RESEARCH PLAN (RWE SCIENTIFIC PROJECT)

TITLE:	NATIONWIDE REGISTRY-BASED COHORT STUDY OF TREATMENT PATTERNS, DISEASE PROGRESSION, AND OUTCOMES IN HORMONE RECEPTOR-POSITIVE, HER2-NEGATIVE BREAST CANCER IN NORWAY: NOR-REAL-BC
PROJECT NUMBER:	SL46796
VERSION NUMBER:	2.0
STUDIED MEDICINAL PRODUCTS:	Not Applicable
RESEARCH INITIATOR:	Roche Norge AS Brynsengveien 6b 0667 Oslo Norway
AUTHOR:	[redacted], M.Sc.Pharm., Ph.D. [redacted] NordicRWE Tel.: [redacted] E-mail: [redacted]
DATE FINAL:	See electronic date stamp below

FINAL RESEARCH PLAN APPROVAL

CONFIDENTIAL

This research is managed by F. Hoffmann-La Roche Ltd of Basel, Switzerland. However, it may be implemented in individual countries by Roche's local affiliates, including Genentech, Inc. in the United States. The information contained in this document, especially any unpublished data, is the property of F. Hoffmann-La Roche Ltd (or under its control) and therefore is provided to you in confidence. It is understood that this information will not be disclosed to others without written authorization from Roche except to the extent necessary.

TABLE OF CONTENTS

1. LIST OF ABBREVIATIONS 4
2. RESEARCH TEAM 6
3. AMENDMENTS AND UPDATES 8
4. MILESTONES 8
5. RESEARCH QUESTION AND OBJECTIVES 8
   - 5.1 Background and Rationale 8
   - 5.2 Research Question 10
   - 5.3 PRIMARY OBJECTIVE 10
   - 5.4 SECONDARY OBJECTIVES 10
6. RESEARCH METHODS 11
   - 6.1 Research Design 11
   - 6.2 Setting 12
     - 6.2.1 Study period 12
     - 6.2.2 Data sources 13
     - 6.2.3 Study population 14
   - 6.3 Variables and Endpoints 14
     - 6.3.1 Overview 14
     - 6.3.2 Baseline variables 15
     - 6.3.3 Exposure variables (treatment patterns) 15
     - 6.3.4 Primary Endpoints 16
     - 6.3.5 Secondary and Exploratory Endpoints 18
     - 6.3.6 Covariates 19
   - 6.4 Data Source(s) 20
     - 6.4.1 Overview of Data Sources 20
     - 6.4.2 Cancer Registry of Norway (CRN) 20
     - 6.4.3 Norwegian Patient Registry (NPR) 21
     - 6.4.4 Norwegian Prescribed Drug Registry (NorPD) 21
     - 6.4.5 Norwegian Primary Care Registry (KPR) 21
     - 6.4.6 Data Linkage and Data Handling 22

6.4.7 Data Completeness and Validity 22
6.4.8 Limitations of Data Sources 22
6.5 Research Size 23
6.6 Data Management 23
6.6.1 Data Quality Assurance 23
6.7 Data Analysis and other statistical considerations 24
6.7.1 General Analytical Approach 25
6.7.2 Data Processing and Analysis Pipeline 25
6.7.3 Descriptive Analyses 25
6.7.4 Time-to-Event Analyses 26
6.7.5 Estimation of Incidence and Rates 26
6.7.6 Healthcare Resource Utilization (HCRU) analyses 26
[redacted] 28
6.7.8 Subgroup and Stratified Analyses 29
6.7.9 Handling of Missing Data 29
6.7.10 Bias and Confounding 29
6.7.11 Sensitivity Analyses 30
6.8 Research Documentation 30
6.9 Limitations of the research method 31
7. PROTECTION OF HUMAN PATIENTS 32
7.1 Informed Consent 32
7.2 Confidentiality 32
7.3 Institutional Review Board or Ethics Committee 32
8. PLANS FOR DISSEMINATION AND COMMUNICATION OF RESEARCH RESULTS 33
9. MANAGEMENT AND REPORTING OF ADVERSE EVENTS 33
10. RETENTION OF RECORDS 33
11. REFERENCES 35

1. LIST OF ABBREVIATIONS

Abbreviation	Definition
ALND	Axillary lymph node dissection
BC	Breast Cancer
BCT	Breast Conserving Therapy
BCSS	Breast Cancer-Specific Survival
CCI	Charlson Comorbidity Index
CRN	Cancer Registry of Norway
DRFS	Distant Recurrence-Free Survival
DRG	Disease Related Group
eBC	Early Breast Cancer
ER	Estrogen Receptor
ESR1	Estrogen Receptor 1 (gene)
GP	General Practitioner
HER2	Human Epidermal Growth Factor Receptor 2
HCRU	Healthcare Resource Utilization
[redacted]	
ICD-10	International Classification of Diseases, 10th Revision
KPR	Norwegian Primary Care Registry
LHRH	Luteinizing-hormone-releasing hormone agonist
LoT	Line of Therapy
LRR	Locoregional recurrence
mBC	Metastatic Breast Cancer
NBCG	Norwegian Breast Cancer Group
NCSP	Nordic Medico-Statistical Committee Classification of Surgical Procedures
NICE	National Institute for Health and Care Excellence
NorPD	Norwegian Prescription Database
NPR	Norwegian Patient Registry
OFS	Ovarian function suppression
OS	Overall Survival
PI3K	Phosphoinositide 3-kinase
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

Abbreviation	Definition
PR	Progesterone Receptor
RWE	Real-World Evidence
SAP	Statistical Analysis Plan
SERD	Selective Estrogen Receptor Degrader
SN	Sentinel node biopsy
TNM	Tumor, Node, Metastasis staging system

2. RESEARCH TEAM

Roche Responsible Lead

[redacted]
Roche Norge AS
Brynsengveien 6b
0667 Oslo
Norway

Roche Project Team Members

[redacted]
Roche Norge AS
Brynsengveien 6b
0667 Oslo
Norway

[redacted]
Roche Norge AS
Brynsengveien 6b
0667 Oslo
Norway

RWE Data Science Responsible and Project Lead

[redacted]
NordicRWE AS
Gaustadalléen 21
0349 Oslo
Norway

RWE Project Team Members

[redacted]
NordicRWE AS
Gaustadalléen 21
0349 Oslo
Norway

[redacted], Statistician
NordicRWE AS
Gaustadalléen 21
0349 Oslo
Norway

[redacted]
NordicRWE AS
Gaustadalléen 21
0349 Oslo
Norway

Breast Cancer Clinical Experts

[redacted]
Affiliated to Ålesund sykehus, Helse Møre og Romsdal HF
Åsehaugen 1
6026 Ålesund
Norway

[redacted]
Affiliated to Sykehuset Østfold HF
Kalnesveien 300
1714 Grålum
Norway

3. AMENDMENTS AND UPDATES

Amendment/ Update Number	Date	Section of Research Plan	Amendment or Update	Reason
1	25-Jun-2026	2-Research team	The research team description has been changed from version 1 to version 2 to include the extended research team and align the role with the application to the regional ethical committee.	As per request from the regional ethical committee

4. MILESTONES

Milestone	Planned Date
Final protocol	April 2026
Ethical committee approval	June 2026
Data Permit Authority approval	September 2026
Dataset complete	November 2026
First Interim report	February 2027
Final report of research results	April 2027

5. RESEARCH QUESTION AND OBJECTIVES

5.1 BACKGROUND AND RATIONALE

Breast cancer is the most common cancer among women worldwide and represents a significant public health burden.[1] Although rare, breast cancer also occurs in men and is predominantly hormone receptor-positive and HER2-negative (HR+/HER2−), with similar biological characteristics but limited dedicated evidence.[2]

The HR+/HER2− subtype accounts for the majority of breast cancer cases and is characterized by a heterogeneous disease course, ranging from early-stage disease with curative intent to advanced or metastatic disease requiring long-term systemic treatment.[3]

The treatment landscape in HR+/HER2− breast cancer has evolved substantially in recent years, with the introduction of novel endocrine therapies and targeted agents, including CDK4/6 inhibitors and therapies targeting key molecular pathways such as PI3K/AKT.[4-6] In addition, emerging therapies targeting specific genomic alterations (e.g., PIK3CA mutations) and next-generation selective estrogen receptor degraders (SERDs), including agents targeting ESR1-driven resistance mechanisms, are expected to further transform treatment strategies.[5-9] While these advances have improved outcomes, they have also increased the complexity of treatment pathways across disease stages and lines of therapy.

In early breast cancer (eBC), treatment aims to reduce the risk of recurrence through multimodal strategies, including surgery, radiotherapy, and systemic therapies. Despite this, a proportion of patients will experience recurrence and progress to advanced disease.[3] In the metastatic setting, treatment is generally not curative, and management focuses on prolonging survival and maintaining quality of life through sequential lines of therapy.

Randomized clinical trials provide evidence on efficacy and safety but are conducted in selected populations under controlled conditions. Real-world evidence (RWE) is therefore needed to complement trial data by describing clinical characteristics, treatment patterns and outcomes in routine clinical practice, as well as identifying potential unmet medical needs.[10-11]

Norway's nationwide health registries provide a unique opportunity to generate population-based RWE with comprehensive longitudinal follow-up and linkage across healthcare settings.[12] These data enable detailed characterization of patient trajectories from diagnosis through treatment and progression, including transitions from early to advanced disease.

Despite these strengths, there is limited comprehensive real-world evidence describing treatment pathways in HR+/HER2− breast cancer in Norway, including treatment patterns in early disease, progression to advanced disease, and first-line treatment in the metastatic setting.

This study aims to address these evidence gaps by leveraging nationwide registry data to characterize treatment patterns, patient trajectories, and outcomes in HR+/HER2− breast cancer, providing a robust evidence base to inform clinical practice and future research in an increasingly complex and evolving therapeutic landscape.

5.2 RESEARCH QUESTION

The primary aim of this study is to describe the real-world clinical characteristics, treatment patterns, and outcomes of patients with hormone receptor-positive (HR+), HER2-negative breast cancer in Norway, across both early and metastatic settings.

Specifically, the study seeks to address the following research questions:

1. **Early breast cancer (eBC):**
   What are the treatment patterns, disease trajectories, and clinical outcomes - including distant recurrence free survival (DRFS), locoregional recurrence (LRR), overall survival (OS), and breast cancer specific survival (BCSS) - among patients diagnosed with HR+/HER2- early breast cancer in Norway?
2. **Metastatic breast cancer (mBC):**
   What are the characteristics, treatment patterns, and outcomes of patients with metastatic breast cancer, including both:
   - patients who progress from early breast cancer, and
   - patients with *de novo* metastatic disease at diagnosis?
3. **Clinical and decision-relevant populations:**
   What is the size, characteristics, and outcomes of patient subpopulations relevant for clinical decision-making and [redacted], including patients meeting eligibility criteria for novel targeted therapies.

By leveraging nationwide registry data with complete longitudinal follow-up, the study aims to provide a comprehensive real-world characterization of the breast cancer disease continuum, from initial diagnosis through recurrence and metastatic disease.

5.3 PRIMARY OBJECTIVE

The primary objective for this research is as follows:

- To describe the characteristics, real-world treatment patterns and clinical outcomes among patients with HR+/HER2- early breast cancer in Norway.

5.4 SECONDARY OBJECTIVES

Secondary objectives are intended to provide additional characterization of disease progression, treatment patterns, and outcomes across the HR+/HER2- breast cancer continuum.

The secondary objectives for this research are as follows:

- To describe disease progression from early breast cancer to locoregional and/or metastatic disease and characterize the treatment phase among patients who relapse.

- To describe the characteristics, treatment patterns, and outcomes of patients diagnosed with metastatic breast cancer at initial diagnosis (*de novo* mBC).

- To describe healthcare resource utilization (HCRU) for patients with HR+/HER2- breast cancer across disease stages.

- To estimate the incidence and timing of endocrine resistance among patients with HR+/HER2− early breast cancer, based on predefined real-world proxy definitions.

- To estimate the size, characteristics, and outcomes of patient populations eligible for novel therapies based on alignment with clinical trial eligibility criteria.

6. RESEARCH METHODS

6.1 RESEARCH DESIGN

This study is a retrospective, population-based observational cohort study based on secondary use of nationwide Norwegian health registry data.

The study is designed to characterize the full disease continuum of breast cancer, including:

1. Patients diagnosed with early breast cancer (eBC; stage I–III)

2. Patients diagnosed with *de novo* metastatic breast cancer (mBC; stage IV at diagnosis)

3. Patients who progress from early breast cancer to locoregional and/or metastatic disease

Overall design framework

The study applies a multi-cohort longitudinal design with three parallel entry pathways:

- Early breast cancer cohort (eBC cohort):
  Patients diagnosed with stage I–III breast cancer at initial diagnosis
- *De novo* metastatic cohort (de novo mBC cohort):
  Patients diagnosed with metastatic (stage IV) breast cancer at initial diagnosis
- Progression-defined cohort (recurrent BC cohort):
  Patients initially diagnosed with early breast cancer who subsequently develop locoregional recurrence or metastatic disease

Each cohort is defined by a distinct index event, and patients are followed longitudinally to assess treatment patterns and clinical outcomes.

Index dates and follow-up

- For the eBC cohort, the index date is defined as the date of initial breast cancer diagnosis

- For the *de novo* mBC cohort, the index date is defined as the date of breast cancer diagnosis with metastatic disease at presentation

- For the recurrent BC cohort, the index date is defined as the date of first evidence of recurrent disease (both non-distant and distant recurrence), based on predefined registry-based events

Follow-up will begin at the relevant index date and continue until the earliest of:
- death
- emigration
- end of study period

Design rationale

A retrospective registry-based cohort design is appropriate due to:
- Nationwide coverage and high completeness of Norwegian health registries
- Ability to capture longitudinal patient trajectories across care settings
- Availability of detailed information on tumor characteristics, treatments and outcomes

The inclusion of both early and metastatic cohorts enables a comprehensive assessment of disease progression and treatment pathways, reflecting real-world clinical practice.

6.2 SETTING

The study will be conducted in Norway, using nationwide health registries covering the entire population.

Norway has a universal healthcare system with mandatory reporting to national registries, enabling comprehensive capture of cancer diagnoses, treatments, and outcomes.

6.2.1 Study period

Inclusion period:

Patients diagnosed with breast cancer during the period 2012 to the most recent available data cut (expected up to 31.12.2025, depending on registry updates)

Baseline period. Historical data will be used to:
- identify incident cases
- assess other primary cancers, prior comorbidities and medical history

For comorbidity and comedication assessment, data from the Norwegian Patient Registry and Norwegian Prescribed Drug Registry will be used from 2008 onwards, corresponding to registry availability.

Follow-up period. Patients will be followed from index date until:

- death
- emigration
- end of study period

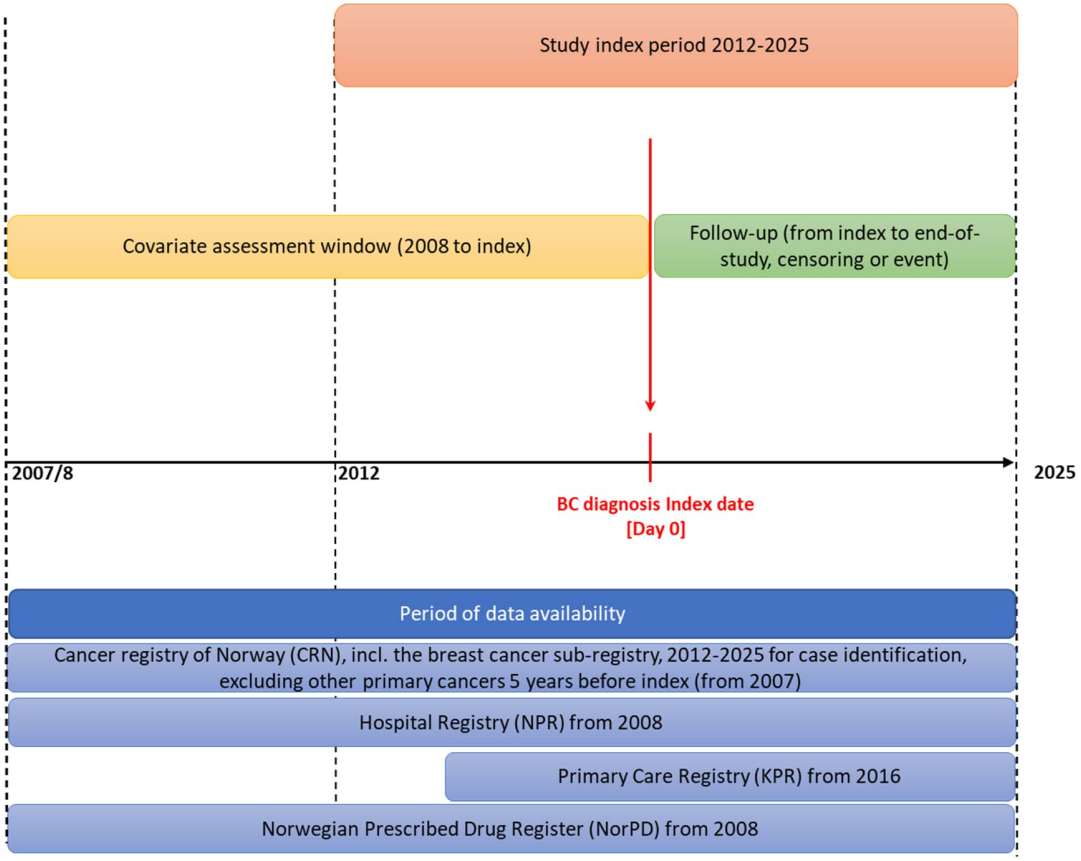

Figure 1: Study design and data availability. The study index period spans 2012–2025, with the index date defined as the date of breast cancer diagnosis (Day 0). Baseline covariates are assessed from 2008 up to index, subject to registry-specific data availability. Patients are followed from index until end of study, event, or censoring. Data sources include the Cancer Registry of Norway (CRN; 2012–2025) for case identification, as well as excluding other prior primary cancers within 5 years before index (2007), the Norwegian Patient Registry (NPR; from 2008), the Primary Care Registry (KPR; from 2016), and the Norwegian Prescribed Drug Registry (NorPD from 2008).

6.2.2 Data sources

Cancer Registry of Norway (CRN), including the national breast cancer quality registry, provides comprehensive information on cancer diagnoses, including:

- date of diagnosis
- other primary cancers

- tumor characteristics (e.g., histology, tumor size, nodal status, grade, TNM stage)
- biomarker information including hormone receptor status (ER and PR), HER2 status, and Ki-67 where available, as well as gene expression profile test such as Prosigna (PAM50) and OncotypeDX where available; somatic biomarkers such as PIK3CA and ESR1 mutation status will be included if available in the CRN
- information on primary treatment modalities, including surgery and radiotherapy
- data on hospital-administered systemic anti-cancer therapies enabling identification of treatments such as chemotherapy, immunotherapy and targeted therapies administered in specialist care settings

6.2.3 Study population

The source population consists of all individuals diagnosed with breast cancer in Norway during the study period (1.1.2012 to 31.12.2025 (or the latest data update)).

The study population will be defined based on the following criteria:

Inclusion criteria:

- Females and males aged ≥18 years at diagnosis
- Incident breast cancer (ICD-10 C50) diagnosed during the study inclusion period
- Hormone receptor-positive (HR+) and HER2-negative breast cancer at diagnosis
- Histologically verified breast cancer

Exclusion criteria:

- Prior history of other primary malignancy within 5 years prior to the breast cancer index date, with the exception of non-melanoma skin cancer and benign in situ neoplasms
- Breast cancer diagnosis identified only at autopsy

6.3 VARIABLES AND ENDPOINTS

6.3.1 Overview

This study will utilize linked nationwide registry data to define:

- Baseline characteristics
- Exposure variables (treatment patterns)
- Clinical outcomes and endpoints
- Covariates

All variables will be derived using predefined algorithms based on registry-specific coding systems (e.g., ICD-10, ATC, procedure codes), ensuring reproducibility and transparency. Detailed operational definitions will be specified in the Statistical Analysis Plan (SAP).

6.3.2 Baseline variables

Baseline characteristics will be assessed prior to the index date and will include, but are not limited to:

Sociodemographic variables

- Age at diagnosis
- Sex (M/F)
- Region of residence (county) / treating region (RHF)
- Year of diagnosis

Clinical characteristics

- TNM stage at diagnosis
- Tumor characteristics:
    - Tumor size
    - Histological grade
    - Hormone receptor status (ER/PR) and ER/PR expression (%)
    - HER2 status
    - Ki67 (where available) and gene expression profiles (Prosigna PAM50 /OncotypeDX where available)
    - PIK3CA mutation status and ESR1 (where available)
- Nodal involvement:
    - Number of positive lymph nodes
- Metastatic status at diagnosis (for cohort definition)
- Menopausal status, derived using proxy definitions based on available data, including age and treatment-related indicators (e.g., bilateral oophorectomy, use of LHRH agonists).

Comorbidities

Assessed using diagnoses from Norwegian Patient Registry (NPR) and Norwegian Primary Care Registry (KPR). Comorbidity burden will be summarized using Charlson Comorbidity Index (CCI) or equivalent predefined comorbidity measure.

6.3.3 Exposure variables (treatment patterns)

Treatment patterns will be defined using data from CRN, NPR, and NorPD.

Early breast cancer setting

- Surgery:

    - Type of surgery (breast-conserving or mastectomy, lymph node axillary dissection (ALND) or sentinel node biopsy (SN))
- Radiotherapy:
    - Timing (neoadjuvant/adjuvant)
- Systemic therapies:
    - Chemotherapy
    - CDK4/6 inhibitors
    - Endocrine therapy (with or without ovarian function suppression (OFS))

Metastatic setting

- Systemic therapies:
    - Chemotherapy
    - CDK4/6 inhibitors
    - Endocrine therapies (e.g., fulvestrant) with or without ovarian ablation/function suppression
    - Targeted therapies (e.g., everolimus and alpelisib)

6.3.4 Primary Endpoints

Distant Recurrence Free Survival (DRFS) *(early breast cancer cohort only)*

DRFS is defined as the time from index date (date of early breast cancer diagnosis) to the earliest occurrence of any of the following events:

- Death from any cause
- Diagnosis of distant metastasis:
    - ICD-10 codes C78–C79, or
    - ICD-10 code C771, C772, C774, C775, C778 or C779
- Initiation of systemic therapy indicative of metastatic disease, including:
    - CDK4/6 inhibitors (palbociclib, ribociclib, abemaciclib) excluding adjuvant use
    - Fulvestrant
    - Everolimus
    - Alpelisib
    - Chemotherapy
- Radiotherapy indicating advanced disease:
    - Radiotherapy administered at sites other than the breast or marked as palliative intent

Patients will be censored at:

- emigration
- end of study period

Locoregional Recurrence (LRR)

A locoregional recurrence event is defined as the first occurrence of a breast cancer recurrence confined to the ipsilateral breast, chest wall, or regional lymph nodes,

occurring after primary treatment and in the absence of prior or simultaneous distant metastatic disease.

LRR will be identified using a combination of procedure codes, radiotherapy records, and diagnosis codes, following an 12 months disease-free interval from primary treatment (to avoid misclassification of residual disease).

Operational definition:
LRR is defined by the earliest occurrence of any of the following:
1. Chest wall recurrence (post-mastectomy) among patients with prior mastectomy:
- Re-excision surgery involving chest wall or scar area
- Radiotherapy to chest wall not part of primary treatment, potentially with:
    - Breast cancer diagnosis code (C50)
    - Secondary malignancy codes (excluding distant metastases)

2. Regional lymph node recurrence
- Lymph node dissection (e.g., axillary, supraclavicular) not part of primary treatment
- Radiotherapy targeting regional lymph nodes, if not previously administered

3. Ipsilateral breast recurrence among patients treated with breast-conserving therapy (BCT):
- Mastectomy following prior BCT
- Re-excision (new BCT) combined with radiotherapy, where radiotherapy was not part of initial treatment

Systemic treatment initiation may be used as supportive information in combination with locoregional procedures or radiotherapy, but not as a standalone criterion

Timing constraint: Events must occur after a predefined disease-free interval (e.g., ≥12 months after primary treatment completion). Events occurring before this window will be considered part of primary treatment.

Exclusion criteria: Any evidence of distant metastatic disease (e.g., ICD-10 C78–C79) prior to or within a defined window of the event excludes classification as LRR

Overall Survival (OS)

OS is defined as the time from index date to death from any cause.
- Applicable to:
    - early breast cancer cohort
    - *de novo* metastatic cohort
    - recurrent metastatic cohort

Patients emigrated and alive at the end of follow-up will be censored.

Breast Cancer Specific Survival (BCSS)

BCSS is defined as the time from index date to death attributed to breast cancer.

- Cause of death will be obtained from national mortality data
- Death from other causes will be treated as censoring events

6.3.5 Secondary and Exploratory Endpoints

Endocrine resistance

Endocrine resistance will be defined using a registry-based composite algorithm combining information on endocrine therapy exposure and subsequent indicators of disease progression or treatment escalation. The definition is conceptually aligned with clinical trial settings, where endocrine resistance is reflected by disease progression on endocrine therapy.

Resistance events may include occurrence of distant metastasis, as well as initiation of systemic therapies indicative of endocrine-resistant disease (e.g., CDK4/6 inhibitors, fulvestrant, alpelisib, or everolimus-based combinations). Additional criteria and refinements may be applied based on data availability and clinical relevance.

Primary endocrine resistance will be defined as occurrence of a resistance event within approximately 24 months of initiation of adjuvant endocrine therapy. Secondary endocrine resistance will be defined as occurrence of a resistance event after this period, or within approximately 12 months after discontinuation of adjuvant endocrine therapy.

The operational definition may be further refined in the Statistical Analysis Plan (SAP), based on detailed assessment of available data and clinical input.

Metastatic disease identification

Metastatic disease (for recurrent cohort definition) will be identified based on:

- ICD-10 codes for metastasis, defined as:
  - C78–C79, or
  - C771, C772, C774, C775, C778 or C779
- Initiation of metastatic-specific systemic therapies
- Radiotherapy indicative of advanced disease, e.g. localization outside breast or palliative intent radiotherapy

Treatment pattern endpoints

Treatment patterns will be characterized by, but not limited to:

- Time to treatment initiation from clinically relevant index dates
- Treatment persistence
- Treatment sequences and lines of therapy (e.g., neoadjuvant, adjuvant, and metastatic setting including 1st, 2nd, 3rd line and beyond 3rd line)
- Treatment switching patterns across and within drug therapy classes

Definitions

Where multiple index dates are applicable for a patient, separate time-to-initiation measures will be constructed for each relevant treatment setting.

Treatment persistence will be defined as the duration of time from treatment initiation to treatment discontinuation, allowing for permissible gaps between treatment episodes.

Treatment switching will primarily be defined as a change from one drug class to another and will be captured as part of treatment sequences and lines of therapy.

Changes within the same drug class will be captured separately as within-class treatment modifications. While such changes will not, by default, constitute a new line of therapy, they may be considered indicative of treatment progression or line advancement on a case-by-case basis, depending on clinical context and data-driven patterns.

Treatment exposure, sequencing, persistence, discontinuation, and switching will be derived using predefined algorithms based on available registry data, including information on drug dispensing, administration, and treatment gaps.

Detailed operational definitions of treatment patterns will be specified in the Statistical Analysis Plan (SAP).

6.3.6 Covariates

Covariates will include baseline patient and disease characteristics used to describe the study population and to support stratified and subgroup analyses.

These may include, but are not limited to:

- Age and age categories
- Menopausal status (proxy derived)
- Endocrine resistance status
- Adjuvant risk group status (e.g., low, high, or not classifiable) will be derived based on available clinicopathological factors (e.g., nodal status, tumor size, grade, Ki-67, Prosigna/OncotypeDX), in alignment with relevant clinical guidelines (e.g., NBCG) and established risk categories
- Tumor characteristics (grade, stage, biomarkers)
- Comorbidity burden
- Year of diagnosis
- Prior treatments

These variables will be used for:

- Descriptive characterization of the study population
- Stratified analyses

- Subgroup analyses

6.4 DATA SOURCE(S)

6.4.1 Overview of Data Sources

This study will be based on secondary use of linked, nationwide Norwegian health registries. These registries provide comprehensive, longitudinal individual-level data on cancer diagnoses, treatments, healthcare utilization, and outcomes across the entire Norwegian population.

Data will be obtained from the following sources:

6.4.2 Cancer Registry of Norway (CRN)

The Cancer Registry of Norway (CRN) is a nationwide registry with mandatory reporting of all cancer diagnoses. It provides high-quality data on cancer incidence and clinical characteristics and is considered a complete and reliable source for cancer epidemiology in Norway.

The CRN **incidence registry** dataset used in this study includes, but is not limited to, the following variables:

- Age at diagnosis and year of birth
- Sex
- Health region at diagnosis
- Date of diagnosis
- Tumor site and morphology (ICD-O)
- TNM staging
- Laterality
- Information on primary treatments, including:
  - surgery
  - radiotherapy

The **national breast cancer quality registry**, providing detailed clinical information, including:

- Tumor size
- Histological grade
- Lymph node involvement
- Biomarker status (ER, PR, HER2, and Ki67, Prosigna, PIK3Ca and ESR1 where available)
- Information on hospital-administered systemic anti-cancer therapies, enabling identification of:
  - chemotherapy

- other systemic treatments administered in specialist care, e.g. immunotherapy and targeted therapies

6.4.3 Norwegian Patient Registry (NPR)

The Norwegian Patient Registry (NPR) contains nationwide data on specialist healthcare services, including both inpatient and outpatient care.

The NPR provides, but is not limited to:

- Diagnoses coded using ICD-10 (both primary and secondary codes)
- Procedure codes (e.g., surgical and radiotherapy procedures)
- Dates of hospital contacts and treatments (admission, discharge and length of stay)
- Type of contact (e.g. inpatient, outpatient, emergency)
- Type of hospital and department
- Health region
- Disease Related Groups (DRGs)

The NPR will be used to:

- Identify diagnoses of metastatic disease (e.g., ICD-10 C77–C79)
- Capture procedures administered in specialist care, e.g. radiotherapy, surgery
- Assess comorbidities and healthcare utilization

6.4.4 Norwegian Prescribed Drug Registry (NorPD)

The Norwegian Prescribed Drug Registry (NorPD) contains data on all prescription drugs dispensed at pharmacies in Norway.

The NorPD provides, but is not limited to, information on:

- Drug dispensing data coded using ATC classification
- Dates of dispensing
- Information on drug strength, amount (DDD), number of packs and package size, and reimbursement codes

The NorPD will be used to:

- Identify use of endocrine therapies and targeted therapies
- Characterize treatment patterns and persistence of therapy
- Capture medications used in both early and metastatic settings
- Assess co-medications

6.4.5 Norwegian Primary Care Registry (KPR)

The Norwegian Primary Care Registry (KPR) contains data on primary healthcare services, with focus on general practitioner (GP) consultations.

The KPR provides, but is not limited to, information on:

- Diagnoses and clinical information from primary care
- Claim codes

- Data on healthcare utilization in primary care

The KPR will be used to:
- Assess comorbidities diagnosed in primary care
- Capture healthcare utilization not recorded in specialist registries

6.4.6 Data Linkage and Data Handling

At the time of data delivery, registry holders will pseudonymize the data by replacing the personal identification number with a study-specific unique individual identifier, which will be used for data linkage across registries.

All data will be stored and analyzed within a secure research environment, in accordance with applicable data protection regulations and governance requirements.

6.4.7 Data Completeness and Validity

The Norwegian health registries used in this study are characterized by:
- Nationwide population coverage
- Mandatory reporting
- Longitudinal follow-up across healthcare settings

These features ensure a high degree of:
- completeness of data capture
- population representativeness
- external validity, reflecting real-world clinical practice in Norway

As no sampling from a source population is undertaken, the study population is expected to be highly representative of all eligible patients diagnosed with breast cancer in Norway during the study period.

6.4.8 Limitations of Data Sources

Despite the strengths of the registries, certain limitations should be acknowledged:
- Some clinical variables (e.g., Ki67, detailed biomarker data) may be incomplete
- Information on disease progression (e.g., recurrence) is not directly recorded and must be inferred using proxies
- Treatment lines are not directly recorded and must be constructed from available drug exposure data
- Certain treatments administered in hospital settings may be incompletely captured depending on registry coverage and time period
- Potential misclassification of diagnoses or procedures may occur
- Incomplete reporting

These limitations will be addressed through:

- By a rule-based classification of BC treatment lines based on the patients´ full treatment history
- Sensitivity analyses where appropriate
- Transparent reporting of assumptions and definitions

6.5 RESEARCH SIZE

The study will include all eligible patients meeting the inclusion and exclusion criteria within the nationwide Norwegian registries during the study period. As no sampling will be applied, the study size is determined by the underlying population of breast cancer patients in Norway.

Breast cancer incidence in Norway has increased over time, with approximately 3,300 - 4,000 new cases annually during the study period. Over the planned inclusion period (2012 to the most recent available data cut), this corresponds to an estimated total of approximately 50,000 breast cancer cases.

Of these, approximately 70 - 80% are expected to be hormone receptor-positive (HR+) and HER2-negative, resulting in a primary study population of approximately 35,000 patients with HR+/HER2− breast cancer.

Within the overall study population:

- The majority of patients are expected to be diagnosed with early breast cancer (stage I–III) at initial diagnosis

- A smaller proportion will represent:
    - *de novo* metastatic breast cancer (stage IV at diagnosis)
    - patients progressing from early to locoregional and/or metastatic disease during follow-up

The exact size of these sub-cohorts will be determined empirically based on registry data.

Formal sample size calculations are not required, as the study includes the entire eligible population rather than a sampled subset.

6.6 DATA MANAGEMENT

6.6.1 Data Quality Assurance

NordicRWE will be responsible for the data management and data quality assurance of this study, including quality control of all data processing and analyses. The research initiator will maintain oversight of data management activities throughout the study, but not have access to individual-level data.

Data used in this study are obtained from national health registries and are subject to established validation and quality control procedures at the level of the registry holders prior to data delivery. Following delivery, additional study-specific data management and quality assurance procedures will be implemented.

All data will be received in pseudonymized form and handled within a secure research environment, in accordance with applicable data protection regulations and governance requirements. Access to data will be restricted to authorized study personnel.

Upon data receipt, quality control procedures will be performed to ensure that:

- Delivered datasets are complete
- Data are in agreement with the approved variable list and data specifications
- Data structure and formats are consistent with expectations

Any discrepancies will be documented and resolved in dialogue with the registry holders where necessary.

Data processing and quality assurance activities will include:

- Verification of data completeness and internal consistency
- Identification and handling of missing, implausible, or inconsistent values
- Harmonization of variables across data sources, if needed
- Logical validation of key variables, including temporal consistency (e.g., diagnosis, treatment initiation, outcomes)
- Construction and validation of derived variables (e.g., cohort definitions, treatment lines, and endpoints) using predefined algorithms

All statistical programming and data processing will follow standardized quality control procedures, including:

- Structured review of analysis code for completeness and correctness
- Review of program logs to identify errors or warnings
- Verification of key results through independent replication or second-person review where applicable

All data transformations, derivations, and analyses will be fully documented, ensuring transparency, reproducibility, and auditability.

Audit trails and version control will be maintained for all relevant datasets and analysis programs. Data storage, backup, and retention will be performed in accordance with institutional procedures and applicable regulatory requirements.

6.7 DATA ANALYSIS AND OTHER STATISTICAL CONSIDERATIONS

The study will follow established RWE best practices, including NICE framework principles, ensuring transparency, reproducibility, and appropriate handling of missing data, bias, and sensitivity analyses, as detailed in the SAP.

6.7.1 General Analytical Approach

All analyses will be conducted using a predefined statistical analysis plan (SAP) developed prior to data analysis.

The analytical approach is primarily descriptive and longitudinal, aimed at estimating:

- measures of occurrence (e.g., incidence rates)
- treatment patterns and trajectories
- time-to-event outcomes (DRFS, OS, BCSS)

Estimates will be presented with appropriate measures of uncertainty, including confidence intervals (CIs) where relevant.

Statistical analyses will be performed using validated statistical software, e.g. R.

6.7.2 Data Processing and Analysis Pipeline

The analytical workflow will follow a structured and reproducible pipeline:

- Data receipt and validation (Section 6.6)
- Data cleaning and harmonization
- Derivation of cohorts, exposures, and endpoints
- Construction of analysis datasets
- Statistical analyses and reporting
- All steps will be documented to ensure reproducibility

6.7.3 Descriptive Analyses

Baseline characteristics and treatment patterns will be summarized using:

- Continuous variables:
  - mean, standard deviation (SD), median, interquartile range (IQR), minimum, and maximum
- Categorical variables:
  - frequencies and percentages

Comparisons between groups (where relevant) will be performed using:

- Chi-square or Fisher's exact test for categorical variables
- Parametric or non-parametric tests for continuous variables, depending on distribution

All statistical tests will be two-sided, with a significance level of 0.05 where hypothesis testing is applied. Given the primarily descriptive nature of the study, results will be interpreted cautiously without formal adjustment for multiplicity.

6.7.4 Time-to-Event Analyses

Time-to-event endpoints will be analyzed using standard survival analysis methods.

Kaplan–Meier estimation

- Used to estimate:
    - DRFS (eBC cohort)
    - OS (all cohorts)
    - BCSS (all cohorts)
    - Time to treatment discontinuation
- Results will be presented as:
    - survival curves
    - survival probabilities at clinically relevant time points (e.g. 1, 2 and 5 years)
    - median survival where estimable

Model adequacy for parametric survival analyses will be assessed using standard diagnostic approaches, including visual comparison of Kaplan–Meier and fitted curves, as well as evaluation of landmark survival estimates.

Group comparisons

- Survival curves will be compared using the log-rank test, where relevant

For breast cancer-specific survival (BCSS), death from causes other than breast cancer will be treated as a competing event. In addition to cause-specific survival analyses, competing risk methods will be applied where appropriate, including estimation of cumulative incidence functions for breast cancer-specific death.

6.7.5 Estimation of Incidence and Rates

Where relevant, incidence rates (e.g., recurrence rates, endocrine resistance) will be estimated as:

- Number of events divided by person-time at risk

Results will be reported with corresponding confidence intervals.

6.7.6 Healthcare Resource Utilization (HCRU) Analyses

Healthcare resource utilization (HCRU) will be assessed to characterize patterns of healthcare use associated with breast cancer across disease stages and to support health economic evaluations.

HCRU measures will be derived from nationwide registries, including the Norwegian Patient Registry (NPR), the Norwegian Primary Care Registry (KPR), and the Norwegian Prescription Database (NorPD), and will include, but not be limited to:

- Hospital admissions, including number and length of stay
- Outpatient specialist visits
- Primary care contacts, including general practitioner visits
- Selected procedures and interventions
- Dispensed prescription drugs



- Primary care, including general practitioner visits, physiotherapy, and selected municipal services where available
- Specialist care, including inpatient admissions, outpatient hospital contacts, oncology-related visits, and surgical procedures relevant to breast cancer management
- Diagnostic procedures, including imaging (e.g., CT, MRI, PET-CT) and biopsy-related procedures
- Acute events and unplanned visits, including complications and hospitalizations associated with disease progression or treatment (e.g. sepsis, neutropenia, infections)

Resource utilization will be identified using registry-specific coding systems, including ICD-10 diagnosis codes, NCSP procedure codes, and reimbursement-based activity data. Where direct measurement is not available in registry data, proxy-based approaches will be applied. This may include, for example, the use of outpatient hospital contacts as proxies for specific provider types or biopsy procedures as proxies for molecular diagnostic testing.

Costing will be based on standard Norwegian reimbursement frameworks, including diagnosis-related group (DRG) tariffs for specialist care, national tariffs for primary care services, drug costs, and relevant municipal cost sources. Where required, selected components may be estimated using supplementary bottom-up costing approaches informed by clinical and laboratory input.

Where feasible, HCRU will be assessed:

- From index date and over follow-up
- Stratified by disease stage, including early breast cancer (eBC), *de novo* metastatic breast cancer (mBC), and recurrent mBC

- In relation to key clinical events, such as recurrence and treatment initiation

Results will be summarized using:

- Counts and rates, including per patient-year estimates
- Descriptive statistics, including mean, median, and distributional measures

The [redacted] and costing assumptions will be detailed in the Statistical Analysis Plan (SAP) [redacted].

[redacted]

[redacted]

[redacted]

6.7.8 Subgroup and Stratified Analyses

Analyses will be conducted overall and stratified by clinically relevant subgroups, including:

- By different tumor characteristics (e.g., grade, stage, biomarkers)
- By adjuvant setting risk strata (e.g., low, high, or not classifiable)
- By degree of lymph node involvement (pathological node status)
- By treatment status, e.g. drug class
- By time period of diagnosis
- By menopausal status

Subgroup analyses are exploratory and will be interpreted accordingly.

6.7.9 Handling of Missing Data

The extent of missing data will be assessed for all variables.

- Missingness will be described and reported transparently
- No formal imputation is planned for primary analyses
- Analyses will be based on available data

Sensitivity analyses may be conducted where missing data are considered to potentially influence results.

6.7.10 Bias and Confounding

As an observational study based on routinely collected data, several potential sources of bias will be considered:

- Misclassification of exposures and outcomes
- Incomplete capture of clinical variables
- Algorithm-based (proxies) identification of recurrence and treatment lines

To mitigate these risks:

- Predefined and clinically informed algorithms will be used
- Sensitivity analyses will be performed where appropriate
- Results will be interpreted in light of these limitations

Given the primarily descriptive objectives of the study, no formal causal inference framework is applied.

6.7.11 Sensitivity Analyses

Sensitivity analyses may be conducted to assess the robustness of key findings to alternative definitions, assumptions, and analytical choices.

These may include, but are not limited to:

- **Alternative definitions of recurrence or metastatic disease**, including variation in:
  - use of diagnosis codes (e.g., inclusion/exclusion of selected codes)
  - treatment-based proxies (e.g., inclusion of specific systemic therapies or radiotherapy)
- **Variations in treatment exposure definitions**, including:
  - different assumptions for treatment persistence (e.g., allowable gaps between dispensations) as alternative definitions of treatment discontinuation
- **Variations in line of therapy algorithms**, including:
  - alternative rules for defining treatment lines and switches
  - handling of combination therapies and overlapping treatments
- **Variation in time interval definitions,** including**:**
  - alternative time windows for assigning events (e.g., recurrence, including locoregional recurrence, or progression)

The specific sensitivity analyses to be conducted will be defined in the Statistical Analysis Plan (SAP) and may be further refined following detailed assessment of the available data.

6.8 RESEARCH DOCUMENTATION

The research initiator will maintain adequate and accurate records to ensure that the conduct of the research is fully documented. This includes, but is not limited to:

- The study protocol and any protocol amendments
- The statistical analysis plan (SAP) and any updates
- Documentation of data access approvals from relevant authorities (e.g., ethics committees, registry holders, and governmental bodies, where applicable)
- Data management and analysis documentation, including variable definitions and derivation algorithms

The research initiator will ensure that:

- Analysis datasets and derived datasets used in the study are retained in electronic format

- Statistical programs used to generate study results are documented, version-controlled, and retained
- All relevant materials are available for audit and inspection, in accordance with applicable regulatory and institutional requirements

6.9 LIMITATIONS OF THE RESEARCH METHOD

This study is based on routinely collected healthcare data and is therefore subject to several inherent limitations.

First, certain clinical variables may be incomplete or variably recorded over time, including detailed biomarker information (e.g., Ki67) and other pathology-related variables. In addition, some treatments, particularly those administered in hospital settings, may be incompletely captured depending on registry coverage and time period.

Second, key clinical events such as disease recurrence and progression are not directly recorded in the registries and must be inferred using predefined proxy algorithms based on diagnoses, treatments, and procedures. While these algorithms are clinically informed, some degree of misclassification may occur.

Third, treatment patterns—particularly lines of therapy—are not explicitly recorded and must be constructed from available drug exposure data. This requires algorithm-based definitions that may introduce uncertainty, especially in complex treatment sequences.

Fourth, as an observational study based on routinely collected data, the results may be influenced by:

- Differences in patient characteristics and treatment selection (confounding by indication)
- Unmeasured or incompletely captured variables
- Misclassification of exposures or outcomes

These factors may affect the interpretation of observed patterns and associations and will be considered when interpreting the results.

Fifth, when aligning the study population to clinical trial populations, the use of selection criteria or re-weighting approaches depends on the availability and granularity of published trial data. Differences in variable definitions and data availability may limit the degree of alignment that can be achieved.

Finally, while the nationwide coverage of Norwegian registries supports high representativeness and external validity, the findings may be most directly generalizable to healthcare systems with similar structure and clinical practice patterns.

7. PROTECTION OF HUMAN PATIENTS

7.1 INFORMED CONSENT

This study is based on secondary use of data from nationwide, law-regulated health registries in Norway.

In accordance with applicable Norwegian legislation, informed consent is not required for this type of research, as the study uses pseudonymized registry data and has obtained the necessary regulatory approvals.

The study will be conducted subject to:

- Approval from the Regional Committees for Medical and Health Research Ethics (REK)
- Data access approvals from relevant registry holders
- Authorization from the Norwegian Health Data Services (Helsedataservice, HDS), including exemption from the duty of confidentiality granted to the research responsible institution

All data will be:

- Pseudonymized prior to access by the registry holders
- Processed and analyzed within a secure research environment
- Handled in accordance with applicable data protection regulations and governance requirements

7.2 CONFIDENTIALITY

All data used in this study will be pseudonymized prior to access by the research team and processed within a secure research environment in accordance with applicable data protection regulations.

Results from the analyses will be reported only in aggregated and anonymized form, ensuring that no individual patients can be identified.

No attempts will be made to re-identify individuals.

This research will be conducted in full conformance with the Guidelines for GPP published by the International Society of Pharmacoepidemiology and the laws and regulations of the country in which the research is conducted.

7.3 INSTITUTIONAL REVIEW BOARD OR ETHICS COMMITTEE

This research plan and relevant supporting information must be submitted to the IRB/EC to be reviewed and approved by the IRB/EC before the research is initiated.

8. PLANS FOR DISSEMINATION AND COMMUNICATION OF RESEARCH RESULTS

The results of this study will be communicated in accordance with good scientific practice and applicable guidelines for transparent and responsible research dissemination.

Study findings are intended to be disseminated through:

- Peer-reviewed scientific publications
- Presentations at national and international scientific conferences
- Other relevant scientific or stakeholder communications, where appropriate

Authorship of publications will be determined in accordance with the recommendations of the International Committee of Medical Journal Editors (ICMJE).

All results will be reported in an accurate, objective, and balanced manner, irrespective of the nature or direction of the findings.

Where applicable, publication plans and dissemination strategies will be agreed between the study collaborators prior to submission.

9. MANAGEMENT AND REPORTING OF ADVERSE EVENTS

This study is based on secondary use of pseudonymized data from nationwide health registries and does not involve any direct patient contact or intervention.

As no specific medicinal product is being studied and no identifiable patient or reporter information is available, no adverse event reporting or safety signal collection will be conducted as part of this study.

Accordingly, the research report and any resulting publications will not include a summary of adverse events

10. RETENTION OF RECORDS

All records and documentation related to the conduct of this study will be retained by the research-responsible institution in accordance with applicable regulatory requirements and institutional policies.

This includes, but is not limited to:

- Study protocol and amendments
- Statistical Analysis Plan (SAP)
- Data management and analysis documentation
- Derived datasets and analysis datasets (within the secure research environment)
- Statistical programs and outputs
- Final study reports and publications

- Approvals and correspondence with regulatory and data access authorities, including the Regional Committees for Medical and Health Research Ethics (REK) and the Norwegian Health Data Permit Authority (Helsedataservice, HDS)

Pseudonymized individual-level data will be available within the secure research environment for the duration of the approved study period. Following study completion, the data will be retained for an additional five years for quality control and audit purposes, in accordance with applicable regulations and data access agreements.

Study documentation will be retained for at least 10 years after completion of the study, or longer if required by applicable national regulations or data access agreements.

Data will be stored and retained within approved secure research environments in accordance with data protection regulations and governance requirements. Retention and deletion of data obtained from registry sources will be subject to the terms and conditions specified by the relevant data providers and authorities.

No study-related records will be destroyed without appropriate authorization and in accordance with applicable regulations and institutional procedures.

11. REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide. CA Cancer J Clin 2021;71:209–49.
2. Giordano SH. Breast cancer in men. N Engl J Med 2018;378:2311–20.
3. Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol 2023;34:33–57.
4. Finn RS, Martin M, Rugo HS, et al. Palbociclib and letrozole in advanced breast cancer. N Engl J Med 2016;375:1925–36.
5. André F, Ciruelos EM, Juric D, et al. Alpelisib plus fulvestrant for PIK3CA-mutated, HR-positive, HER2-negative advanced breast cancer. N Engl J Med 2019;380:1929–40.
6. Turner NC, Oliveira M, Howell SJ, et al. Capivasertib in hormone receptor-positive advanced breast cancer. N Engl J Med 2023;388:2058–70.
7. Bidard FC, Kaklamani VG, Neven P, et al. Elacestrant in ER-positive, HER2-negative advanced breast cancer. J Clin Oncol 2022;40:3246–56.
8. Jhaveri KL, et al. Advances in oral selective estrogen receptor degraders in breast cancer. Nat Rev Clin Oncol 2023
9. Juric D, et al. Inavolisib plus palbociclib and fulvestrant in PIK3CA-mutated HR-positive, HER2-negative advanced breast cancer. N Engl J Med 2024
10. Schneeweiss S. Learning from big health care data. N Engl J Med 2014;370:2161–63.
11. Booth CM, Karim S, Mackillop WJ. Real-world data: towards achieving the achievable in cancer care. *Nat Rev Clin Oncol*. 2019
12. Furu K, Wettermark B, Andersen M, et al. The Nordic countries as a cohort for pharmacoepidemiological research. Basic Clin Pharmacol Toxicol 2010;106:86–94.

Signature Page for Protocol - SL46796 - GIREDESTRANT - v2 - Norway (Published)
System identifier: RIM-CLIN-1222835

Approval Task	[redacted] 29-Jun-2026 06:34:02 GMT+0000