RESEARCH PLAN (RWE SCIENTIFIC PROJECT)

TITLE:	**Assessment of the association of progression to high-risk varices with death and liver transplantation rates in patients with compensated cirrhosis**
PROJECT NUMBER:	SG46805
VERSION NUMBER:	1.1
STUDIED MEDICINAL PRODUCT:	n/a
RESEARCH INITIATOR:	[redacted] Genentech, a member of Roche [redacted]
AUTHOR:	QUANTIFY RESEARCH [redacted]
DATE FINAL:	Refer to the date on the signature page

QUANTIFY
part of athαgoras

Assessment of the association of progression to high-risk varices with death and liver transplantation rates in patients with compensated cirrhosis

Prepared for: Roche
9 June 2026

Version:	**4.0**
Date:	**9 June 2026**
Status:	**Final**
Roche Study number:	**SG46805**

Authors	[redacted]
Study objectives	The objective of this study is to assess the occurrence of incident liver transplantation and death at pre-specified timepoints in patients with compensated cirrhosis (any etiology) who develop high risk varices for the first time, patients with compensated cirrhosis who experience decompensation events (i.e. variceal bleed, clinically significant ascites or hepatic encephalopathy) or MELD progression for the first time or patients with compensated cirrhosis who do not experience these events. Specifically, the study will: 1. Identify and characterize patients with compensated cirrhosis (any etiology) and first event of any of the below protocol-defined events (PDE), separately for each of the following subcohorts: a. High-risk varices b. Variceal bleed c. Clinically significant ascites d. Hepatic encephalopathy e. Progression from MELD < 12 to MELD ≥ 15 f. Esophagogastroduodenoscopy (EGD) without high-risk varices 2. Estimate risk of liver transplant and all-cause, cardiovascular (CV)-specific, and liver-specific mortality, from index date, defined as first PDE, in the following cohorts: a. Cohort 1: ● High-risk varices ● Variceal bleed ● Clinically significant ascites ● Hepatic encephalopathy ● Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15 b. Cohort 2: ● Variceal bleed ● Clinically significant ascites ● Hepatic encephalopathy ● Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15 c. Reference cohort: ● Esophagogastroduodenoscopy (EGD) without high-risk varices
Country of study	Sweden

Study institutions	Quantify Research AB
Study sponsor	F. Hoffmann-La Roche Ltd
Investigatory substance	NA
Disease area	Liver disease, cirrhosis

Table of Contents

List of abbreviations

Abbreviation	Definition
AASLD	American Association for the Study of Liver Diseases
ALD	Alcohol-associated liver disease
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomic Therapeutic Chemical classification system
BMI	Body mass index
CCI	Charlson Comorbidity Index
CDM	Common data model
CDR	Cause of Death Register
CV	Cardiovascular
DDD	Defined daily dose
DRG	Diagnosis-related group
EASL	European Association for the Study of the Liver
EGD	Esophagogastroduodenoscopy
EHR	Electronic Health Record
EMA	European Medicines Agency
EVL	Endoscopic variceal ligation
FDA	U.S. Food and Drug Administration
FIB-4	Fibrosis-4 Index
GPP	Guidelines for Good Pharmacoepidemiology Practice
HERALD	HEalth outcomes and Risk Assessment in chronic Liver Disease
HRQoL	Health-related quality of life
HRU	Health care resource use
HRV	High-risk varices
ICD-10	The 10th revision of the International Statistical Classification of Diseases and Related Health Problems
IEC	Independent ethics committee
INR	International Normalized Ratio
IQR	Interquartile range
IRB	Institutional review board
ISPE	International Society of Pharmacoepidemiology
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
LISA	Longitudinal Integrated Database for Health Insurance and Labour Market Studies
LSM	Liver stiffness measure
MALO	Major Adverse Liver Outcomes
MASH	Metabolic dysfunction-Associated Steatohepatitis
MASLD	Metabolic dysfunction–associated steatotic liver disease
MELD	Model for End-Stage Liver Disease
NASH	Non Alcoholic SteatoHepatitis
NBHW	National Board of Health and Welfare
NCSP	NOMESCO Classification of Surgical Procedures
NDR	National Quality of Care Diabetes Register
NPR	The National Patient Register
NQR	National Quality Registers
NSBB	Non-selective beta blocker
PBC	Primary biliary cholangitis

PDE	Protocol-defined event
PDR	The Prescribed Drug Register
PEth	Phosphatidylethanol
PPV	Positive predictive value
PSC	Primary sclerosing cholangitis
QC	Quality control
RWD	Real-world data
RWE	Real-world evidence
SAP	Statistical analysis plan
SB	Serum bilirubin
SC	Serum creatinine
SD	Standard deviation
SOPs	Standard Operating Procedures
VCTE	Vibration-Controlled Transient Elastography

Abstract

Study title: Assessment of the association of progression to high-risk varices with death and liver transplantation rates in patients with compensated cirrhosis

Version and date: v 4.0, 9 June 2026

Study background and rationale: The objective of this study is to assess the occurrence of incident liver transplantation and death in patients with compensated cirrhosis (any etiology) who develop high risk varices for the first time, patients with compensated cirrhosis who experience decompensation events (i.e. variceal bleed, clinically significant ascites or hepatic encephalopathy) or MELD progression for the first time or patients with compensated cirrhosis with evidence of esophagogastroduodenoscopy (EGD) without high-risk varices or other events. The aim is to provide population-level real-world evidence (RWE) on the clinical significance of high-risk varices among patients with cirrhosis.

Objective(s): The objective of this study is to assess the occurrence of incident liver transplantation and death at pre-specified timepoints in patients with compensated cirrhosis (any etiology) who develop high risk varices for the first time, patients with compensated cirrhosis who experience decompensation events (i.e. variceal bleed, clinically significant ascites or hepatic encephalopathy) or MELD progression for the first time, or patients with EGD without high-risk varices.

Specifically, the study will:

1. Identify and characterize patients with compensated cirrhosis (any etiology) and first event of any of the below protocol-defined events (PDE), separately for each of the following subcohorts:
    a. High-risk varices
    b. Variceal bleed
    c. Clinically significant ascites
    d. Hepatic encephalopathy
    e. Progression from MELD < 12 to MELD ≥ 15
    f. EGD without high-risk varices
2. Estimate risk of liver transplant and all-cause, cardiovascular (CV)-specific, and liver-specific mortality, from index date, defined as first PDE, in the following cohorts:

Cohort 1:

- High-risk varices
- Variceal bleed
- Clinically significant ascites
- Hepatic encephalopathy
- Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15

Cohort 2:

- Variceal bleed
- Clinically significant ascites
- Hepatic encephalopathy
- Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15

Reference cohort:

- EGD without high-risk varices

Study design: A cohort study utilizing data from 2005 to 2022 in Swedish national and regional registers.

Study population: Adults with compensated cirrhosis are eligible for inclusion in the study. Individuals in each cohort or subcohort will be followed from the first PDE until death, liver transplant, end of data, or emigration.

Exposures and outcomes: Exposure is defined as the first PDE. Outcomes defined as occurrence of liver transplant, CV-related mortality, liver-related mortality, and all-cause mortality. Covariates may include age, sex, and cirrhosis etiology.

Data sources: The study will conduct secondary data analysis on data from the HEalth outcomes and Risk Assessment in chronic Liver Disease (HERALD) research platform. Data in the HERALD research platform comes from the Swedish national registers, including the National Patient Register (NPR), Prescribed Drug Register (PDR), National Quality of Care Diabetes Register (NDR), Cause of Death Register (CDR), and the Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA). The national registers contain data for all residents of Sweden and are linked at the individual level using the national identification number. Follow-up is complete without attrition of individuals over time. Additionally, the HERALD research platform includes patient-level data from individuals who received care in the Stockholm (capital) region, where data has been extracted from electronic health records (EHR). In contrast to the nationwide sources, data from the Stockholm region includes lab data. The HERALD research platform includes data up to the end of 2022.

Study size: As of June 2023, the HERALD research platform includes over 2 million individuals. There are approximately 92,363 patients with cirrhosis of whom approximately 58,197 (63%) have at least one PDE other than 'EGD with no high-risk varices' on or after the date of cirrhosis diagnosis in the HERALD dataset.

Data analysis: Descriptive statistics will be generated to summarize patient characteristics.

[redacted]

1. Amendments

Amendment number	Date	Section of protocol	Reason for amendment
1	n/a		

2. Study milestones

Milestone	Planned date	Comment
1 Draft protocol	2025-12-12	
2 Final protocol	2026-06-12	
3 Preliminary results	2026-07-10	
4 Study report	2026-09-30	

3. Rationale and background

3.1 Introduction

The objective of this study is to assess the occurrence of incident liver transplantation and death at pre-specified timepoints in patients with cirrhosis (any etiology) who develop high risk varices for the first time. Comparison will be made between cohorts containing individuals who experience traditional decompensating events, MELD progression, or high-risk varices (HRV) for the first time and those with compensated cirrhosis who have not developed high-risk varices or decompensating events or MELD progression. This comparison will then be contrasted with another comparison made between patients experiencing strictly traditional decompensating events or MELD progression and those without HRV or traditional decompensating events. The aim is to provide population-level real-world evidence (RWE) on the clinical significance of developing high-risk varices in patients with cirrhosis.

This combined protocol and statistical analysis plan (SAP) detail the study plan. The protocol will be finalized prior to initiation of data analysis. Any changes or amendments will be documented.

3.2 Study rationale

Cirrhosis represents the end stage of chronic liver disease and is associated with significant morbidity and mortality. Once patients experience traditional decompensating events, their prognosis worsens considerably, with median survival dropping from over 12 years in compensated cirrhosis to approximately 2 years in decompensated disease (1). Key decompensating events—such as ascites, variceal hemorrhage, and hepatic encephalopathy as well as worsening liver function reflected by Model for End-Stage Liver Disease (MELD) progression, are well established predictors of poor outcomes, including increased risk of liver transplantation and death (2).

[redacted]

The goal of this study is to provide such evidence using data from the Swedish population registers from 2005 to 2022 in the HERALD (HEalth outcomes and Risk Assessment in chronic Liver Disease) research platform. The HERALD research platform, an RWE research initiative by Quantify Research, in collaboration with Karolinska Institute, focuses on chronic liver diseases in Sweden and is uniquely positioned to provide RWE through its extensive data. Results using different parts of the HERALD data have been presented as abstracts and posters at European Association for the Study of the Liver (EASL) and American Association for the Study of Liver Diseases (AASLD) (5-10), as well as in several peer-reviewed scientific papers already published (11, 12) or currently under review.

The HERALD research platform includes data from the national and regional registers in Sweden. Swedish national registers have complete data coverage of all Swedish residents with attrition only in the instance of emigration. Data includes all inpatient and specialized outpatient care visits from 1998/2001 onward, as well as all outpatient pharmacy prescription use from 2005, up to Q3 2022. Regional data from Stockholm contains information on laboratory measurements from January 1st, 2010, to October 26th, 2020. The study period ending in 2022 is also consistent with clinical interpretability of exposure and outcome definitions in the context of evolving standards of care.

[redacted]

[redacted]

Register data is available at the individual level, as it is linked across registers using the unique identification number assigned to each resident. Validation studies have shown that the data is of very high quality with positive predictive values (PPV) ranging from 72-93% for diagnosis codes and 97% for procedural codes (15). Capture of liver disease and conditions is particularly high. Indeed, validation studies have shown that the PPV for cirrhosis is over 90% in Swedish register data (17) while PPV for metabolic dysfunction-associated steatohepatitis (MASH) ranges from 82-95% (16).

[redacted]

This approach will enable contextualization of high-risk variceal progression relative to decompensating events and to compensated cirrhosis without high-risk varices, providing evidence on the prognostic value of progression to high-risk varices in the context of natural history of cirrhosis.

4. Objectives

4.1 Primary objectives

The objective of this project is to assess the risk of incident liver transplantation and death at pre-specified timepoints in patients with cirrhosis (any etiology) who develop high risk varices for the first time and individuals who experience decompensating events for the first time (i.e., variceal bleed, clinically significant ascites, hepatic encephalopathy) or those who do not develop high-risk varices. Specifically, the study will:

1. Identify and characterize patients with compensated cirrhosis (any etiology) and first event of any of the below protocol-defined events (PDE):
    a. High-risk varices
    b. Variceal bleed
    c. Clinically significant ascites
    d. Hepatic encephalopathy
    e. Progression from MELD < 12 to MELD ≥ 15
    f. Esophagogastroduodenoscopy (EGD) without high-risk varices
2. Estimate risk of liver transplant and all-cause, cardiovascular (CV)-specific, and liver-specific mortality, from index date, defined as first PDE, in the following cohorts:

    Cohort 1:
    - High-risk varices
    - Variceal bleed
    - Clinically significant ascites
    - Hepatic encephalopathy
    - Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15

    Cohort 2:

- Variceal bleed
- Clinically significant ascites
- Hepatic encephalopathy
- Progression from Model for End-Stage Liver Disease (MELD) < 12 to MELD ≥ 15

Reference cohort:

- Esophagogastroduodenoscopy (EGD) without high-risk varices

4.2 Exploratory objectives

[redacted]

5. Research methods

5.1 Study design

This is a Swedish, observational, cohort study on the risk of liver transplantation and mortality in adult patients with compensated and decompensated cirrhosis. The study will conduct secondary data analysis on data from the HERALD research platform. Data in the HERALD research platform origins from the Swedish national registers, including the National Patient Register (NPR), Prescribed Drug Register (PDR), National Quality of Care Diabetes Register (NDR), Cause of Death Register (CDR), and the Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA). Follow-up is complete without attrition of individuals over time and extends up to the end of 2022. Additionally, the HERALD research platform includes patient-level data from individuals who received care in the Stockholm (capital) region, where data has been extracted from electronic health records (EHR). In contrast to the nationwide sources, data from the Stockholm region includes lab data. Further details on data sources are described in section 5.3.

In this study, two populations will be investigated: a national population utilizing national administrative registers and a Stockholm-based population, consisting of patients who received care in the Stockholm region, for whom regional primary care and laboratory data are available in addition to the national registers. Patients with cirrhosis will be assigned an index date at study entry, which will be the date of first incident PDE (see section 5.1.2.1) observed during the inclusion period. For the national population, patients are identified in national administrative registers and the regional EHR where possible in Sweden between 01 January 2006 and 01 November 2021 and studied from 01 January 2005 to 31 October 2022 (to allow for a minimum of one year look back and one year follow-up with complete coverage of national drug prescriptions). The Stockholm population will be identified from the national administrative registers and the regional EHR and laboratory data between 01 January 2011 and 26 October 2020 and studied between 01 January 2010 and 31 October 2022, to reflect the maximal span of regional lab data available for stratification and allow for a minimum of 2 years of follow-up. Notably, liver- and CV-related death will be followed until 31 December 2021 to account for completeness concerns for the last year of cause of death data. [redacted]

[redacted]

Baseline patient characteristics including demographics, and biomarkers data in the case of the Stockholm population, will be observed during corresponding baseline periods, see section 5.1.1 for definitions for all study periods and additional details. Mortality and risk of liver transplantation will be assessed during prespecified time windows (1-, 3-, and 5-years after index) and overall in patients experiencing the different PDE using survival analysis. Risks of developing outcomes will be estimated in adjusted survival models. A sensitivity analysis to account for patients with multiple contemporaneous PDE (i.e., within 30 days of each other) during follow-up may be conducted within a competing risks survival analysis framework if at least 5% of the cohort has multiple contemporaneous PDEs. The study design is summarized in Figures 1 and 2.

Figure 1. Study design overview

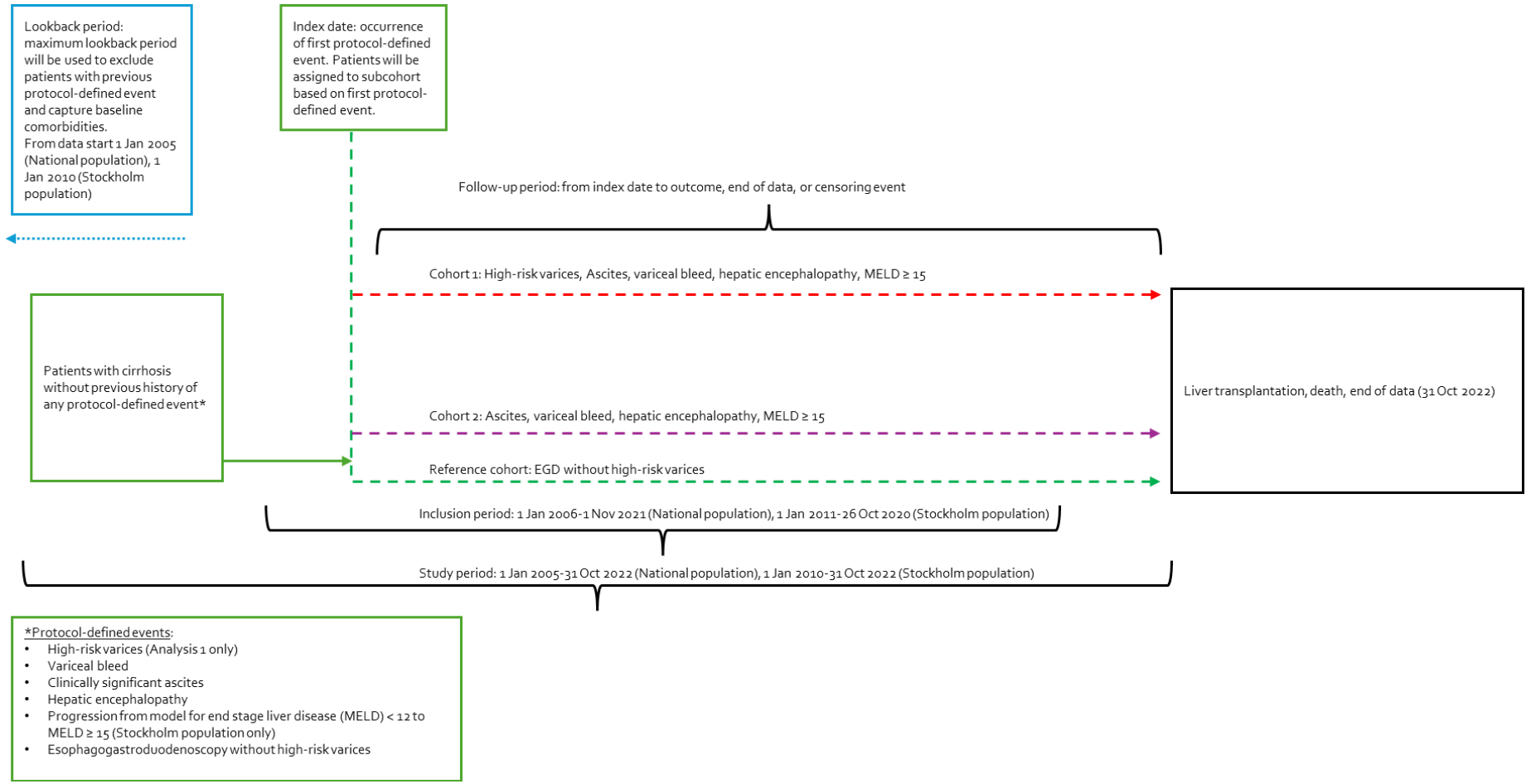
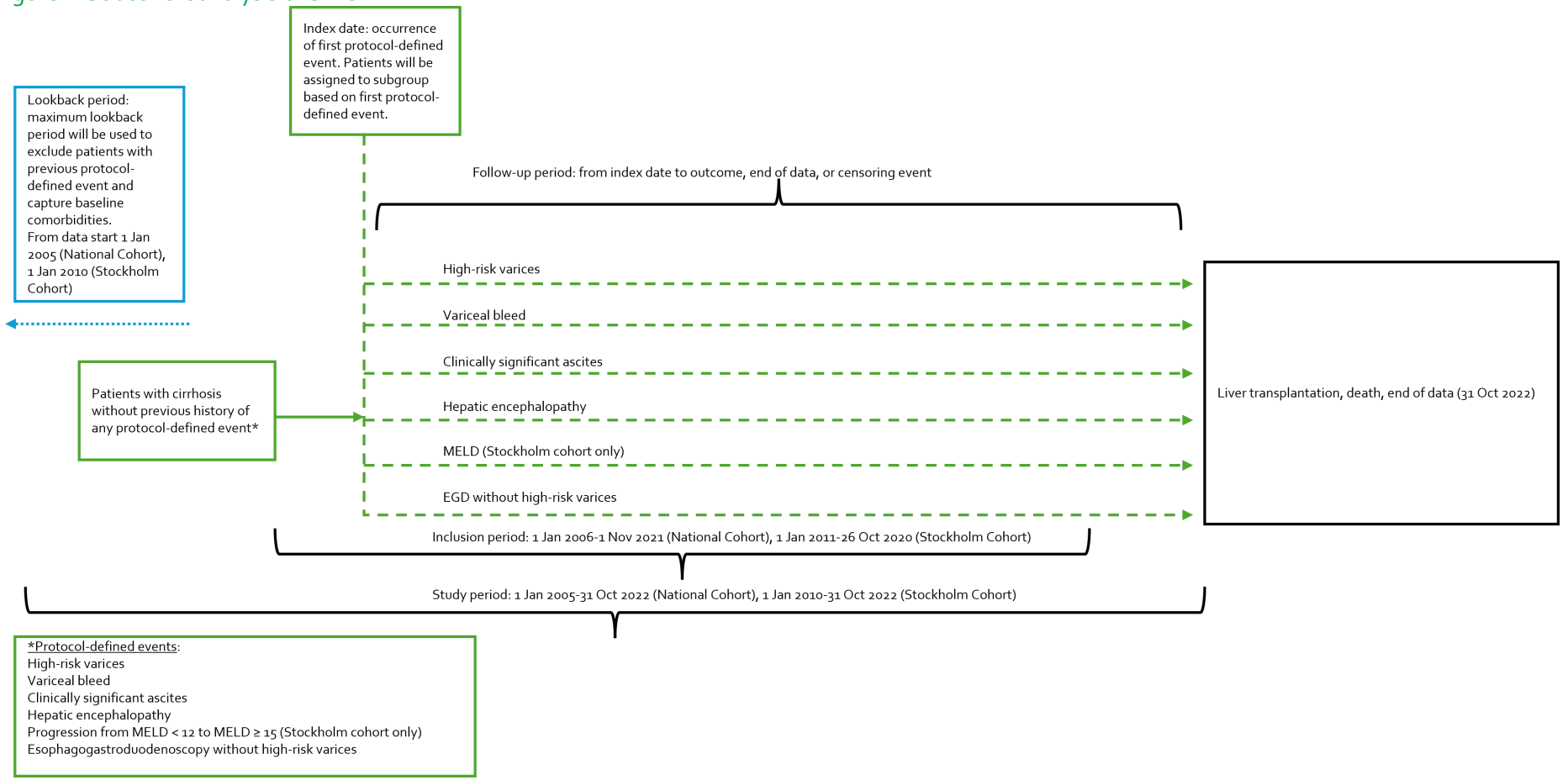


Figure 1. Subcohort analysis overview



5.2 Study setting

5.2.1 Study periods

Study periods are detailed in Table 1. HERALD data was extracted beginning in 1998, however, the Swedish PDR was initiated in 2005, thus the study period will begin then, to allow for capture of relevant medication use. Data in the Stockholm region is available from 2001, but coverage of laboratory measures is poor prior to 2010, thus the study period for the Stockholm population will begin in 2010. The end of extraction in the HERALD platform is 2022, thus the end of the study period. For an overview of the study time periods, see Table 1.

Table 1. Study periods

Term	Definition	Value
Study period for national population	The date of the first data record to the date of the last data record.	01 January 2005 – 31 October 2022
Study period for Stockholm population	The date of the first data record to the date of the last data record.	01 January 2010 – 31 October 2022
Inclusion period for national population	The time period in which patients in the national population are screened for inclusion to the study and during which the patient selection criteria should be met. Ensures a minimum of one year look back period and one year follow up period.	01 January 2006 – 01 November 2021
Inclusion period for Stockholm population	The time period in which patients in the Stockholm population are screened for inclusion to the study and during which the patient selection criteria should be met. Ensures a minimum of one year look back period and two-year follow up period.	01 January 2011 – 26 October 2020
Index date	The date of first recorded PDE (see section 5.2.2.1) during the inclusion period. • [redacted] • [redacted]	Any date within an inclusion period

Term	Definition	Value
	• [redacted]	
Lookback period	The time period prior to the index date during which covariate data are collected.	Lookback period dependent on condition
Lookback period for comorbidities	The time period prior to the index date during which data on comorbidities are collected.	Maximum lookback period to be used; minimum required lookback period is one year prior to index [-365,0)
Lookback period for prior to PDE	The time period prior to the index date during which non-incident patients are excluded from the cohort.	01 January 2005 (National)/01 January 2010 (Stockholm) to one day prior to index date.
Follow-up period	The time period on and after the index date during which outcome data are collected.	From index date up until end of study period, emigration or death (censoring events)

**The inclusion period for the Stockholm population is truncated to the exact cut-off of availability of laboratory data in Stockholm, based on HERALD extraction dates, which is 26 October 2020. Maximizing the inclusion period ensures exhaustive identification of patients with [redacted] data, see section 5.1.2.1.*

5.2.2 Study population and study cohorts

5.2.2.1 Inclusion criteria

Study patients must fulfil each of the inclusion criteria outlined below. Individuals will be considered for inclusion if they have compensated cirrhosis (see Table 2) and no prior evidence of any of the PDEs. For the operational definitions of each criterion, see Table 2 and Table 3. All codes also include their respective subcodes. For PDEs with multiple conditions (e.g., diagnosis code plus treatment code), the date of the event highest in the multi-level hierarchy of the bulleted list determines the index date.

1.1.1.1.1 National Population

- Patient has at least one (≥1) primary or secondary diagnosis of compensated cirrhosis (defined in Table 2) recorded in the national in- or outpatient secondary care registers or in the regional EHRs during the inclusion period between 01 January 2006 and 01 November 2021:

AND

- Patient has a first occurrence of any of the following PDEs (complete definitions in Table 3), during the inclusion period between 01 January 2006 and 01 November 2021:
  - [redacted]
    - [redacted]
      - [redacted]

- [redacted]
- Variceal bleed
  - Patient has variceal bleeding.
    - Patient has variceal bleeding OR
    - Patient has gastric varices
      - AND has hematemesis, melena or gastrointestinal hemorrhage
- Clinically significant ascites:
  - Patient has ascites from any cause.

  AND
    - Patient does not have diagnosis of heart failure or cancer within ± 30 days of the diagnosis of the ascites.
- Hepatic encephalopathy (proxy):
  - Hospitalization

  AND on the date thereof, or within ± 30 days
    - Dispensation for rifaximin
- [redacted]
  - [redacted]

AND

- Patient is at least eighteen (≥18) years of age on the date of the first PDE used for inclusion.

1.1.1.1.2 Stockholm population

- Patient has at least one (≥1) primary or secondary diagnosis of compensated cirrhosis (see Table 2) recorded in the national in- or outpatient secondary care registers or in the regional EHRs on the index date or during lookback

AND

- Patient has any of the following PDEs (see Table 3), with the first occurrence thereof defined as the index date, during the inclusion period between 1 January 2011 and 26 October 2020:
  - [redacted]
    - [redacted]
    - [redacted]
    - [redacted]
  - Variceal bleed
    - Patient has variceal bleeding.
      - For gastric variceal bleed:
        - Patient has gastric varices
        - AND has hematemesis, melena or gastrointestinal hemorrhage
  - Clinically significant ascites:
    - Patient has ascites from any cause.

    AND
    - Patient does not have diagnosis of heart failure of cancer at or within ± 30 days of the diagnosis for ascites.
  - Hepatic encephalopathy (proxy):
    - Hospitalization

    AND on the date thereof, or during the preceding or following month
      - Dispensation of rifaximin

- [redacted]
  - [redacted]
  - [redacted]
- [redacted]
  - [redacted]

AND

- Patient has at least one observation in the TakeCare data

AND

- Patient is at least eighteen (≥18) years of age at the date of the first diagnosis used for inclusion.

5.2.2.2 *Exclusion criteria*

Study subjects must not fulfil any of the following exclusion criteria:

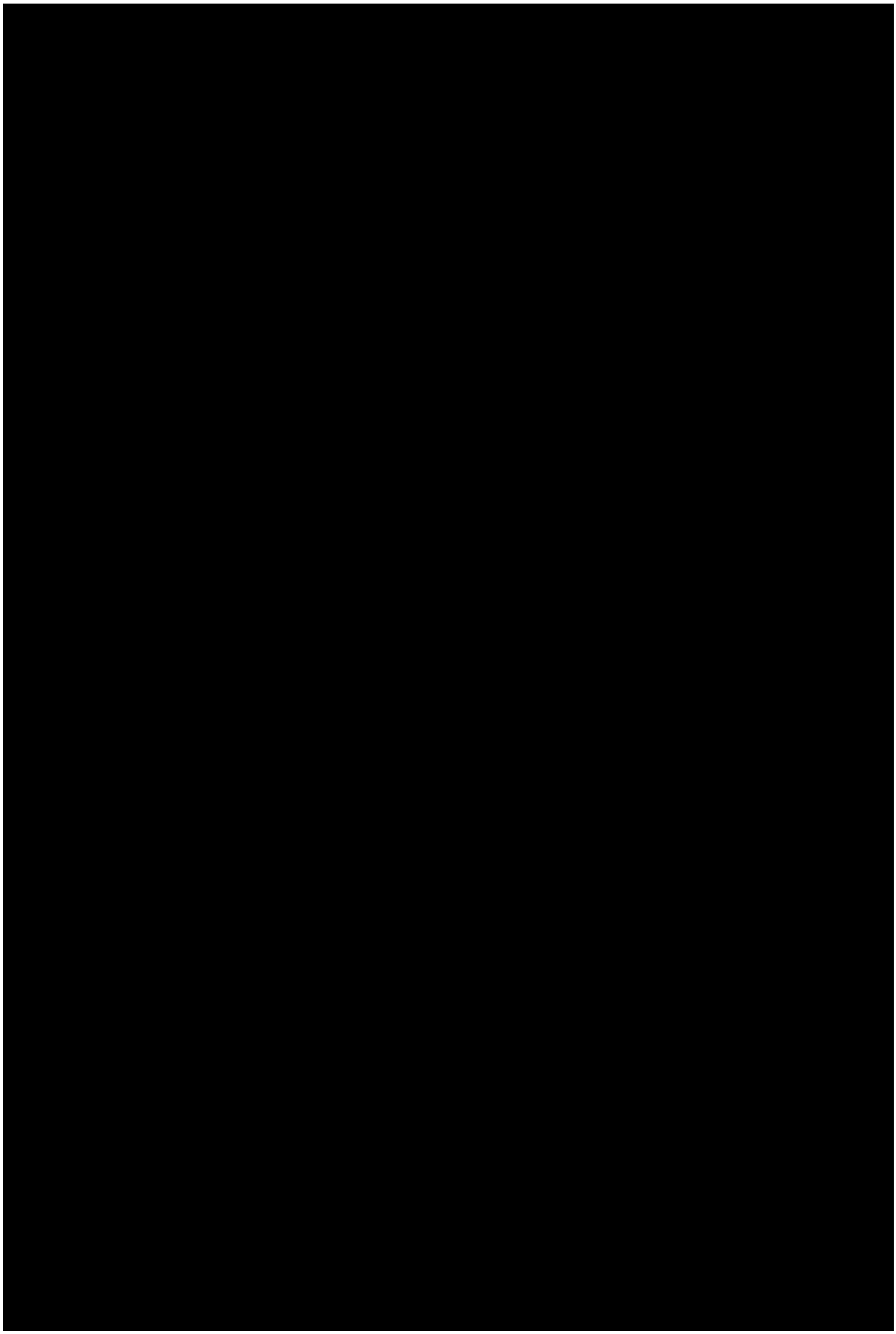
1.1.1.1.3 National population

- Patient has evidence of decompensated cirrhosis [redacted] prior to their first PDE.

1.1.1.1.4 Stockholm population

- Patient has evidence of decompensated cirrhosis [redacted] prior to their first PDE.

[redacted]



5.2.2.3 Cohorts

The incident PDE of the study patient will determine their subcohort, i.e. the PDEs and the corresponding subcohorts are mutually exclusive, with no overlapping patients. Accordingly, the cohorts of the study are patients with compensated cirrhosis and first occurrence of the following (see [redacted] for operational definition of PDEs):

- Cohort 1 group consisting of patients whose first PDE during follow-up was any of:
  - High-risk varices: [redacted]
  - Variceal bleeding
  - Clinically significant ascites:
  - Hepatic encephalopathy
  - Progression from MELD <12 to ≥15 ([redacted])
- Cohort 2 group consisting of patients whose first PDE during follow-up was any of:
  - Variceal bleeding
  - Clinically significant ascites:
  - Hepatic encephalopathy
  - Progression from MELD <12 to ≥15 ([redacted])
- Reference cohort
  - EGD without high-risk varices.

Each objective (i.e. patient characteristics, liver transplantation rates, and mortality) will be reported separately for each of the three cohorts. A further subcohort analysis will be conducted where the objectives will be individually analysed for each of the PDEs comprising the cohorts.

5.3 Variables

This section describes the variables used in this study, whereas the data sources are described in section 5.3. For a detailed list of all variables and their operational definitions, see Table 4. Pharmacy dispensed medications and hospital administrations will be identified using ATC codes and diagnoses will be defined using ICD-10 codes. Procedures are identified using the NOMESCO Classification of Surgical Procedures (NCSP) codes.

5.3.1 Outcomes

This study is based on descriptive and survival outcomes. Patient characteristics will be captured in Objective 1 and the specific variables to be described are defined in Table 4. The outcomes captured in Objective 2, which investigates occurrence of liver transplant, all-cause mortality, CV-related mortality, and liver-related mortality in survival analyses, are defined in Table 5. Outcomes relating to the exploratory objective are defined in Table 6.

Table 4. Objective 1: Definitions of patient characteristics

Variable	Data source(s)	Operational definition
Age	NPR/EHR	Age in years at index date (continuous; years).
Sex		Sex (dichotomous; male/female) at index date.
Region of residence		Region (Svealand, Norrland, Götaland)
Cirrhosis etiology		
Autoimmune	NPR/EHR	ICD-10 code: K76.0 any point prior to or on the date of the cirrhosis diagnosis
Hemochromatosis		ICD-10 code: E83.1 any point prior to or on the date of the cirrhosis diagnosis
Alcohol-associated liver disease (ALD)		ICD-10 code: K70 any point prior to or on the date of the cirrhosis diagnosis
Metabolic dysfunction–associated steatotic liver disease (MASLD)/Metabolic dysfunction-associated steatohepatitis (MASH)		ICD-10 code: K76.0 or K75.81 any point prior to or on the date of the cirrhosis diagnosis
Viral hepatitis		ICD-10 code: B18.1 E/G, B18.2 E/G, or B18 any point prior to or on the date of the cirrhosis diagnosis
Primary biliary cholangitis (PBC)		ICD-10 code: K74.3 any point prior to or on the date of the cirrhosis diagnosis
Primary sclerosing cholangitis (PSC)		ICD-10 code: K83.01 any point prior to or on the date of the cirrhosis diagnosis
Comorbid conditions		
Body mass index (BMI)	EHR, NDR, NPR	BMI within 1 year prior to index date (continuous; score) (kg/m2)
Overweight		BMI > 25 kg/m2 but ≤ 30 kg/m2 within 1 year prior to index date index date
Obesity		BMI > 30 kg/m2 or ICD-10 code E66 within 1 year prior to index date index date
Baseline comorbidities (Charlson Comorbidity Index [CCI])		Captured with ICD-10 codes at any point prior to or on the date of cirrhosis diagnosis. See Section 5.3.3 for further details.
Biomarkers (Stockholm population only)		
Alanine aminotransferase (ALT)	EHR	ALT value closest to index date, up to 1-year prior and 3-months after index
Aspartate aminotransferase (AST)		AST value closest to index date, up to 1-year prior and 3-months after index
AST/ALT ratio		$\frac{AST_{\left[\frac{U}{L}\right]}}{ALT_{\left[\frac{U}{L}\right]}}$
Fibrosis-4 Index (Fib-4)		$\frac{age_{[years]} \times AST_{\left[\frac{U}{L}\right]}}{PLT_{\left[\frac{109}{L}\right]} \times ALT_{\left[\frac{U}{L}\right]} \times \frac{1}{2}}$

Variable	Data source(s)	Operational definition
Liver stiffness measure (LSM)**	FibroScan machine*/EHR	Vibration-Controlled Transient Elastography (VCTE) value closest to index date, up to 1-year prior and 3-months after index
		<8 kPa
		8-12 kPa
		≥12 kPa
		Missing
Platelet count	EHR	Platelet count closest to index date, up to 1-year prior and 3-months after index
[redacted]		[redacted]
Phosphatidylethanols (PEth) level		PEth value closest to index date, up to 1-year prior and 3-months after index

**Available for a subset of the Stockholm population.*

Table 5. Objective 2: Definitions of clinical outcomes

Variable	Data source(s)*	Operational definition		Description
		Classification system	Code	
All-cause mortality	CDR	ICD-10	Any	All-cause death
CV-related mortality		ICD-10 including main and contributing codes	I05-I09	Chronic rheumatic heart disease
			I11	Peripheral arterial disease
			I13	Hypertensive heart disease with heart and kidney failure
			I20	Angina pectoris
			I21	Acute heart failure
			I22	ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction. Subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction
			I23	Certain current complications following ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (within the 28-day period)
			I24	Other acute ischemic heart disease
			I25	Chronic ischemic heart disease
			I27	Other pulmonary heart disease
			I34	Nonrheumatic valve disorders
			I35	Nonrheumatic aortic valve disorders
			I36	Nonrheumatic tricuspid valve disorders
			I37	Pulmonary valve disorders

Variable	Data source(s)*	Operational definition		Description
		Classification system	Code	
			I38	Endocarditis, valve unspecified
			I39	Endocarditis and heart valve disorders in diseases classified elsewhere
			I42	Cardiomyopathy
			I50	Heart failure
			I51-59	Other forms of heart disease
			I60-69	Cerebrovascular disease
			I70-79	Diseases of arteries, arterioles, and capillaries
			I80-89	Diseases of veins, lymphatic vessels, and lymph nodes, not elsewhere classified
Liver-related mortality		ICD-10 including main codes	K72.0	Acute and subacute liver failure
			K72.1	Chronic hepatic failure
			K72.9	Hepatic failure, unspecified
			K74.6	Cirrhosis, unspecified
			K76.6	Portal hypertension
			K76.7	Hepatorenal syndrome
			R18.9	Ascites
			I85.0, I98.3	Esophageal varices with bleeding
			I85.9, I98.2	Esophageal varices without bleeding
Liver transplantation	NPR, EHR	NCSP	JJC00	Homologous liver transplant
			JJC10	Homologous partial liver transplant
			JJC20	Homologous liver transplant, living donor
			DJ005	Homologous auxiliary liver transplantation
			DJ006	Homologous auxiliary liver transplantation, with living donor
			JJC30	Heterologous liver transplantation
			JJC40	Heterologous partial liver transplant
		ICD-10	T86.4	Complications of liver transplant (malfunction and rejection of liver transplant)
			Z94.4	Liver transplant status

Table 6. Exploratory objectives: Definitions of proxy quality of life outcomes

Variable	Data source(s)	Operational definition		Description
		Classification system	Code	
Work absence	LISA	Swedish Social Insurance Agency registration	NA	Registered sick leave or early retirement
Depression	NPR, EHR	ICD-10	F32-F33	≥1 Primary or secondary diagnosis of Major depressive disorder
	PDR, EHR	ATC	N06AA	≥1 delivered prescription of Tricyclic antidepressants
			N06AB	≥1 delivered prescription of Selective serotonin reuptake inhibitors

Variable	Data source(s)	Operational definition		Description
		Classification system	Code	
			N06AF/AG	≥1 delivered prescription of Monoamine oxidase inhibitors
			N06Ax	≥1 delivered prescription of Other antidepressants
Anxiety disorders	NPR, EHR	ICD-10	F41	≥1 Primary or secondary diagnosis of Generalized anxiety/panic
	PDR, EHR	ATC	N05B	≥1 delivered prescription of Anxiolytics
Adjustment disorder/stress reaction	NPR, EHR	ICD-10	F43	≥1 Primary or secondary diagnosis of Reaction to severe stress
Fatigue	NPR, EHR	ICD-10	R53	≥1 Primary or secondary diagnosis of Fatigue/asthenia
Sleep disturbance	NPR, EHR	ICD-10	F51	≥1 Primary or secondary diagnosis of Nonorganic sleep disorders
			G47	≥1 Primary or secondary diagnosis of Sleep disorder
	PDR, EHR	ATC	N05C	≥1 delivered prescription of Hypnotics and sedatives

5.3.2 Exposures

Exposure is the first PDE experienced by the individual patient. PDEs are defined by diagnosis, procedure, and ATC codes in nationwide secondary care and regional EHRs, dispensed prescriptions in the PDR and administered medications in the EHRs, and laboratory values recorded in the TakeCare data (Stockholm population), see section 5.1.2 for the definition of the PDEs and Table 3 for the technical operationalization of the definitional criteria thereof.

5.3.3 Covariates

5.3.3.1 Charlson Comorbidity Index (CCI)

A version of the CCI developed for use in Swedish register data (17) will be used to assess comorbidities at index date. A lookback period using all available data prior to index will be applied. The CCI will be presented as mean, median, standard deviation, interquartile range and each condition will be presented independent as count and percent. The conditions, codes, and weighting are presented in Table 7.

Table 7. Charlson Comorbidity Index conditions, codes, and weighting

Condition	ICD-10	Charlson index score
Myocardial infarction	I21, I22, I25.2	1
Congestive heart failure	I11.0, I13.0, I13.2, I25.5, I42.0, I42.6-I42.9, I43, I50	1
Peripheral vascular disease	I70, I71, I73.1, I73.8, I73.9, I77.1, I79.0, I79.2, K55	1
Cerebrovascular disease	G45, I60, I61-I64, I67, I69	1
Dementia	F00-F03, F05.1, G30, G31.1, G31.9	1
Chronic obstructive pulmonary disease	J43, J44	1
Other chronic pulmonary disease	J41, J42, J45-J47, J60-J69, J70	1

Condition	ICD-10	Charlson index score
Rheumatic disease	M05, M06, M12.3, M07.0-M07.3, M08, M13, M30, M31.3-M31.6, M32-M34, M35.0, M35.1, M35.3, M45, M46	1
Peptic ulcer disease	K25-K28	1
Mild liver disease	B15-B19, K70.3, K70.9, K73, K74.6, K75.4	1
Diabetes without chronic complication	E10.0, E10.1, E11.0, E11.1, E12.0, E12.1, E13.0, E13.1, E14.0, E14.1	1
Diabetes with chronic complication	E10.2-E10.5, E10.7, E11.2-E11.7, E12.2-E12.7, E13.2-E13.7, E14.2-E14.7	2
Hemiplegia or tetraplegia	G11.4, G80-G82, G83.0-G83.3, G83.8	2
Renal disease	I12.0, I13.1, N03.2-N03.7, N05.2-N05.7, N11, N18, N19, N25.0, Z49, Z94.0, Z99.2, Q61.1-Q61.4	2
Any malignancy	C00-C09, C10-C19, C20-C29, C30-C39, C40-C49, C50-C59, C60-C69, C70-C76, C81-C86, C88-C89, C90-C97	2
Moderate to severe kidney disease	N03.2-N03.7, N05.2-N05.7, N11, N18, N19, N25.0, I12.0, I13.1, Q61.1-Q61.4, Z49, Z94.0, Z99.2	2
Moderate to severe liver disease	R18, I85.0, I85.9, I98.2, I98.3	3
Metastatic cancer	C77-C79, C80	6
HIV/AIDS	B20-B24, F02.4, O98.7, R75, Z21.9, Z71.1	6

Certain covariates (e.g., age, sex) may be considered in adjusted survival models to allow for comparison of outcomes accounting for these potentially confounding factors. Details on adjusted survival models are described in section 5.7.2.2. Any covariates included will be defined as described in Table 4.

5.4 Original data sources

The Swedish national health care registries offer high quality data for research with nation-wide coverage and close to complete data for most available variables. Patient-level data from each register is available for research upon prior ethics approval and individual patients can be linked between all data sets via the personal identifier given to all Swedish citizens. The national health care registries to be used in this study are the NPR, PDR, CDR, and LISA.
Regional data for the Stockholm population has been extracted from the primary and specialist care EHR system named TakeCare, and data from the FibroScan (VCTE) machine at the Hepatology clinic at Karolinska University hospital in Huddinge, Stockholm. The data sources are summarized in Table 8.

Table 8. Data sources

Type of register	Register (data holder)	Main variables [examples]
In- and outpatient specialist care	National Patient Register, Swedish Cancer Register (NBHW).	Age, sex, district, primary- and secondary diagnosis codes (ICD-10) and procedure codes (KVÅ), dates and type(s) of visit, admission and discharge date, diagnoses related groups (DRG) and liver cancer related information from the Cancer Register
Primary care	EHR – Take Care (SLSO)	Age, sex, district, primary- and secondary diagnosis codes (ICD-10) and procedure codes, dates and type(s) of visits
Laboratory data and lifestyle factors	EHR – Take Care (SLSO)	Age, sex, liver-related laboratory values (e.g. ALT, AST, platelet count) and lifestyle factors (e.g. weight) in both primary and secondary care settings

FibroScan data	EHR - Take Care (SLSO) and FibroScan machines at the hepatology clinic at Karolinska University Hospital	FibroScan test scores
Mortality	Cause of death register (NBHW)	Age, sex, date of death, cause of death
Prescription medicines	Prescription register (NBHW)	Age, sex, district, dispensing and prescription date, ATC-codes and type of medication, number of packages, package size and strength, prescribed and package defined daily dose (DDD) value and unit, pharmacy retail price and product number
Socioeconomics	LISA (Statistics Sweden)	Emigration status, immigration status
Care in diabetes patients	National Diabetes Register	Diagnoses, care, prescriptions, and lab data available nationally for patients with diabetes.

**NBHW, National Board of Health and Welfare; SLSO, Stockholm region*

5.4.1 Electronic health records for primary and specialist care

A majority of health care providers (both publicly- and privately-owned providers) in the Stockholm region (SLSO) utilize the TakeCare EHR system. These health care providers report in- and outpatient specialist care data (e.g., diagnoses [ICD-10], procedures [KVÅ]), and information on types of visits in both primary- and secondary care), prescription medicines and lifestyle data, such as BMI (height and weight), FibroScan data, etc. In addition, TakeCare stores information on patient laboratory values that were taken at the three largest laboratory organizations in the Stockholm region: Karolinska University lab, Unilabs, and Aleris. These laboratories automatically report laboratory values into TakeCare. TakeCare subsequently stores all EHR data on a server named Intelligence, which is both coordinated and administered by SLSO. The Intelligence server contains data from 2001 and onwards, although registration in the system was low in the first decade (2001-2009). The Stockholm population will be required to have observations in the EHR data.

5.4.1.1 Information on FibroScan test scores

In HERALD, FibroScan test scores have been obtained from a FibroScan machine located at the Hepatology clinic at Karolinska University hospital in Huddinge (Stockholm). The FibroScan machine contains a register with FibroScan test scores. For this study, the LSM (liver stiffness measure) will be used, and it measures the level of liver stiffness in kPa.
VCTE data was extracted from a single M-probe device FibroScan (Echosense) at the Unit of Hepatology (specialist care) at Karolinska University Hospital and linked to the HERALD dataset together with additional VCTE data available in the EHR from the full Stockholm region. Data from the single FibroScan machine at the Karolinska University Hospital were extracted from 2015 and onwards, while data from the EHR were extracted from 2010 and onwards, when coverage was more complete. The data extracted from these two different sources are not expected to be complete for all patients treated in the region.

5.4.2 The National Patient Register

The National Patient Register held by the National Board of Health and Welfare (NBHW) contains geographical, administrative, and medical patient data for both inpatient and outpatient specialist care in Sweden (18). The register contains main and contributing diagnosis codes (ICD-10), which refer to both the primary reason for the care visit and contributing or underlying conditions for each

admission and outpatient specialist visit as well as procedure codes. The outpatient register includes physician visits only, and records of surgical procedures based on the NOMESCO Classification of Surgical Procedures (NCSP) are included in the register. The register has nationwide coverage from 1987 as well as complete outpatient data from 2001. The study will encompass complete data on all inpatient and outpatient hospital care between 2001 and the date of latest available data extracted onto the HERALD platform, i.e., 31 October 2022 (including diagnosis codes related to each hospital admission and specialist visit, all health care procedures performed in a hospital setting). Data on e.g., comorbidities, risk factors, and demographics will be captured in the study by utilizing this register. The register will be used to identify the patient cohorts and subcohorts, as well as assess the rates of liver transplantation.

5.4.3 The Prescribed Drug Register

The Prescribed Drug Register (PDR) held by the National Board of Health and Welfare contains information on all filled prescriptions (i.e., dispensed at a pharmacy in Sweden). The register does not include over-the-counter medicines or hospital-administered drugs. Data can be captured on prescribed drugs using ATC codes, prescription date, dispensing date, dose, pack size, healthcare professional issuing the prescription, as well as sex, age, and residency of the patients. The register is updated monthly and complete data from the register is available from 1 July 2005. This study will encompass complete data on all pharmacy dispensed drugs between 2005 and the date of the latest available data (30 September 2022). The register will be used to evaluate treatment patterns, potential protective factors and estimate resource use and costs in terms of drug use.

5.4.4 The Cause of Death Register

The CDR is held by the NBHW and will be used to assess all-cause, cardiovascular-, and liver-related mortality, as well as to remove patients (censor) from the analysis once they are no longer under observation. The statistics on causes of death comprise all deaths, covering Swedish residents, irrespective of whether the deaths occurred in Sweden or not. The quality of the statistics varies, depending on the examinations made to define the underlying cause of death and due to changes in the classification system or the processing methods. Data is available from 1961.

5.4.5 The Longitudinal Integrated database for health insurance and labour market studies (LISA)

The LISA register is held by Statistics Sweden and includes yearly data for the adult Swedish population from 1990 onwards, concerning socioeconomic factors. Statistics Sweden also provide data regarding migration dates and region of residency at the end of each year used for censoring in the present study.

5.4.6 The National Diabetes Register (NDR)

The NDR is a national quality of care register held and maintained by Region Västra Götaland. It was established in 1996 and covers more than 90% of individuals with diabetes in Sweden (19). The register includes demographic, lab value, body composition (e.g., BMI), treatment, comorbidity, and hereditary data. The register is updated annually, and this study will utilize data from 1998-2021.

5.5 Data linkage

The research data is pseudonymized by Statistics Sweden and NBHW. Data is linked at the individual level using the personal identification number unique to every resident of Sweden. Data is linked

across all sources, including data from NBHW, TakeCare, Cosmic and the Fibroscan machine, and Statistics Sweden.

5.6 Study size

The HERALD research platform includes over 2 million individuals. There are approximately 92,363 patients with cirrhosis of whom approximately 58,197 (63%) have a PDE other than 'EGD with no high-risk varices' on or after the date of cirrhosis diagnosis in the HERALD dataset from 2001 to 2022. A feasibility assessment was conducted in conjunction with this study protocol to ascertain preliminary counts for the subcohorts identified in this study (see section 5.1.2.3), to ensure that a sufficient sample size is reached and the conducted statistical analyses are adequately powered. Preliminary estimates for the subcohort sizes are given below in [redacted] depicts the sequence of PDEs based on first observed occurrence and transition for those patients who experience multiple different events over the follow-up.

[redacted]

[redacted]

5.7 Data management and quality control

The study will use both commonly and less commonly employed data sources for register-based research. Frequently used Swedish administrative population-based registers (e.g., health care visits, prescriptions, and mortality registers) are known to have high completeness and validity due to mandatory reporting; thus, missing data from these sources are expected to be minimal (15, 16, 20). Additional data will be sourced from EHR systems (TakeCare/Cosmic) and FibroScan measurements from the Hepatology Clinic at Karolinska University Hospital, where certain variables may be less consistently recorded (e.g., VCTE measures). All data sources will undergo comprehensive quality control to verify completeness and to identify corrupt or ambiguous entries prior to data management.

As part of Quantify's ongoing HERALD platform initiative, raw data used in this study have been cleaned from their original format. Data management has been conducted using Quantify's Nordic Common Data Model (CDM), Qniform, a proprietary software solution that synthesizes raw data into standardized analysis-ready datasets. Pseudonymized analytical datasets will be compiled to include all necessary observations and variables required for the planned analyses, as specified in Section 5.8. These analytic datasets will be at the person level. All data management has been completed and will continue to be performed by Quantify.

Data cleaning, management, and statistical analyses will be conducted using R. All steps in data processing and analysis will be documented through scripts and logs to ensure reproducibility and facilitate future data updates. A continuously maintained change log will document any deviations from the original combined protocol/statistical analysis plan (SAP), including the rationale for each modification. After cleaning and management, analyses will be carried out according to this document, and results will be populated into pre-specified table shells.

Quality control (QC) procedures will follow Quantify's internal QC protocol. This includes independent code review, where a second programmer evaluates all data cleaning and analysis steps. Errors, inconsistencies, or ambiguities identified during QC will be communicated back to the original analyst, addressed, and documented. The national health registers included in the study are continually validated by national authorities, ensuring high-quality, comprehensive data. Numerous scientific publications based on these registers further support their reliability.

5.7.1 Missingness

Although missing data from the national administrative registers are expected to be minimal, completeness will be assessed across all data sources, and the extent of missingness will be reported where relevant. Missing data may be more common in EHR-derived variables (e.g., VCTE results, diagnosis codes, treatment data, or dates). If missing, certain data points — diagnosis codes, or treatment information—will be excluded from analyses, as imputation would be difficult to justify given the complete coverage of the national register data. Therefore, missing diagnoses or treatment signal absence not missingness.

If a date of event (e.g., diagnosis date) is completely missing, the event will be excluded without imputation. If a date of death is completely missing, the patient will be excluded. For partially missing event dates, the following rules apply:

- If year or month is missing: the entire date is set to missing.
- If only the day is missing: the date will be set to the 15th of the respective month.

In the extremely unlikely situation that a health care visit, prescription, or dispensation is entirely absent from the national registers, such missingness would not be detectable.

5.8 Statistical analysis plan

5.8.1 Statistical software

All analyses will be performed in Stata or R.

5.8.2 Statistical methods

5.8.2.1 Descriptive statics

For all included patients, descriptive statistics will be generated separately in each subcohort (by PDE, see section 5.1.2.3) including for all stratified analyses and sensitivity analyses. Continuous covariates will be described by mean, standard deviation (SD), and interquartile range (IQR). Categorical covariates and continuous covariates that were also categorized will be described by proportion and frequency in each category. The number of missing observations will be also described where applicable.

5.8.2.2 Time-to-event analyses

Time-to-event analyses will be based on a landmark analysis, corresponding to their first PDE. In each cohort adults with cirrhosis with no evidence of prior PDE will be followed from the first PDE (index date). [redacted]

An event-landmark design will be implemented for the primary analysis in which time zero is set to the date of first PDE (index date). Occurrence of outcomes (liver transplant, all-cause mortality, liver-related mortality, CV-related mortality) will be assessed both together and separately after the first protocol-defined event. Patients will be censored at emigration, end of data, or death in cases where the outcome of interest is liver transplantation separately or cause of death specific.

- Unadjusted survival from index (landmark) will be estimated using Kaplan-Meir curves stratified by cohort assigned based on first PDE type with 95% confidence intervals and prespecified time-point survival estimates (1-, 3-, and 5-years). Median time to survival and 95% CI will also be estimated.
- Relative hazards of outcomes (hazard ratio and 95% CI) will be estimated using Cox proportional hazard regression comparing cohort 1 to the reference cohort and cohort 2 to the reference cohort separately (see section 5.1.2.3 for cohorts).
- [redacted]
- [redacted]

Proportional hazards will be assessed using Schoenfeld residuals and visual log(-log) plots. If strong violations are observed, time-varying hazard ratios using flexible interaction with log(time) or present restricted mean survival time differences over clinically relevant horizons may be conducted as a complementary estimand.

To account for competing risks of competing events, sensitivity analyses will be conducted applying Fine-Gray subdistribution hazards or cause-specific hazards, reporting cumulative incidence of death. In models examining all-cause mortality, liver transplant will be treated as a censoring event. In models examining liver transplant, all-cause mortality will be treated as a censoring event. In models examining liver- or CV-related mortality, all-cause mortality and liver transplant will be treated as a censoring event.

The definitions of the PDEs are given in sections 5.1.2.1– 5.1.2.3. Clinical endpoints are listed in Table 5. Baseline characteristics are defined in section 5.2.3. Events and outcomes of interest are defined in section 5.2.1.

5.8.3 Data analysis

An overview of the planned analyses by objectives are presented below. Any major modifications of study outcomes or their analyses would be reflected in a protocol amendment or change log.

5.8.3.1 Objective 1: Patient characteristics

Table 10. Describe patient characteristics (per subcohort)

Objective 1 – Outcome 1	**Describe patient characteristics**
Design	Cohort
Analysis type	Descriptive analysis

Outcomes	• Demographics:
	  ○ Age
	  ○ Sex
	• Cirrhosis etiology
	  ○ Autoimmune
	  ○ Hemochromatosis
	  ○ ALD
	  ○ MASLD/MASH
	  ○ Viral hepatitis
	  ○ PBC
	  ○ PSC
	  ○ Other
	• Lifestyle
	  ○ BMI
	  ○ Overweight BMI in [25, 30)
	  ○ Obesity BMI in [30, Inf)
	• Baseline comorbidities (CCI)
	• Biomarkers
	  ○ Alanine aminotransferase (ALT)
	  ○ Aspartate aminotransferase (AST)
	  ○ AST/ALT ratio
	  ○ Fibrosis-4 Index (FIB-4)
	  ○ Fibrosis stage (Fo-F_4)

Measures	Number, proportion of patients, or mean, SD, median, IQR, min, max.
Study populations	National population, Stockholm population
Subcohorts	Patient characteristics will be presented separately for each subcohort.

Objective 1 – Outcome 1	Describe patient characteristics
Reporting interval	During lookback period and at index date (depending on outcomes)

5.8.3.2 Objective 2: Risk of liver transplantation and mortality

Table 11. Estimate risk of liver transplantation/mortality

Objective 2.1	Estimate liver transplantation/mortality
Design	Cohort
Analysis type	Time-to-event analysis
Outcomes	Liver transplantation/mortality
Measure	• Counts (proportions) and cumulative incidence and 95% CI stratified by cohort and subcohort at 1-, 3-, 5-years post-index and overall. Median time to survival and 95% CI by cohort and subcohort.
	• Cumulative incidence plots by cohort and subcohort
	• Unadjusted and adjusted hazard ratios (95% CI) comparing 1) cohort 1 and reference cohort and 2) cohort 2 and reference cohort
Study populations	National population, Stockholm population
Reporting interval	Follow-up period

Table 12. Estimate risk of mortality (all-cause, CV-related, liver-related)

Objective 2.2	Estimate mortality
Design	Cohort
Analysis type	Time-to-event analysis
Outcomes	All-cause mortality Cardiovascular-related mortality Liver-related mortality
Measures	• Counts (proportions) and cumulative incidence and 95% CI stratified by cohort and subcohort at 1-, 3-, 5-years post-index and overall. Median time to survival and 95% CI by cohort and subcohort.
	• Cumulative incidence plots by cohort and subcohort
	• Unadjusted and adjusted hazard ratios (95% CI) comparing 1) cohort 1 and reference cohort and 2) cohort 2 and reference cohort
Sensitivity analyses	• Sensitivity analysis: Competing risks estimates (Fine-Gray subdistribution hazards (95% CIs) or cause specific hazards accounting for competing risks (95% CIs)
Study populations	National population, Stockholm population
Reporting interval	Follow-up period
Specification	Time-to-event analyses will be run separately for each outcome

Table 13. Estimate risk of liver transplantation

Objective 2.3	Estimate mortality
Design	Cohort
Analysis type	Time-to-event analysis
Outcomes	Liver transplant
Measures	• Counts (proportions) and cumulative incidence and 95% CI stratified by cohort and subcohort at 1-, 3-, 5-years post-index and overall. Median time to survival and 95% CI by cohort and subcohort.
	• Cumulative incidence plots by cohort and subcohort
	• Unadjusted and adjusted hazard ratios (95% CI) comparing 1) cohort 1 and reference cohort and 2) cohort 2 and reference cohort
Study populations	National population, Stockholm population
Reporting interval	Follow-up period

5.8.3.3 *Exploratory objectives: quality of life outcomes*

Table 14. Estimate quality of life burden via proxy

Exploratory objective	Estimate quality of life burden via proxy
Design	Cohort
Analysis type	Descriptive, Time-to-event analysis
Outcomes	Work absence Depression Anxiety disorders Stress disorders Fatigue Sleep disturbance Work absence
Measures	• Cumulative incidence curves and 95% CI (figure) stratified by cohort and subcohort (excluding work absence)
	• Kaplan-Meier survival by cohort and subcohort
	• Summary of days of annualized days of work absence by cohort and subcohort at 1-, 3-, and 5-years after index
Study populations	National population, Stockholm population
Reporting interval	Follow-up period

5.8.4 Additional details on analyses

For all studied cohorts and subcohorts, patients will be followed from respective index date to occurrence of outcome or censoring (death, if relevant, or emigration).

5.8.5 Sensitivity analyses

A selection of the results may be evaluated in sensitivity analyses to characterize the impact of key study population definitions and limitations – namely, if a significant proportion (at least 5%) of the given patient subcohort is found to experience multiple contemporaneous (within 30 days of each other) PDEs during follow-up, the 1, 3, and 5 -year cumulative incidence for liver transplantation and all-cause, cardiovascular-, and liver-related mortality may be re-evaluated in a competing risks framework where for each PDE, the alternative events serve as the competing risks.

[redacted]

If sample sizes allow (i.e., minimum 40 patients in each subcohort), all analyses will be performed among the subset of patients with cirrhosis due to MASH. These patients will be identified as those with a MASH/MASLD diagnosis (K75.81/K76.0) prior to or at cirrhosis diagnosis. This will provide further information regarding risk among patients with cirrhosis due to MASH.

[redacted]

[redacted]

5.9 Limitations of the research methods

5.9.1 Internal and external validity of the study design

[redacted]

The HERALD research platform integrates data from Sweden's national and regional registers, which together provide an exceptionally robust foundation for both internal and external validity. The Swedish national registers offer virtually complete population coverage, encompassing all residents with no attrition other than emigration. This comprehensive inclusion minimizes selection bias and ensures a population-level sampling frame that enhances generalizability (external validity).

Individual-level linkage across registers is enabled by the unique personal identification number assigned to all Swedish residents. This linkage supports accurate temporal sequencing of exposures and outcomes, reduces misclassification, and strengthens the internal validity of longitudinal analyses. The data include all inpatient and specialized outpatient care, as well as all outpatient pharmacy dispensations, ensuring near-complete capture of clinically relevant events.

Extensive validation studies confirm the high quality of these data, with PPVs between 72–93% for diagnosis codes and approximately 97% for procedural codes (15). Liver-related conditions are particularly well captured: PPVs exceed 90% for cirrhosis and range from 82–95% for non-alcoholic fatty liver disease (16, 20). These high validity metrics further reinforce the reliability of the data for both exposure and outcome ascertainment, supporting strong internal validity while enabling findings to be confidently generalized to the broader Swedish population, as well as to similar health care settings, e.g., with similar access to care and treatment guidelines.

5.9.1.1 External validity and regulatory relevance of Swedish register data for liver disease

Swedish national health registers constitute regulatory-grade data sources, given their nationwide coverage, linkage via personal identifiers, and documented validity for chronic liver disease and cirrhosis diagnoses (20, 22). Indeed, multiple Swedish registers are listed in the EMA's RWD catalogue, where over 100 studies have been conducted using this data for this purpose (23-25) and the impact of the value of individual linkage between the registers has been identified (26).

For liver disease specifically, Swedish register-based epidemiology has been repeatedly incorporated into European-level assessments of cirrhosis burden, etiology, and outcomes, including EMA-relevant scientific and policy documents that inform regulatory deliberations (27, 28). These Europe-

wide syntheses include Sweden alongside other EU Member States and show that incidence, mortality, and etiological patterns of cirrhosis in Sweden fall within the broader European distribution, suggesting the generalizability of Swedish registry data to a broader European population (27, 28).

5.9.2 Limitations due to missing/incomplete data

As this is an observational study, potential confounding factors cannot be ruled out. Data collection will reflect routine clinical practice rather than mandatory assessments at prespecified time points, which may have an impact on the amount of data and its interpretation.
This study will to a large extent rely on register-based data, which are known to have a high degree of completeness. The reporting of many variables in the national health care registers used in this study is mandatory and there is therefore almost complete coverage of the information related to health care visits and prescriptions. For some specific variables incomplete data is however expected, including, all lab measures and follow-up data in the Quality of Care registers (e.g. BMI). However, missing data is expected to be similar across all subcohorts, thus any bias is expected to be nondifferential.

6. Protection of human subjects

6.1 Patient information

This study involves data that exist in pseudonymized structured format and contain no direct identifiable patient information.

6.2 Patient consent

The study will be carried out in compliance with the protocol, the principles laid down in the Declaration of Helsinki, Guidelines for Good Pharmacoepidemiology Practice (GPP), and the relevant Standard Operating Procedures (SOPs). Standard medical care (prophylactic, diagnostic, and therapeutic procedures) remains the responsibility of the physician treating the patient.
Per design, this non-interventional study utilizes secondary data. Thus, the study does not affect the treatment or health outcomes of the study individuals. The study individuals will not be contacted during any phase of the study. For our research purposes, per Swedish law on the use of secondary health data, individual consent is not necessary to collect due to the retrospective non-interventional nature of the study.

6.3 Institutional review board (IRB)/Independent ethics committee (IEC)

The study will be conducted in accordance with applicable legal and regulatory requirements. Ethical approval from the Independent Ethics Committee in Sweden was previously obtained for completion of RWE studies conducted using the HERALD data (2021-12-06, Dnr 2021-04422; 2022-06-15, Dnr 2022-02875-02).

6.4 Ethical conduct of the study

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

Guidelines for GPP issued by the International Society for Pharmacoepidemiology (ISPE) and Good Practices for Outcomes Research issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

7. Plans for disseminating and communicating study results

7.1 Publications

Authorship of any publications resulting from this study will be determined based on the International Committee of Medical Journal Editors authorship requirements and Sponsor contract.

8. References

1. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol. 2006;44(1):217-31.

2. Garcia-Tsao G, Abraldes JG, Berzigotti A, Bosch J. Portal hypertensive bleeding in cirrhosis: Risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases. Hepatology. 2017;65(1):310-35.

3. Kumar R, Kerbert AJC, Sheikh MF, Roth N, Calvao JAF, Mesquita MD, et al. Determinants of mortality in patients with cirrhosis and uncontrolled variceal bleeding. J Hepatol. 2021;74(1):66-79.

4. Nigatu A, Yap JE, Lee Chuy K, Go B, Doukky R. Bleeding Risk of Transesophageal Echocardiography in Patients With Esophageal Varices. J Am Soc Echocardiogr. 2019;32(5):674-6.e2.

5. Toresson Grip E, Ström O, Hagström H. Health outcomes and risk assessment in chronic liver disease (HERALD): a large Swedish research platform [Abstract]. The European Association for the Study of the Liver (EASL); London: Journal of Hepatology; 2022.

6. Shang Y, Toresson Grip E, Ntanios F, Modica A, Ström O, Hagström H. ADDITIONAL COMPONENTS OF THE METABOLIC SYNDROMES INCREASE THE RISK FOR SEVERE LIVER DISEASE IN PATIENTS WITH TYPE 2 DIABETES IN A DOSE-DEPENDENT MANNER [Abstract]. American Association for the Study of Liver Diseases (AASLD); Washington: Hepatology; 2022. p. S620.

7. Toresson Grip E, Skröder H, Shang Y, Ström O, Hagström H. Prevalence of NAFLD and advanced fibrosis in Stockholm, Sweden The European Association for the Study of the Liver (EASL) [Abstract]; Vienna: [Accepted]; 2023.

8. Bergquist A, Kachru N, Aldvén M, Ström O, Knutsson D, Toresson Grip E, et al. Healthcare resource use and costs among patients with primary sclerosing cholangitis in Sweden - a retrospective population-based cohort study [Abstract]. The European Association for the Study of the Liver (EASL); Vienna: [Accepted]; 2023.

9. Toresson Grip E, Scheffer Apecechea N, Basuroy R, Johansen P, Kant Mangla K, Skröder H, et al. Characteristics and risk of major adverse liver outcomes & major cardiovascular events among Swedish patients with diagnosed and high risk of metabolic dysfunction-associated steatohepatitis – a REVEAL-MASH study. EASL [Abstract]. 2024.

10. Toresson Grip E, Scheffer Apecechea N, Basuroy R, Johansen P, Kant Mangla K, Skröder H, et al. Health care resource utilization for patients diagnosed and at high risk of metabolic dysfunction-associated steatohepatitis in Stockholm, Sweden – a REVEAL-MASH study. EASL [Abstract]. 2024.

11. Shang Y, Grip ET, Modica A, Skroder H, Strom O, Ntanios F, et al. Metabolic Syndrome Traits Increase the Risk of Major Adverse Liver Outcomes in Type 2 Diabetes. Diabetes Care. 2024;47(6):978-85.

12. Toresson Grip E, Shang Y, Skröder H, Modica A, Ström O, Ntanios F, et al. Number of metabolic syndrome traits adds little to prediction of major adverse liver outcomes on top of FIB-4 in patients with type 2 diabetes. Diabetes Epidemiology and Management. 2025;19-20:100276.

13. de Franchis R, Bosch J, Garcia-Tsao G, Reiberger T, Ripoll C. Baveno VII - Renewing consensus in portal hypertension. J Hepatol. 2022;76(4):959-74.

14. European Association for the Study of the Liver. EASL Clinical Practice Guidelines for the management of patients with decompensated cirrhosis. J Hepatol. 2018;69(2):406-60.

15. Everhov Å H, Frisell T, Osooli M, Brooke HL, Carlsen HK, Modig K, et al. Diagnostic accuracy in the Swedish national patient register: a review including diagnoses in the outpatient register. Eur J Epidemiol. 2025;40(3):359-69.

16. Åström H, Wester A, Hagström H. Administrative coding for non-alcoholic fatty liver disease is accurate in Swedish patients. Gastrodagarna; Örebro, Sweden2023.

17. Ludvigsson JF, Appelros P, Askling J, Byberg L, Carrero JJ, Ekström AM, et al. Adaptation of the Charlson Comorbidity Index for Register-Based Research in Sweden. Clin Epidemiol. 2021;13:21-41.

18. Socialstyrelsen (National Board of Health and Welfare). Statistical register's production and quality: National Patient Register. 2022.

19. Eeg-Olofsson K, Svensson AM, Franzén S, Ahmed Ismail H, Törnblom M, Levrat-Guillen F. Real-world study of flash glucose monitoring among adults with type 2 diabetes within the Swedish National Diabetes Register. Diab Vasc Dis Res. 2023;20(1):14791641211067418.

20. Bengtsson B, Askling J, Ludvigsson JF, Hagström H. Validity of administrative codes associated with cirrhosis in Sweden. Scand J Gastroenterol. 2020;55(10):1205-10.

21. de Franchis R. Expanding consensus in portal hypertension: Report of the Baveno VI Consensus Workshop: Stratifying risk and individualizing care for portal hypertension. J Hepatol. 2015;63(3):743-52.

22. Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaëlsson K, Neovius M, et al. Registers of the Swedish total population and their use in medical research. Eur J Epidemiol. 2016;31(2):125-36.

23. European Medicines Agency (EMA). RWD Catalogues: Sweden National Prescribed Drugs Register / Läkemedelsregistret 2024 [Available from: https://catalogues.ema.europa.eu/node/968/administrative-details.

24. European Medicines Agency (EMA). RWD Catalogues: Sweden National Cancer Register / Cancerregistret 2024 [Available from: https://catalogues.ema.europa.eu/data-source/1111194.

25. European Medicines Agency (EMA). RWD Catalogues: Swedish Cause of Death Register 2024 [Available from: https://catalogues.ema.europa.eu/data-source/1000000245.

26. European Medicines Agency (EMA). Health Impact - Swedish Population Evidence Enabling Data-linkage 2025 [Available from: https://catalogues.ema.europa.eu/data-source/1000000562.

27. Blachier M, Leleu H, Peck-Radosavljevic M, Valla DC, Roudot-Thoraval F. The burden of liver disease in Europe: a review of available epidemiological data. J Hepatol. 2013;58(3):593-608.

28. Pimpin L, Cortez-Pinto H, Negro F, Corbould E, Lazarus JV, Webber L, et al. Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. J Hepatol. 2018;69(3):718-35.

Signature Page for Redacted Protocol - SG46805 - PEGOZAFERMIN - v2 -
System identifier: RIM-CLIN-1231152

Approval Task	[redacted] Company Signatory 24-Jul-2026 15:30:20 GMT+0000