- and to evaluate whether there are any regional differences in prescription practice.

Prescription length will be evaluated by calculating the proportion of prescription lasting for 1, 2, 3, or more than 3 months for selected treatments, by geographic region and prescription number per patient (i.e.: first prescription, second prescription, third prescription etc.). The analysis will be restricted to one or a few anti-cancer drugs used for ALK+, ROS1+, or EGFR+ NSCLC (presented in Table 3).

Drug waste (and potential public savings of changing prescription practice) can be estimated by identifying markers for when a patient has stopped his or her treatment. Examples of such markers are the initiation of a new treatment regiment, treatment of a specific adverse event or death.

For this research question, the analyses will be based on the drugs delivered to patients (number of packs and package size collected by the patient) and the expected time the drugs will last to evaluate the length of each prescription. The drugs collected by the patients will be used as a proxy for the length of the prescription.

9.6. Variables

The table below (Table 4) provides a list of key variables from the national registries where individual data will be collected with a personal identification number (PID) attached. For all

dates (date of diagnosis, treatment date, death date), the year and days will be used since an unknown reference date (exact dates will not be included due to privacy considerations).

Table 4: List of Key Variables

Variable	Role	Data source(s)	Level of detail
Kommunenummer (Municipality code)	Grouping variable / covariate	Cancer Registry of Norway	
Histologi (Histology)	Grouping variable	Cancer Registry of Norway	
Statusdato (Date of status)	Identification variable	Cancer Registry of Norway	Days from unknown reference date
Fødselsår (Year of birth)	Grouping variable / covariate	Cancer Registry of Norway	
Metastase (Metastases)	Grouping variable	Cancer Registry of Norway	
Diagnosens sikkerhet	Quality control	Cancer Registry of Norway	
Kategorisering av underliggende dødsårsak (Categorization of cause of death)	Control variable	Cancer Registry of Norway	Cancer deaths or other cause of death
Morfologi (Morphology)	Grouping variable	Cancer Registry of Norway	
Basis for diagnose (Basis for diagnosis)	Grouping variable	Cancer Registry of Norway	
Kjønn (Gender)	Grouping variable / covariate	Cancer Registry of Norway	
Diagnoseår (Diagnosis year)	Measure/outcome variable	Cancer Registry of Norway	
Statusår (Status year)	Measure/outcome variable	Cancer Registry of Norway	
Kjemoterapi (Chemotherapy)	Outcome variable	Cancer Registry of Norway	
Strålebehandling (Radiation)	Outcome variable	Cancer Registry of Norway	Including dose-level (dosenivå)
Diagnosekode for underliggende dødsårsak (Diagnosis code for cause of death)	Control variable	Cancer Registry of Norway	Cancer deaths or other cause of death
Kirurgi (Surgery)	Outcome variable	Cancer Registry of Norway	
Morfologisk verifisert (Morpohologically verified)	Quality control	Cancer Registry of Norway	
Sykehus (Hospital)	Grouping variable / covariate	Cancer Registry of Norway	
Stadium (Stage)	Grouping variable / covariate	Cancer Registry of Norway	
Personstatus (Patient status)	Grouping variable / covariate	Cancer Registry of Norway	Alive, dead or loss of follow-up
Diagnosedato (Date of diagnosis)	Measure/outcome variable	Cancer Registry of Norway	Days from unknown reference date
ALK	Grouping variable / covariate	Lung Cancer Registry	Positive, negative, not specified
ATC-kode (ATC code)	Outcome variable	Lung Cancer Registry	5-digit level
Behandlingssyklus (Treatment cycle)	Grouping variable / covariate	Lung Cancer Registry	
Behandlingssyklusdato (Treatment cycle date)	Outcome variable	Lung Cancer Registry	Days from unknown reference date
Dato for behandlingsbeslutning (Treatment decision date)	Outcome variable	Lung Cancer Registry	Days from unknown reference date
Dose virkestoff (Dose chemical substance)	Measure/outcome variable	Lung Cancer Registry	
Doseenhet (Unit for chemical substance)	Measure/outcome variable	Lung Cancer Registry	
EGFR	Grouping variable / covariate	Lung Cancer Registry	Positive, negative,

			inconclusive, unclearly described, reply forwarded later, not specified
Inklusjon i klinisk studie (Inclusion in clinical study)	Control variable	Lung Cancer Registry	
Operasjon av primærtumor (Surgery of primary tumor)	Descriptive	Lung Cancer Registry	
Operasjonsdato for primærtumor (Surgery date for primary tumor)	Descriptive	Lung Cancer Registry	Days from unknown reference date
Operasjonsår (Year of surgery)	Descriptive	Lung Cancer Registry	
Opererende sykehus (Hospital performing the surgery)	Descriptive	Lung Cancer Registry	
PD-L1	Grouping variable / covariate	Lung Cancer Registry	Unknown, negative, positive (unknown %), 1-49% or >= 50%
Preparattype (type of drug)	Measure/outcome variable	Lung Cancer Registry	
ROS1	Grouping variable / covariate	Lung Cancer Registry	Test performed/ not performed
Startdato for strålebehandling (Start date for radiation)	Measure/outcome variable	Lung Cancer Registry	Days from unknown reference date
Strålebehandling (Radiation)	Descriptive	Lung Cancer Registry	Including dose-level (dosenivå)
Strålebehandling med kurativ intensjon (Radiation with curative intent)	Descriptive	Lung Cancer Registry	
Type medikamentell kreftbehandling (Type of cancer drug used)	Measure/outcome variable	Lung Cancer Registry	
Utleveringsdato for H-resept (Dispensing date for hospital financed prescription)	Measure/outcome variable	Lung Cancer Registry	Days from unknown reference date
Utredende sykehus (Investigative hospital)	Descriptive	Lung Cancer Registry	
Varenavn H-resept (Name of drug from hospital financed prescription)	Measure/outcome variable	Lung Cancer Registry	
Virkestoff (Chemical substance)	Measure/outcome variable	Lung Cancer Registry	
Virkestoff, rapportert fra sykehus (Hospital reported chemical substance)	Measure/outcome variable	Lung Cancer Registry	
cM	Grouping variable / covariate	Lung Cancer Registry	
cN	Grouping variable / covariate	Lung Cancer Registry	
cT	Grouping variable / covariate	Lung Cancer Registry	
pM	Grouping variable / covariate	Lung Cancer Registry	
pN	Grouping variable / covariate	Lung Cancer Registry	
pT	Grouping variable / covariate	Lung Cancer Registry	
ypM	Grouping variable / covariate	Lung Cancer Registry	
ypN	Grouping variable / covariate	Lung Cancer Registry	
ypT	Grouping variable / covariate	Lung Cancer Registry	
Pasientens løpenummer (Patient ID number)	Identification	Norwegian Drug Registry	

Vare pakningsenhet (Unit for package size)	Measure/outcome variable	Norwegian Drug Registry	
Ordinasjonens antall definerte døgndoser (Defined daily dose of dispensing)	Measure/outcome variable	Norwegian Drug Registry	
Ordinasjon løpenummer (Dispensing ID)	Identification	Norwegian Drug Registry	
ATC-kode (ATC-code)	Measure/outcome variable	Norwegian Drug Registry	
Utleveringsdato (Dispensing date)	Identification	Norwegian Drug Registry	Days from unknown reference date
Varenavn (Name of drug)	Measure/outcome variable	Norwegian Drug Registry	
Definert døgndose Verdi (Defined daily dose value)	Descriptive	Norwegian Drug Registry	
Utleveringsår (Year of dispensing)	Measure/outcome variable	Norwegian Drug Registry	
Forskriver løpenummer (Prescriber ID)	Grouping variable / covariate	Norwegian Drug Registry	
Vare pakningsstyrke (Package dose)	Measure/outcome variable	Norwegian Drug Registry	
Definert døgndose måleenhet (Unit for defined daily dose)	Measure/outcome variable	Norwegian Drug Registry	
Forskriver profesjon (Work title of the prescriber)	Grouping variable / covariate	Norwegian Drug Registry	
Vare pakningsstørrelse (Package size)	Measure/outcome variable	Norwegian Drug Registry	
Ordinasjonens refusjonskode ICPC (Dispensing reimbursement code from ICPC)	Grouping variable / covariate	Norwegian Drug Registry	
Forskriver spesialitet (Speciality of the subscriber)	Grouping variable / covariate	Norwegian Drug Registry	
Ordinasjonens utsalgspris (Dispensing sale price)	Measure/outcome variable	Norwegian Drug Registry	
Ordinasjonens refusjonskode ICD (Dispensing reimbursement code from ICD)	Grouping variable / covariate	Norwegian Drug Registry	
Ordinasjonens antall pakninger (No. of packages redeemed)	Measure/outcome variable	Norwegian Drug Registry	
Utleveringsmåned (Month of dispensing)	Measure/outcome variable	Norwegian Drug Registry	Days from unknown reference date
Liggetid (Length of hospital stay)	Identification	Norwegian Patient Registry	
Prosedyrer (Procedure)	Measure/outcome variable	Norwegian Patient Registry	
DRG-type (Type of DRG)	Identification	Norwegian Patient Registry	
Tidspunkt for start av tjenesten (Start time for hospital service)	Identification	Norwegian Patient Registry	Days from unknown reference date
Omsorgsnivå (Level of care)	Identification	Norwegian Patient Registry	
Spesifikk DRG (Specific DRG)	Identification	Norwegian Patient Registry	
Cytostatika (Cytotoxic agents)	Measure/outcome variable	Norwegian Patient Registry	
DRG	Identification	Norwegian Patient Registry	

ATC-kode (ATC-code)	Measure/outcome variable	Norwegian Patient Registry	
Kontakttype (Type of contact)	Identification	Norwegian Patient Registry	
Aktivitetstype (Type of activity)	Identification	Norwegian Patient Registry	
Hoveddiagnosegruppe (Group of principal diagnosis)	Identification	Norwegian Patient Registry	
Bidiagnosegruppe (Group of secondary diagnosis)	Identification	Norwegian Patient Registry	
Omsorgsnivå (Level of care)	Identification	Norwegian Patient Registry	
DRG-vekt (DRG weight)	Identification	Norwegian Patient Registry	

9.7. Data sources

The study will be based on data extracted from 3 National Health registries. Data from these registries will be linked at the individual patient level to create a single, unified dataset. The following registries will be included in the study:

The Cancer Registry of Norway (CRN)
The Cancer Registry of Norway (CRN) contains health information about all individuals in Norway who have been diagnosed with cancer. All medical doctors in Norway are required by law to notify CRN of any new cancer case and the data are collected from several sources, including physicians, hospitals, laboratories, and by linkage with NPR. Up to now, the CRN has had information on surgical treatments and radiation therapy only. With the recent INSPIRE-project, the CRN now also has data on cancer drug treatment from most hospitals in Norway.

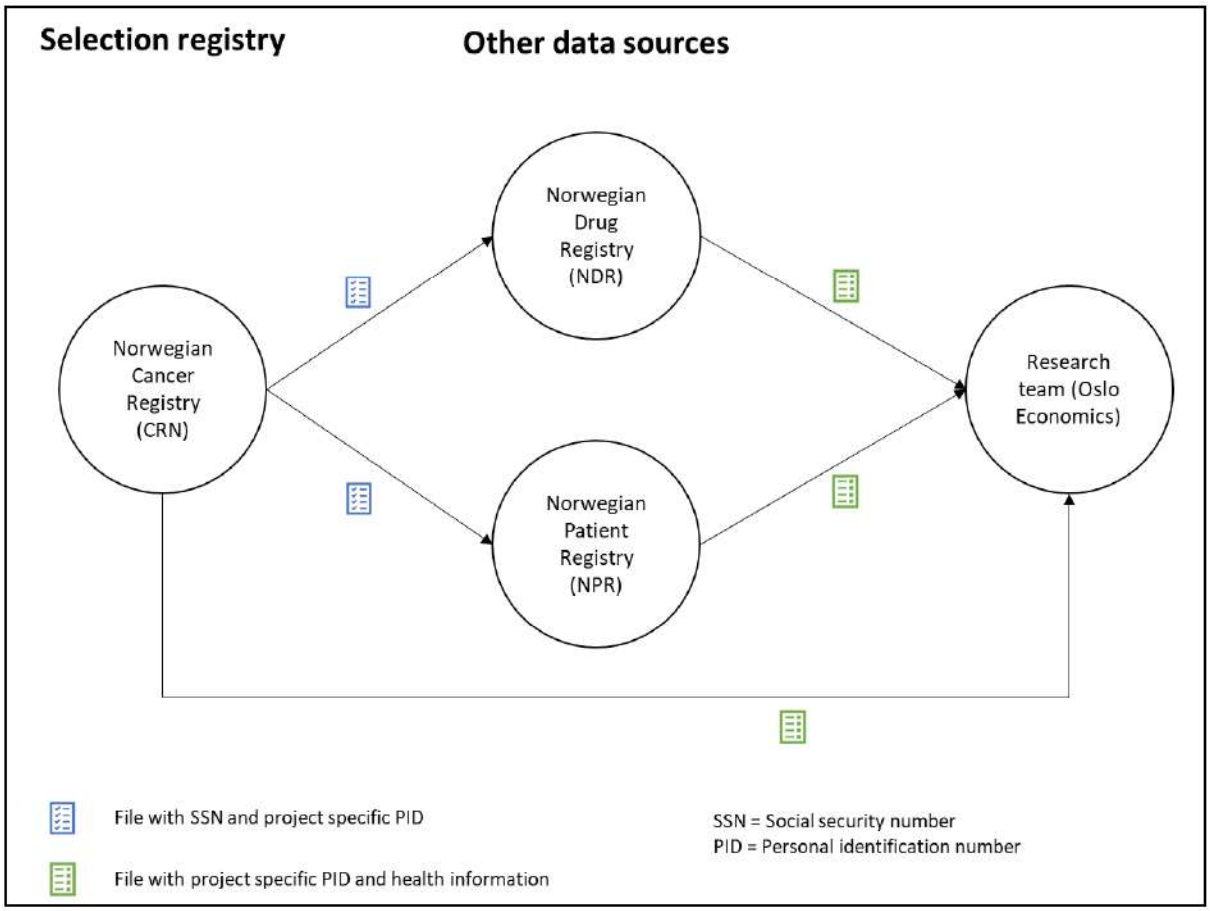
The Norwegian Drug Registry (NDR)
The Norwegian Drug Registry (Legemiddelregisteret, NDR) holds individual patient-level data on all drugs prescribed (reimbursed or not) and dispensed to patients living outside institutions in Norway. All Norwegian pharmacies are required by law to send electronic data to NDR, and the completeness of the data is considered to be close to 100 percent. NDR is under development and is expected to be up and running in 2022. NDR replaces the former Norwegian Prescription Database (Reseptregisteret, NorPD) and will include the same data as NorPD. By linking NDR data to data from CRN we will have data on all drug treatments given outside of hospitals (patient-administered drugs).

The Norwegian Patient Registry (NPR)
The Norwegian Patient Registry (NPR) holds data on all treatment episodes in publicly financed hospitals. As cancer treatment in private hospitals is negligible in Norway, NPR covers virtually all cancer-related episodes of care. From NPR we will be able to identify the use of chemotherapy not reported to CRN.

Linking of data
Linking of datasets will be done by distributed linking, which is the most efficient way of linking data from several sources. The research team at Oslo Economics will continually

receive data from the various sources as data applications are being processed and perform the actual compilation of the data itself based on project-specific identification numbers. Datasets from national registers will be linked according to the process described below Figure 2.

Figure 2: Linking of Registries and Dataflow



9.8. Study size

The primary objective of the study is to improve the collective understanding of the actual treatment patterns for metastatic NSCLC patients in a real-world setting. The Norwegian health care registries cover most Norwegian cancer patients; thus, the data will reflect the current treatment practice in a national setting. Sample size (power) calculations are not applicable since there is no hypothesis testing. Additionally, the study is carried out as a nationwide, population-based study at the national population level utilizing all obtainable data.

The study will include all patients who satisfy the inclusion and exclusion criteria described under subsection 9.2. Based on data on stage distribution at diagnosis presented in the annual

report from Norwegian Lung Cancer Registry [6] and the relative proportion of NSCLC/ SCLC the expected study size is estimated to be 20,605 using data back to 2009 (Table 5).

Table 5: Estimated Study Size

Proportion / Incidence	Number / %	Source / calculation
Annual incidence lung cancer stage III*	538	Cancer Registry of Norway [6] (Table 3.1)
Annual incidence lung cancer stage IV	1,327	Cancer Registry of Norway [6] (Table 3.1)
Proportion NSCLC	85%	[3] and [4]
Annual incidence metastatic NSCLC	1,585	(538+1,327) x 0,85
Expected study size (2009-2021)	**20,605**	**1,585 x 13 years**

**Note: stage IIIc and IIIB cannot be separated from IIIa. The expected study size presented in the table is probably an overestimation and should be regarded as an upper bound.*

9.9. Data Management

The individual registries will provide data on a memory stick, in UNICODE 8 text format with a semi-colon as the delimiter. The data is stored in a password-protected file. The individual registry provides the password by short text message (SMS) to the project lead defined in the data application.

The responsible Partner at Oslo Economics (PPD) will be responsible for data security. PPD will ensure that all data and associated files are stored in a password-protected environment throughout the project. Moreover, the researcher ensures that data archiving is in line with the national and the sponsor regulations; and that deletion is in accordance with regulations stipulated by the individual registry. Researchers from Pfizer (or clinicians involved in the study) will not have direct access to the individual patient level data.

Specifically, data will be stored at the Oslo Economics' secure server, specifically designed for securing and processing sensitive data. The equipment is password protected with 2-factor authentication. It uses encrypted storage, encrypted backups, and encrypted communications.

The main data management steps are:

1. Receipt of data and password from the individual registry
2. Importing data files into IT (information technology) analytical platform
3. Opening password-protected data files
4. Importing text data into a statistical software package

5. Storing the imported text data as "base" files in a password-protected environment according to the regulation governing the data provided by the individual registry. The base files stay unaltered throughout the project.
6. Creating necessary data work files and naming appropriately
7. Linking data using project specific personal identification numbers
8. Cleaning data
9. Check data integrity by random selection
10. Creating necessary variables to support analysis according to project protocol
11. Analysis according to project protocol
12. Check variable creation and programming by second researcher/programmer
13. Test and re-test analysis to ensure reliability
14. Output tables in an appropriate format

The researcher records and documents all data management steps in the Quality System (QS). Additionally, all members of the research team will be able to access the STATA code / R-scripts used in the cleaning, preparation, and analysis stages, as well as all intermediate and final results.

Statistical analyses will be performed using the statistical packages STATA 15 or R for Linux. All data management is documented in STATA do-files or R-scripts to ensure reproducibility of the results.

9.10. Data analysis

Methodology for summary and statistical analyses of data collected in this study are documented in this section.

Statistical analyses

Mean, median, and standard deviation will be provided for continuous variables when performing descriptive analysis of continuous data. Numbers and percentages will be provided for dichotomous and polychotomous variables when performing descriptive analysis of categorical data. Bivariate comparisons of baseline characteristics and outcomes measures will be provided. An unadjusted Kaplan Meier curve will be drawn to illustrate time-to-event.

Handling of missing values

The statistical analyses will be carried out using available data only. The analyses should be subject to little or no missing data as it is based on national register data covering the patients from the whole country. In Norway, notification of new cases of cancer to the cancer registry is mandatory and the registry is proved to have an overall completeness of close to 100 percent[10]. The data are collected from hospitals, physicians, and pathology laboratories, as well as by linkage with the Norwegian Patient Registry (NPR). Thus, it is assumed that a patient who is not registered with the events of interest in the NPR or CRN did not have them.

The coverage of drug treatment for patients enrolled in clinical trials in CRN will be limited [6]. As of October 2021, drug treatment in hospitals in the northern health region (Helse Nord) is not included in the CRN, because the region lacks electronic reporting of drug treatment [6]. For these patients, cancer drug treatment can only be identified using NPR and NDR. Missing data is expected related to reflex testing and cTNM stage classification, especially in the early years. No imputation for missing values will be performed and analyses will be carried out using available data only.

9.11. Quality control

The study involves the collection and analyses of individual patient-level data. The registries will compile the full datasets to be used for this retrospective register (database) study. All data management is documented in STATA do-files or R-scripts to ensure reproducibility of the results.

9.12. Strengths and limitations of the research methods

The strength of the study is that it – as a retrospective registry (database) study utilizing the national registries in Norway – is a study on the population level. This makes the study and its results as complete and generalizable as possible, as well as minimizing the risk of selection bias. Furthermore, the study is – with the use of the national registries – based on the daily clinical practice and not a study set-up with protocol-required data collection.

As cancer treatment in private hospitals is negligible, the public registries cover virtually all cancer related episodes of care. The quality of the data in CRN is found to have a high degree of comparability, accuracy, and timeliness, and the completeness is estimated to be close to 100 percent [10]. Moreover, all Norwegian pharmacies are required by law to send electronic data to NDR, and the completeness is regarded to be good [11]. The registry reported missing or incorrect personal identification (ID) on 0.14 percent of the prescriptions in 2019 (information provided by the registry in 2021).

Since new data is not being collected, the data may miss important information, variables may be inaccurate or unfit for the study objective, or the data may be less detailed than desired. Different registries sometimes use different coding of variables (classification of diagnosis, age groups etc.) which may present a challenge when combining different data sources. The study is limited by the fact that the available data being used is located in the national registries, as well as accepting that the coding reported in the national registries is correct. One main limitation of the study is that information on drug treatment from hospitals in the northern region of Norway is lacking. Information on patients receiving treatment in clinical trials and compassionate use programs is also lacking, which is the case for several patients in the EGFR+ group. The time of introduction of reflex testing and recording of gene alterations will influence the level of completeness for these markers and will make it difficult to identify subpopulations (EGFR+, ALK+, ROS1+, and PD-L1+) for the early years of the study period. As discussed earlier in this protocol, cTNM stage classification is missing for the years 2009-2013, and is not complete for the years 2014-2017. Therefore, the

analyses may be limited to patients with metastatic NSCLC (stage IV) if it is impossible to identify patients with stage IIIb or IIIc without also including II+IIIa.

An important weakness of the study is that in many cases it is not known why the patient has discontinued treatment. This is particularly a challenge for the drug survival analyses for immunotherapies as patients may stop treatment after 2 years even if they respond well to treatment. This because it is believed that the responders to immunotherapies have long lasting response and that it is unnecessary to continue drug treatment after 2 years. Based on the collected registry data, it will be challenging to separate "treatment breaks" from discontinuation of treatment due to adverse events.

9.13. Other aspects

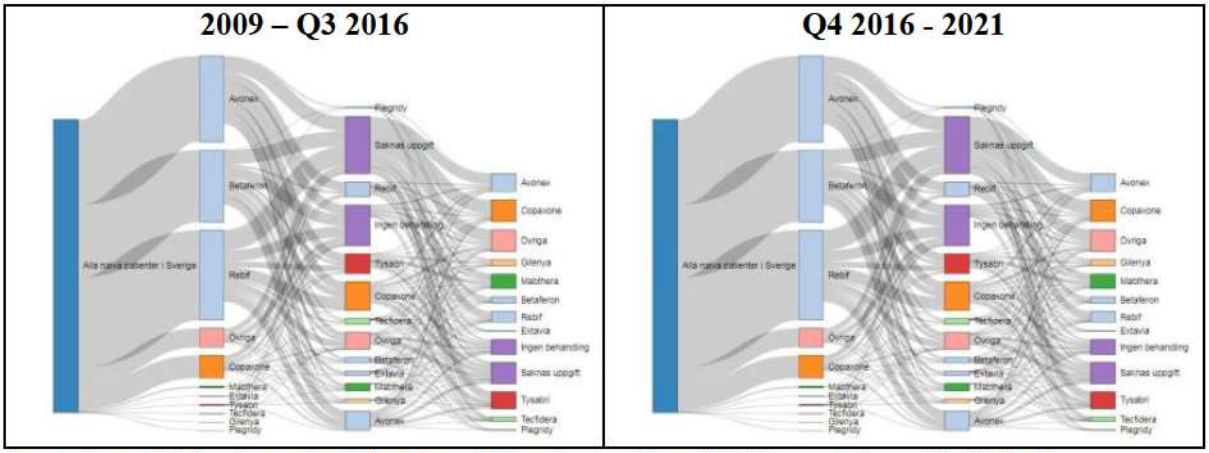
Not applicable

9.14. Table shells

This section provides examples of table shells that will be used to present the results. Note that this is not a complete presentation of all results that will be presented, but examples of some key result tables and figures.

Primary objectives
The first primary objective (A1) is to describe the cancer drug treatment sequence for patients with metastatic NSCLC. This will be done for the 4 study populations (ALK+, EGFR+, PD-L1+, and ROS1+) using Sankey diagrams. The results will be presented for different study periods as illustrated in Figure 3. In this example we use PD-L1+ for illustrative purposes.

Figure 3: Drug Treatment Patterns for PD-L1+ Patients, 2009-2021 (Objective A1)



Source: Example taken from a Swedish study on MS treatment patterns [12]. Time periods used for illustrative purposes only.

Results will also be presented as number and proportion of patients receiving the different treatments as illustrated in Table 6.

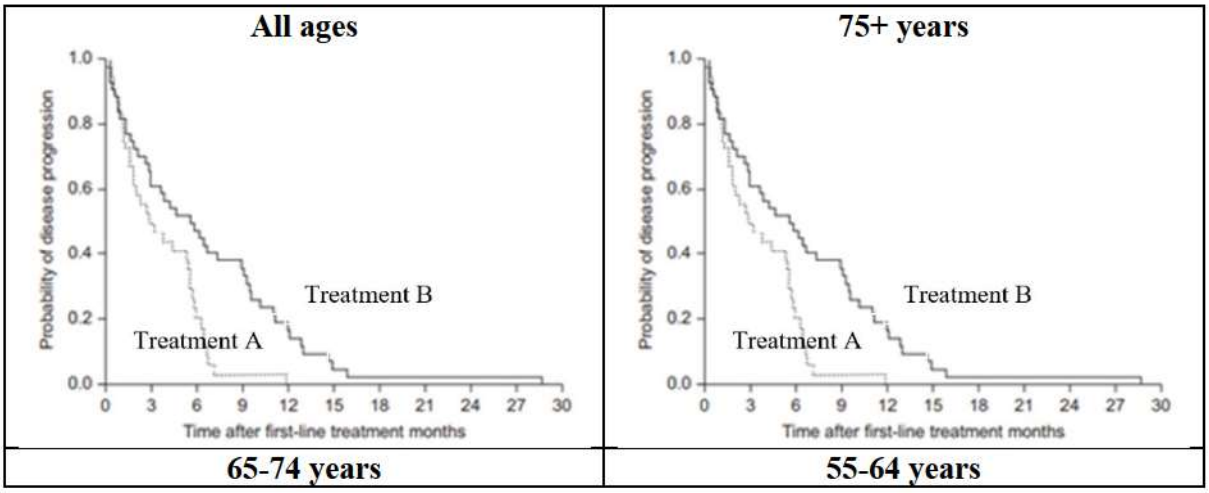
Table 6: Number and Proportion of PD-L1+ Patients Receiving Drug Treatment in 1st and 2nd Line, by Drug and Year (Objective A1)

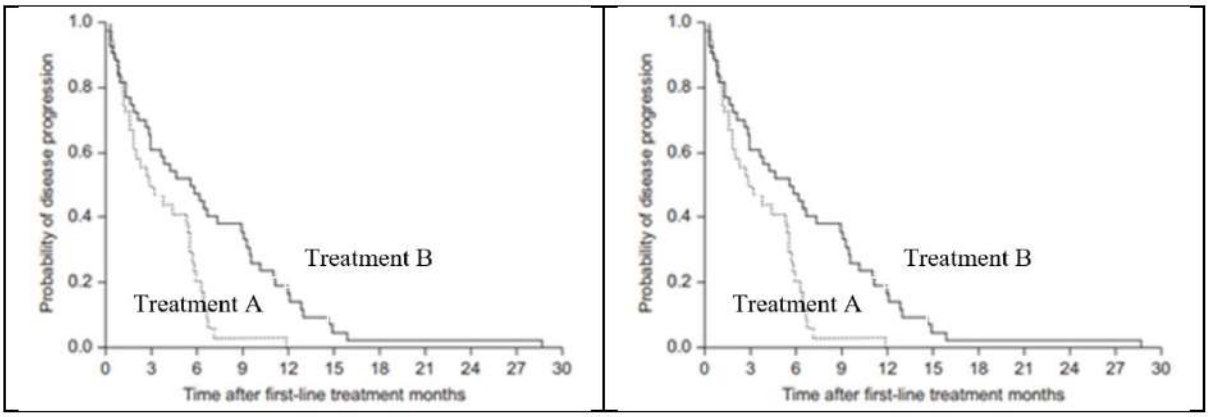
	2009	2010	2011	…	2019	2020	2021
1st line							
Treatment A	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Treatment B	N/A	N/A	N (%)	N (%)	N (%)	N (%)	N (%)
Treatment C	N/A	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
…							
2nd line							
Treatment A	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Treatment B	N/A	N/A	N (%)	N (%)	N (%)	N (%)	N (%)
Treatment C	N/A	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
…							

Table periods used for illustrative purposes only

The next research question (A2) is related to time to treatment discontinuation (retention rate) for different drugs. Results will be presented both as Kaplan-Meier curves and in tables as illustrated in Figure 4 and Table 7 below. In this example we use age as grouping variable for illustrative purposes.

Figure 4: Drug Survival for Patients Diagnosed With ALK+ NSCLC by age Group, 2009-2021 (Objective A2)





Time periods used for illustrative purposes only

Table 7: Length of Treatment for Selected Drugs to Treat ALK+ NSCLC by age Group, 2009-2021 (Objective A2)

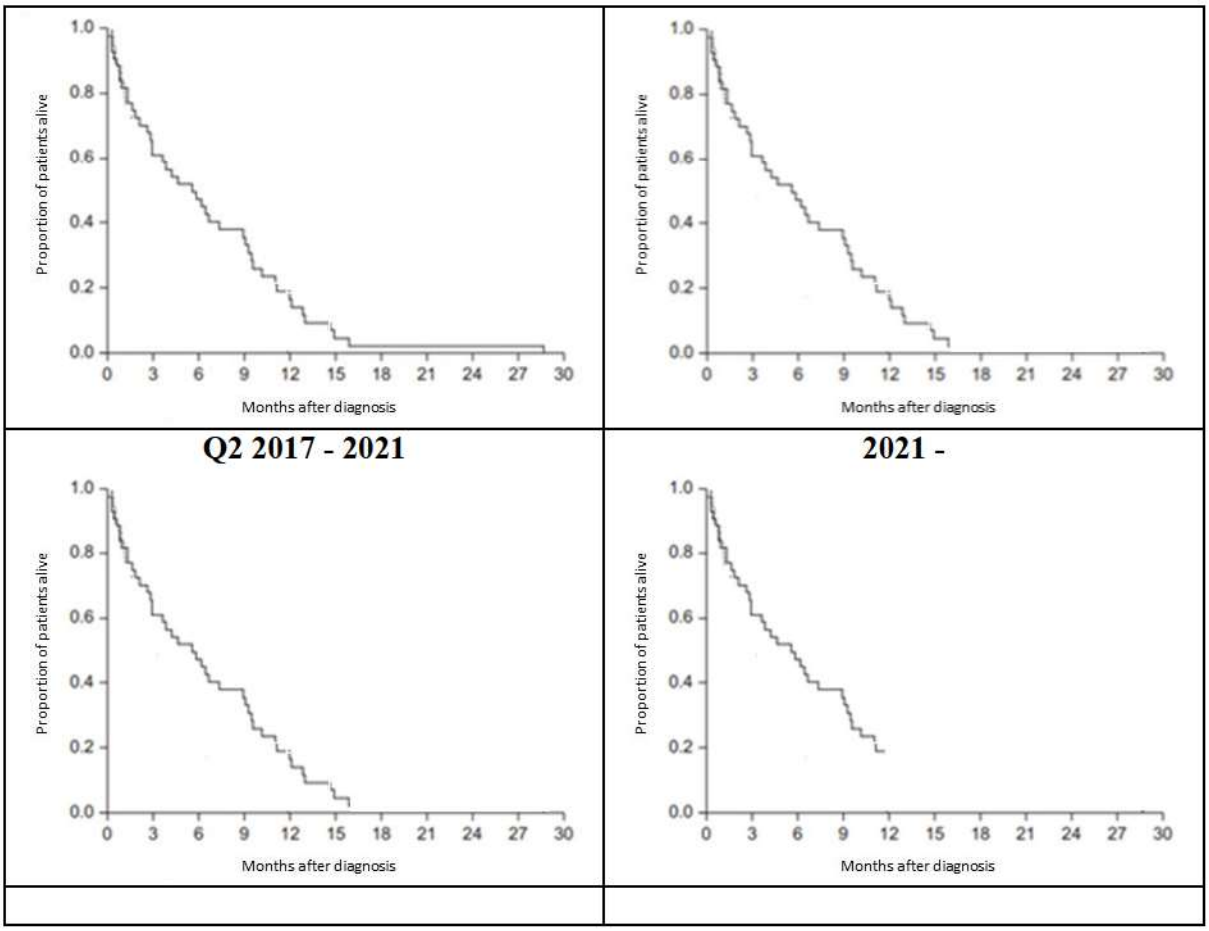
Treatment	Number of patients	Mean treatment length (days)	Share in treatment after 1 month	Share in treatment after 3 months	Share in treatment after 6 months
All ages					
Treatment A					
Treatment B					
Treatment C					
…					
75+ years					
Treatment A					
Treatment B					
Treatment C					
…					
65-74 years					
Treatment A					
Treatment B					
Treatment C					
…					

Table used for illustrative purposes only

The final primary objective (A3) relates to changes in overall survival (OS) over time. Results will be presented both as Kaplan-Meier curves and in tables as illustrated in Figure 5 and Table 8 below. In this example we use the ALK+ group and corresponding study periods for illustrative purposes.

Figure 5: Overall Survival for Patients Diagnosed with ALK+ NSCLC by Year of Diagnosis, 2009-2021 (Objective A3)

2009-2012	2012 -Q1 2017



Time periods used for illustrative purposes only

Table 8: Overall Survival for NSCLC Patients by Year of Diagnosis, 2009-2021 (Objective A3)

Treatment	Number of patients	Mean survival (months)	Proportion alive after 3 months	Proportion alive after 6 months	Proportion alive after 9 months
ALK+					
2009 - 2012					
2012 – Q1 2017					
Q2 2017 - 2021					
2021 -					
EGFR+					
2009 - 2020					
2020 -					
PD-L1+					
2009 Q3 2016					
Q4 2016 -					
ROS1+					
2018-					

Time periods used for illustrative purposes

Secondary objectives

Baseline characteristics of patients diagnosed with metastatic NSCLC will be described in terms of stage at diagnosis, comedication, and other patients' demographics (e.g.: age, gender, place of residence) as illustrated in Table 9 (example for illustrative purposes).

Table 9: Patient Characteristics, 2009 – 2021 (Objective B1)

	Number of patients				Proportion of patients			
	Males	*Females*	*Other/ Unknown*	*Total*	*Males*	*Females*	*Other/ Unknown*	*Total*
Age at diagnosis								
All ages								
75+ years								
65 – 74 years								
55 – 64 years								
45 – 54 years								
> 45 years								
Stage at diagnosis								
IIIb								
IIIc								
IVa								
IVb								
Comedication								
Comedication A								
Comedication B								
Comedication C								
....								

Table used for illustrative purposes only

Prescription practice for patient administered drugs will be described as the proportion of prescriptions lasting for 1, 2, 3, or more than 3 months as illustrated in Table 10.

Table 10: Proportion of Prescriptions by Length of Prescription for Different Drugs, First Prescription Only, 2016-2021 (Objective B2)

Prescription	<31 days	31 – 60 days	61 – 90 days	> 90 days
All regions				
First Prescription				

Second Prescription				
Third Prescription				
…				
South eastern				
First Prescription				
Second Prescription				
Third Prescription				
…				
Central				
First Prescription				
Second Prescription				
Third Prescription				
…				
Western				
First Prescription				
Second Prescription				
Third Prescription				
…				
Northern				
First Prescription				
Second Prescription				
Third Prescription				
…				

Table used for illustrative purposes only

10. PROTECTION OF HUMAN SUBJECTS

10.1. Patient information

All parties will ensure the protection of patient data and will not include patient names on any sponsor forms, reports, publications, or in any other disclosures except where required by law. As detailed, Oslo Economics will only have access to data in de-personalized form (project specific PID) from the national registries as data will be made de-identified by the individual registries before delivery of data files to Oslo Economics. Data delivered by the individual registries will be stored at Oslo Economics' secure server and data will be handled according to Oslo Economics' guidelines for managing sensitive data. Pfizer will not have access to any individual data in connection with the project but only the summary data produced by Oslo Economics.

This study involves data that exists in an anonymized structured format and contains no patient personal information.

10.2. Patient consent

As this study involves anonymized structured data, which according to applicable legal requirements do not contain data subject to privacy laws, obtaining informed consent from patients by Pfizer is not required.
The Norwegian Centre for Research Data (Oslo Economics' Data Protection Officer) will assess the project according to the General Data Protection Regulation (GDPR) and their assessment will be used in the application to each of the registries.

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

Approval from REC and the individual registries is required to access individual patient level data.

No other approvals from IRBs or IECs are necessary by Norwegian law.

10.4. Ethical conduct of the study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigor, and follow generally accepted research practices.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

This study involves: data that exist as structured data by the time of study start.

In these data sources, individual patient data are not retrieved or validated, and it is not possible to link (i.e., identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (AE) (i.e., identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

The results of the study will be published in an Oslo Economics report on Oslo Economics' homepage.

Furthermore, depending on the findings, 1 or 2 scientific article manuscripts will be submitted to an international peer-reviewed journal. Authorship of the article manuscript will follow the requirements set by the International Committee of Medical Journal Editors (ICMJE) (http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html). All members of the research team, including external members and Pfizer employees, fulfilling these requirements with respect to the article manuscript are offered co-authorship of the scientific article.

In the event of any prohibition or restriction imposed (e.g., clinical hold) by an applicable competent authority in any area of the world, or if the party responsible for collecting data

from the participant is aware of any new information which might influence the evaluation of the benefits and risks of a Pfizer product, Pfizer should be informed immediately.

13. REFERENCES

1. Norwegian Institute of Public Health. *Dødsårsaksregisterets statistikkbank [English title: Cause of Death Registry statistics bank]*. 2018 12.12.2018 [cited 2019 20 Nov]; Available from: http://statistikkbank.fhi.no/dar/.
2. Cancer Registry of Norway, *Cancer in Norway 2020 - Cancer incidence, mortality, survival and prevalence in Norway*. 2021, Cancer Registry of Norway: Oslo.
3. Sher, T., G.K. Dy, and A.A. Adjei, *Small cell lung cancer.* Mayo Clin Proc, 2008. **83**(3): p. 355-67.
4. Navada, S., et al., *Temporal trends in small cell lung cancer: analysis of the national Surveillance, Epidemiology, and End-Results (SEER) database.* Journal of Clinical Oncology, 2006. **24**(18_suppl): p. 7082-7082.
5. Brustugun, O.T., et al., *Substantial nation-wide improvement in lung cancer relative survival in Norway from 2000 to 2016.* Lung Cancer, 2018. **122**: p. 138-145.
6. Kreftregisteret, *Årsrapport 2020 med resultater og forbedringstiltak fra Nasjonalt kvalitetsregister for lungekreft*. 2021: Oslo.
7. Cancer Registry of Norway, *INSPIRE:lungekreft. Evaluering av pilotprosjekt.* 2021: Oslo.
8. Cancer Registry of Norway, *Cancer in Norway - Cancer incidence, mortality, survival and prevalence in Norway*. 2019.
9. Cho, Y., C. Hu, and D. Ghosh, *Analysis of regression discontinuity design using cencored data.* 2019.
10. Larsen, I.K., et al., *Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness.* Eur J Cancer, 2009. **45**(7): p. 1218-1231.
11. Furu, K., *Establishment of the nationwide Norwegian Prescription Database (NorPD)--new opportunities for research in pharmacoepidemiology in Norway.* Norsk epidemiologi, 2008. **18**(2).
12. Hagel, E., J. Hillert, and L. Stawiarz. *Real-time monitoring of MS-immunotherapy discontinuation, switching, and restarting with a Visualisation and Analysis Platform (VAP) implemented in the Swedish MS-Registry*. in *MULTIPLE SCLEROSIS JOURNAL*. 2017. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND.

14. LIST OF TABLES

15. LIST OF FIGURES

ANNEX 1. LIST OF STAND-ALONE DOCUMENTS

None.

ANNEX 2. ADDITIONAL INFORMATION

Not applicable.