Study Protocol P2-C1-009

22/03/2024

Version 2.3

TABLE OF CONTENTS

DOCUMENT HISTORY

Version	Date	Description
V1.0	26/01/2024	Initial version for EMA review
V2.0	15/02/2024	Second version including comments from EMA
V2.1	14/03/2024	Refinement of code list in collaboration with EMA
V2.2	22/03/2024	Definitive code list following EMA comments

Study Title	DARWIN EU® – Frailty and polypharmacy among adults with selected cancers at the time of diagnosis
Protocol version identifier	2.3
Date of last version of protocol	22/03/2024
EU PAS register number	EUPAS1000000120
Active substance	N/A
Medicinal product	N/A
Research question and objectives	The aim of this study is to estimate the prevalence of frailty and polypharmacy at the point of diagnosis of selected cancers in people aged 18 and above and to describe their characteristics. The specific objectives of the study are: 1. To estimate the prevalence of frailty and polypharmacy in people aged 18 and above diagnosed with selected cancers at the point of cancer diagnosis. 2. To describe the characteristics of people aged 18 and above diagnosed with selected cancers among different frailty and polypharmacy categories at the point of cancer diagnosis.
Country(-ies) of study	Belgium, Estonia, Germany, Spain, The Netherlands, The United Kingdom
Author	Talita Duarte-Salles (tduarte@darwin-eu.org)

LIST OF ABBREVIATIONS

Abbreviation	Name
ALL	Acute Lymphocitic Leukemia
AML	Acute Myeloid Leukemia
CDM	Common Data Model
CLL	Chronic Lymphocitic Leukemia
CML	Chronic Myeloid Leukemia
CPRD	Clinical Practice Research Datalink
DA	Disease Analyzer
DARWIN EU®	Data Analysis and Real World Interrogation Network
DOI	Declaration Of Interests
DQD	Data Quality Dashboard
GP	General Practitioner
EHR	Electronic Health Record
EMA	European Medicines Agency
EBB	Estonian Biobank
ECOG	Eastern Cooperative Oncology Group
EGCUT	Estonian Genome Center at the University of Tartu
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
IPCI	Integrated Primary Care Information Project
LPD	Longitudinal Patient Database
NSCLC	Non-Small Cell Lung Cancer
PCT	Primary Care Teams
OMOP	Observational Medical Outcomes Partnership
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària
SNOMED	Systematized Nomenclature of Medicine
WHO	World Health Organisation

1. TITLE

DARWIN EU® – Frailty and polypharmacy among adults with selected cancers at the time of diagnosis

2. RESPONSIBLE PARTIES – STUDY TEAM

Study team Role	Names	Organisation
Study Project Manager/Principal Investigator	Talita Duarte-Salles	Erasmus MC
Epidemiologist / Clinical Expert	Julieta Politi	Erasmus MC
Data Scientists	Ross Williams Adam Black Maarten van Kessel Cesar Barbosa	Erasmus MC Erasmus MC Erasmus MC Erasmus MC

Data Partner*	Names	Organisation – Database
Local Study Coordinator/Data Analyst	Cesar Mendosa Mees Mosseveld Laura Pérez Crespo Anna Palomar Lucía Amalia Carrasco Ribelles James Blash Antonella Delmestri Raivo Kolde Marek Oja	Erasmus MC – IPCI Erasmus MC – IPCI IDIAPJGol – SIDIAP IDIAPJGol – SIDIAP IDIAPJGol – SIDIAP IQVIA – DA Germany and LPD Belgium University of Oxford – CPRD-GOLD University Tartu – Estonian Biobank University Tartu – Estonian Biobank

3. ABSTRACT

Title

DARWIN EU® – Frailty and polypharmacy among adults with selected cancers at the time of diagnosis

Rationale and Background

Frailty, polypharmacy, and comorbidities are common and important factors which usually coexist in older patients. Assessment of frailty and polypharmacy is difficult, due to lack of standardized definitions. However, accounting for them is relevant, especially among older adults with cancer, due to their adverse impact on cancer outcomes and treatment. Despite this, studies reporting on the prevalence of frailty and polypharmacy specifically in older adults with cancer remain sparse.

This study intends to investigate the ability to characterise frailty and polypharmacy in real-world data sources, to estimate the prevalence of frailty and polypharmacy in people aged 18 and above with selected cancers at the point of diagnosis and to describe their characteristics. While the focus is on older adults, the study will explore the full age range of adulthood to better contextualise the results.

Research question and Objectives

The aim of this study is to estimate the prevalence of frailty and polypharmacy at the point of diagnosis of selected cancers in people aged 18 and above and to describe their characteristics.

The specific objectives of the study are:
1. To estimate the prevalence of frailty and polypharmacy in people aged 18 and above diagnosed with selected cancers at the point of cancer diagnosis.
2. To describe the characteristics of people aged 18 and above diagnosed with selected cancers among different frailty and polypharmacy categories at the point of cancer diagnosis.

All results will be reported by database and selected cancer type, overall and stratified by age and sex.

Research Methods

Study design

Population-based cohort study.

Population

The study population will include all individuals aged 18 years and above with a primary diagnosis of selected cancers (lung, breast, ovary, endometrium, prostate, pancreas, colorectal cancer, lymphoma, leukemia and myeloma) recorded between 01/01/2017 and 31/12/2022, with at least one year of prior history available before cancer diagnosis. Individuals with a diagnosis of cancer (any, excluding non-melanoma skin cancer) any time prior to the diagnosis of one of the selected cancers will be excluded.

Additional eligibility of a minimum of 1 year of potential follow-up time prior to the end of last database observations will be imposed for the estimation of one-year hospitalisation and mortality rates if the data sources capture this information.

Data sources

1. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom
2. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
3. IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium), Belgium
4. Integrated Primary Care Information Project (IPCI), The Netherlands
5. Estonian Biobank (EBB), Estonia
6. Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain

Variables of interest

The number of concomitant prescriptions will be calculated based on the overall number of prescriptions during the 90 days prior to index date, and polypharmacy will be measured as the maximum number of drug eras that overlap on any day during the 90-day period, using two definitions as the concomitant

prescription of >= 5 and >= 10 medications (ingredient level) anytime during the 90 days prior to cancer diagnosis.

To measure frailty, a score will be created based on the presence of polypharmacy (defined as >=5 prescriptions of medications as mentioned above) and the following conditions included in frailty indexes previously proposed in the literature: Mobility and transfer problems; Housebound; Activity limitation; Visual impairment; Hearing impairment; Requirement for care; Social vulnerability; Falls; Urinary incontinence; Weight loss and anorexia; Memory and cognitive problems; Dizziness; Dyspnoea; Sleep disturbance; Anaemia and haematinic deficiency; Hypertension; Ischaemic heart disease; Heart failure; Cerebrovascular disease; Peripheral vascular disease; Atrial fibrillation; Heart valve disease; Hypotension/syncope; Diabetes; Foot problems; Arthritis; Respiratory disease; Peptic ulcer; Thyroid disease; Chronic kidney disease; Osteoporosis; Fragility fracture; Parkinsonism and tremor; Urinary system disease; Skin ulcer. Their definition will be based on diagnosis codes recorded anytime prior or at cancer diagnosis, except for polypharmacy for which we will need the definition mentioned above (including medicine use 90 days prior to index date). The frailty score will be calculated based on the number of conditions present and polypharmacy (defined as >=5 concominant prescriptions 90 days prior to index date) divided by the total number of conditions/polypharmacy mentioned above (35 conditions and 1 for polypharmacy prevalence). Individuals will then be further categorised into the following levels of severity according to their scores: fit: 0–0.12; mild frailty: >0.12–0.24; moderate frailty= >0.24–0.36; severely frail: >=0.36.

All co-morbidities and medications will be used for large-scale patient characterisation, identified as concept/code and descendants.

Other variables of interest will include number of hospitalisation (SIDIAP and EBB) and mortality (CPRD, IPCI, EBB, SIDIAP) during the year after date of cancer diagnosis.

Sample size

No sample size will be calculated as this is a descriptive Disease Epidemiology Study where we are interested in the characteristics of all incident cases of selected cancers.

Data analyses

The prevalence of the frailty, overall and its categories (as defined based on the above-mentioned score), and polypharmacy (objective 1) will be estimated at the time of cancer diagnosis.
Large-scale patient-level characterisation (objective 2) will be conducted for individuals with different frailty and polypharmacy categories as follows: Age and sex at time of cancer diagnosis will be described; Medical history will assessed for: any time prior to index date; up to 365 days before index date; 365 to 31 days before index date; 30 to 1 day before index date; and at index date; Medication use history will be reported for 365 to 31 days before index date, for 30 to 1 day before index date, and at index date; hospitalization and mortality rates will be calculated for up to 365 days after index date.

For all analyses, absolute and relative frequencies will be reported. A minimum cell count of 5 will be used when reporting results, with any smaller counts reported as “<5” and zero counts as “0”. Analyses will be done separately for each database and selected cancers. Next to overall reporting, stratification by age category (18-44, 45-64, 65-74, 75-84 and 85+) and sex will be conducted when possible (minimum cell count reached). Results for objective one on the estimation of the prevalence of frailty in cancer patients, results will be further stratified by time period (before or after 2020).

4. AMENDMENTS AND UPDATES

Number	Date	Section of studyprotocol	Amendment orupdate	Reason

5. MILESTONES

STUDY SPECIFIC DELIVERABLE	TIMELINE
Draft Study Protocol	26/01/2024
Final Study Protocol	February 2024
Creation of Analytical code	March 2024
Execution of Analytical Code on the data	March 2024
Draft Study Report	April 2024
Final Study Report	April/May 2024

6. RATIONALE AND BACKGROUND

Frailty, polypharmacy, and comorbidities are important considerations for the health of people aged ≥ 65 years. Frailty is an age-associated clinical syndrome characterised by decreased physiological reserves, increased vulnerability to stressors, and diminished capacity to maintain homeostasis due to a cumulative decline in the individuals' physiological systems (1). The estimated pooled prevalence of frailty among older adults (≥ 65) within European community settings has been estimated at 12% (2). Despite being commonly observed in older adults and being associated with increased risks of adverse outcomes, including treatment toxicity, hospitalisation, and mortality, achieving a concise definition remains challenging (3). In addition to frailty, polypharmacy poses a further challenge and is an increasing concern when treating older adults due to the potential contraindications and drug-drug interactions among treatments (4, 5). While there is no universal definition of polypharmacy, one of the most widespread definitions refers to the presence of 5 or more medications, with extreme polypharmacy defined as the presence of 10 or more medications, and its prevalence ranges from 7-45% in community settings (6, 7).

Unlike younger patients with cancer, older adults are at increased risk from cancer treatment, especially among vulnerable older subjects, who are more susceptible to unfavourable health events and complications during the disease course. In turn, there is interest in measuring frailty among older adults with cancer since frailty can be understood as an ageing-associated vulnerability and has been recognised as an important factor to successful cancer treatment in patients of advanced age (8-13).

Integrating frailty assessments into cancer clinical trials allows for a more comprehensive evaluation of patient's health status and helps tailor treatment approaches to individual patients. Identifying frailty enables clinicians to recognise patients who may have a poorer prognosis for adverse outcomes occurrence and be at higher risk of treatment-related complications. Some indexes have been proposed to predict the risk of chemotherapy toxicity and survival among this specific patient population (14). However, most indexes have focused primarily on patients with solid tumours.

Moreover, frailty assessment can allow decision-makers to understand better the patient population participating in trials and potentially account for frailty-related factors when considering the evidence on treatment efficacy and safety. In this sense, clinical trial inclusion and exclusion criteria may lead to uncertainty related to the external validity of trial results to the general patient population (which includes individuals of varying ages and performance statuses). Likewise, despite some comorbidities and polypharmacy being prevalent, many oncology trials do not routinely publish data on the frequency of common comorbid conditions or the extent of polypharmacy in their trial populations (14-18).

Frailty is a multidimensional concept, including physical, psychological, and social constructs, best captured in the context of a geriatric assessment (19). Two methods to measure frailty are generally the most accepted approaches for identifying frailty in older adults (10, 11). The Fried phenotypic frailty approach defines frailty as a clinical syndrome in which three or more of the following criteria are present: unintentional weight loss (10 lbs. in the past year), self-reported exhaustion, weakness (grip strength), slow walking speed, and low physical activity (20). The index proposed by Rockwood et al. defines frailty as an accumulation of deficits, where an individual's health status can be quantified as a proportion of ageing-associated deficits using measures from comprehensive geriatric assessment (21). While the ECOG (Eastern Cooperative Oncology Group) performance status scale is often used in oncology trials to assess cancer patients' functional status and overall health, data on ECOG status may not be routinely collected in electronic health records or administrative databases. Furthermore, some have argued that ECOG status does not reliably capture the full extent of frailty and may be subjective, with marked interobserver variability (22).
Several other frailty measures have been developed and applied to healthcare databases to support clinical care and research. Examples include the electronic frailty index, the modified frailty index, the hospital frailty risk score and the Faurot frailty index, among others (3, 23-25).

Older adults undergoing cancer-related treatment are more likely to experience polypharmacy because they tend to have a greater number of underlying comorbidities requiring treatment, together with the need to prescribe additional cancer-related medications, including cancer therapy and supportive medications (26-29). Despite several adverse outcomes being linked to polypharmacy, studies of polypharmacy in older adults with cancer are limited (30, 31).

This study aims to investigate the ability to measure frailty and polypharmacy in included DARWIN EU data sources, estimate the prevalence of frailty and polypharmacy in adults with selected cancers at the point of diagnosis, and describe their characteristics.

7. RESEARCH QUESTION AND OBJECTIVES

The aim of this study is to estimate the prevalence of frailty and polypharmacy at the point of diagnosis of selected cancers in adults and to describe their characteristics.

The specific objectives of the study are:

1. To estimate the prevalence of frailty and polypharmacy in people aged 18 and above diagnosed with selected cancers at the point of cancer diagnosis.
2. To describe the characteristics of people aged 18 and above diagnosed with selected cancers among different frailty and polypharmacy categories at the point of cancer diagnosis.

While understanding frailty and polypharmacy in older adults (aged 65 years and over) remains a key focus we will also characterise younger adults with the selected cancers to contextualise the results. All results will be reported by database and selected cancer type, overall and stratified by age and sex. Results for objective 1 will be further stratified by study period (before or after 2020) to capture the potential impact of the pandemic in the prevalence of frailty and polypharmacy.

Table 1: Primary and secondary research questions and objective

Objective:	To estimate the prevalence of frailty and polypharmacy at the point of diagnosis of selected cancers in people aged 18 and above and to describe their characteristics.
Hypothesis:	N/A
Population *(mention key inclusion-exclusion criteria):*	The study population will include all individuals aged 18 years and above with a primary diagnosis of any of the selected cancers (lung, breast, ovary, endometrium, prostate, pancreas, colorectal cancer, lymphoma, leukemia and myeloma) recorded between 01/01/2017 and 31/12/2022, who have at least one year of prior history available before cancer diagnosis (index date). Individuals with a history of cancer (any, excluding non-melanoma skin cancer) prior to the diagnosis of one of the selected cancers will be excluded. Additional eligibility of a minimum of 1 year of potential follow-up time prior to the end of last database observations will be imposed for the estimation of one-year hospitalisation and mortality rates.
Exposure:	The number of concomitant prescriptions will be calculated based on the overall number of prescriptions during the 90 days prior to index date, and polypharmacy will be measured as the maximum number of drug eras that overlap on any day during the 90-day period. Polypharmacy will be defined as the concomitant prescription of >= 5 and >= 10 medications (ingredient level) anytime during the year prior to cancer diagnosis.

	A frailty score will be implemented based on the presence of polypharmacy (defined as >=5 prescriptions of medications as mentioned above) and the following conditions included in frailty indexes (i.e.: eFI or eFRAGICAP): Mobility and transfer problems; Housebound; Activity limitation; Visual impairment; Hearing impairment; Requirement for care; Social vulnerability; Falls; Urinary incontinence; Weight loss and anorexia; Memory and cognitive problems; Dizziness; Dyspnoea; Sleep disturbance; Anaemia and haematinic deficiency; Hypertension; Ischaemic heart disease; Heart failure; Cerebrovascular disease; Peripheral vascular disease; Atrial fibrillation; Heart valve disease; Hypotension/syncope; Diabetes; Foot problems; Arthritis; Respiratory disease; Peptic ulcer; Thyroid disease; Chronic kidney disease; Osteoporosis; Fragility fracture; Parkinsonism and tremor; Urinary system disease; Skin ulcer. Their definition will be based on diagnosis codes recorded any time prior or at cancer diagnosis. The score will be calculated based on the number of present conditions and polypharmacy divided by the total number of conditions/polypharmacy mentioned above (35 conditions and 1 for polypharmacy prevalence) and will be further categorised accordingly into the following levels of severity: fit: 0–0.12; mild frailty: >0.12–0.24; moderate frailty= >0.24–0.36; severely frail: >=0.36.
Comparator:	N/A
Outcome:	N/A
Time *(when follow up begins and ends):*	For the estimation of one-year hospitalization and mortality rates, follow-up will start from date of cancer diagnosis (index date) until the earliest of the following: 1) loss to follow-up, 2) end of data availability, 3) date of death, or 4) end of the 365 days follow-up.
Setting:	Outpatient setting from 6 databases currently in DARWIN EU covering 6 European countries.
Main measure of effect:	Proportions

8. RESEARCH METHODS

8.1 Study type and Study Design

This will be a **population-level descriptive epidemiology** and **patient-level characterisation** study classified as "off-the-shelf" and as described in the DARWIN EU® Complete Catalogue of Standard Data Analyses. A population-based cohort study of all incident cases of selected cancers will be conducted.

Table 2. Description of Potential Study Types and Related Study Designs

STUDY TYPE	STUDY DESIGN	STUDY CLASSIFICATION
Population-level descriptive epidemiology	Population-level cohort	Off the shelf
Patient-level characterisation	Cohort analysis	Off the shelf

8.2 Study Setting and Data Sources

This study will be conducted using routinely collected health data from 6 databases in the DARWIN EU network of data partners from 6 European countries. All databases were previously mapped to the OMOP CDM.

Data sources

1. Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom
2. IQVIA Disease Analyzer Germany (IQVIA DA Germany), Germany
3. IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium), Belgium
4. Integrated Primary Care Information Project (IPCI), The Netherlands
5. Estonian Biobank (EBB), Estonia
6. Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain

These databases fulfil the criteria required for a population- and patient-level characterisation study allowing for large-scale characterisation, while covering different regions of Europe. The selection of databases was based on the availability of prior history data at index date on conditions and medicines in order to estimate frailty and polypharmacy, as well as to perform patient-level characterisation. Detailed information on the selected data sources are described in **Table 3**.

When it comes to assessing the reliability of data sources, the data partners are asked to describe their internal data quality process on the source data as part of the DARWIN EU onboarding procedure. To further ensure data quality, we utilize the Achilles software tool (https://ohdsi.github.io/TheBookOfOhdsi/DataQuality.html#data-quality-checks), which systematically characterizes the data and presents it in a dashboard format that is inspected. The generated data characteristics such as age distribution, condition prevalence per year, data density, measurement value distribution can be compared against expectations for the data. Additionally, the data quality dashboard (DQD) provides more objective checks on plausibility consistently across the data sources. In terms of relevance, a more general purpose diagnostic software tool, CohortDiagnostics, was developed. This package evaluates phenotype algorithms for OMOP CDM datasets, offering a standard set of analytics for understanding patient capture including data generation. It provides additional insights into cohort characteristics, record counts and index event misclassification. Furthermore, timeliness is guarded by extracting the release dates for each dataset in the network and monitoring when data are out-of-date with the expected refresh cycle (typically quarterly or half-yearly). In addition, it is important to have clear understanding of the time period covered by each released database, as this can vary across different domains. To facilitate this, the CdmOnboarding (and Achilles) packages contain a 'data density' plot that

are checked. This plot displays the number of records per OMOP domain on a monthly basis. This allows to get insights when data collection started, when new sources of data were added and when until when data was included.

Table 3. Description of the selected Data Sources.

Country	Name of Database	Justification for Inclusion	Health Care setting	Type of Data	Number of active subjects	Data lock for the last update
BE	IQVIA LPD Belgium	Covers primary care setting with information on cancer diagnoses and medical history. Hospitalisations and deaths are not routinely recorded.	Primary care	EHR	1.1 million	31/03/2023
DE	IQVIA DA Germany	Covers primary care setting with information on cancer diagnoses and medical history. Hospitalisations and deaths are not routinely recorded.	Primary care and outpatient specialist care	EHR	42 million	31/03/2023
ES	SIDIAP	Covers primary care setting, data on cancer diagnoses previously validated, information available on history of conditions and medications, number of hospitalisations and date of death.	Primary care with hospital data linkage	EHR	5.8 million	30/06/2023
ET	EBB	Contains information on 200,000 participants with not only genetic information but also health insurance claims, digital prescriptions, discharge information and causes of death through linkage with the national death register. Data is linked to cancer registry.	Biobank wi th hospital and cancer registry linkage	Claims	0.2 million	31/12/2022
NL	IPCI	Covers primary care setting, data on cancer diagnoses previously validated, information available on history of conditions and medications, and date of death. Hospitalisations are not routinely recorded.	Primary care	EHR	2.8 million	01/09/2023
UK	CPRD GOLD	Covers primary care setting, data on cancer diagnoses, information available on history of conditions and medications, and date of death. Hospitalisations are not routinely recorded in the primary care data.	Primary care	EHR	17.3 million	01/07/2023

BE = Belgium, CPRD = Clinical Practice Research Datalink, DA = Disease Analyzer; DE = Germany, EBB = Estonian Biobank, ES = Spain, ET = Estonia, IPCI = Integrated Primary Care Information Project, LPD = Longitudinal Patient Database; NL = The Netherlands, SIDIAP = Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària,

Clinical Practice Research Datalink (CPRD) GOLD, United Kingdom

The Clinical Practice Research Datalink (CPRD) GOLD is a database of anonymised electronic health records (EHR) from General Practitioner (GP) clinics in the UK that use the Vision® software system for their management.(32) The source population encompasses 98% of the UK, registered with GPs responsible for non-emergency care and referrals. Participating GPs provide CPRD EHR for all registered patients who did not specifically request to opt out of data sharing. Covering 4.6% of the current UK population, GOLD includes 4.9% of contributing GP practices, providing comprehensive information within its defined source population. GOLD contains data from all four UK constituent countries and the current regional distribution of its GP practices is 5.7% in England, 55.6% in Scotland, 28.4% in Wales, and 10.2% in Northern Ireland (May 2022).

GOLD data include patient's demographic, biological measurements, clinical symptoms and diagnoses, referrals to specialist/hospital and their outcome, laboratory tests/results, and prescribed medications. GOLD has been assessed and found broadly representative of the UK general population in terms of age, gender, and ethnicity.(32) GOLD has been widely used internationally for observational research to produce nearly 3,000 peer-reviewed publications, making GOLD the most influential UK clinical database so far.(33-35)

In terms of quality checks, the integrity, structure, and format of the data is reviewed. Collection-level validation ensures integrity by checking that data received from practices contain only expected data files and ensures that all data elements are of the correct type, length, and format. Duplicate records are identified and removed.(32) Transformation-level validation checks for referential integrity between records ensure that there are no orphan records included in the database (for example, that all event records link to a patient), while research-quality-level validation covers the actual content of the data. CPRD provides a patient-level data quality metric in the form of a binary 'acceptability' flag.(32) This is based on recording and internal consistency of key variables including date of birth, practice registration date and transfer out date.

IQVIA Disease Analyser (DA) Germany, Germany

IQVIA Disease Analyzer (DA) Germany is a database of de-identified electronic medical records from specialized and general primary practices (GP) in Germany since 1992. This dataset encompasses approximately 3% of all outpatient practices within Germany, ensuring a substantial representation of the national healthcare landscape.(36, 37) The sampling methods used for practice selection, taking into account physician's demographics, specialty focus, community size category and federal state location, was instrumental in constructing a database that accurately mirrors the diverse spectrum of healthcare providers in the country.(36) Consequently, data within IQVIA DA Germany database has been demonstrated to be representative of general and specialised practices throughout Germany.
The database contains demographics records, basic medical data, disease diagnosis according to International Classification of Diseases, 10th revision (ICD-10), and prescription records.(37) While the database partly records information on deaths and procedures, it currently does not support linkage with external data sources meaning death recording is not comprehensive and it does not routinely contain data on hospitalisations. Routine updates are conducted at regular intervals.
IQVIA DA Germany is suitable for pharmacoepidemiologic and pharmacoeconomic studies as previously demonstrated.(37-39) The quality of data is assessed based on several criteria including completeness of information and correctness (e.g. linkage between diagnosis and prescriptions).

IQVIA Longitudinal Patient Database (LPD) Belgium, Belgium

IQVIA Longitudinal Patient Data (LPD) Belgium is a database of pseudonymized electronic medical records from general practices (GPs) in Belgium since 2005. The database encapsulates records of approximately 10% of Belgian patient population.
This patient-level database captures patient demographics, diagnoses (using a specific diagnostic coding system that can be bridged with ICD-10-CM codes). In addition, it encompasses medical history, prescription data (associated with a hard-coded diagnosis), as well as supplementary metrics such as anthropometric measures (height, weight), vital signs (blood pressure) and results from laboratory tests.(40) All patients and GPs in the database are pseudonymized and can be followed longitudinally based on a unique identifier (ID). Strict attention to confidentiality is present at every stage of data collection, storage and analysis in accordance with GDPR and Belgian Ethics Committees recommendations. IQVIA LPD Belgium database is nationally representative in terms of both geographical coverage and patient demographic characteristics, including age and sex.(40)
This database has been widely used in previous drug utilization and epidemiological studies and represents a robust source of information on primary care in Belgium.(41-43) Date of death and hospitalisations are not available.

Estonian Biobank (EBB), Estonia

The Estonian Biobank (EBB) is a population-based biobank of the Estonian Genome Center at the University of Tartu (EGCUT), encompassing close to 20% of the entire adult population of Estonia. The current biobank size is reflecting the age, sex and geographical distribution of the adult Estonian population: Estonians represent 83%; Russians 14%; and other nationalities 3% of all participants.
All participants have undergone a standardized health assessment including provision of blood samples for purification of DNA, white blood cells and plasma, and completed a questionnaire covering various health-related topics, such as lifestyle, diet and clinical diagnoses.(44) Follow-up data are available via linkage with national health-related registries and via re-examination of participants. Furthermore, electronic health records are updated for phenotypic outcome information every half year. The EBB database is regularly linked with national registries (such as Cancer Registry and Causes of Death Registry), hospital databases, and the database of the national health insurance fund, which holds treatment and service bills. Diseases and health problems are recorded as ICD-10 codes and prescribed medicine according to the ATC classification. The Estonian Biobank has been suitable for epidemiological and pharmacogenetic studies as previously demontrated.(45-49)

Information System for Research in Primary Care (SIDIAP), Catalonia, Spain

The Information System for Research in Primary Care (SIDIAP) is a dynamic database of pseudo-anonymized electronic health records of the primary care patient population in Catalonia, Spain.(50) It contains data of approximately 80% of the Catalan population registered in over 280 primary care practices throughout Catalonia since 2005.
The database contains data recorded in primary care centres on a daily basis. Additionally, it integrates data from external sources including biomarkers data from laboratories and records of drug prescription and dispensation. The dataset covers demographics, all-cause mortality, disease diagnoses classified under the International Classification of Diseases 10th revision (ICD-10), prescription and dispensation records of drugs, results of laboratory tests, socio-economic indicators, vaccination records, lifestyle information, parent–child linkage and various clinical parameters. Additional data from other data sources such as hospital discharges, mental health centres or specific disease registries can be obtained through diverse

linkages. The demographic composition within SIDIAP closely mirrors that of the broader Catalan population, encompassing a representative spectrum of geographic distribution, age, and sex proportions. The database is updated every 6 months.
SIDIAP data quality has been previously documented and SIDIAP has proved valuable for epidemiological studies.(39, 51-59) In terms of data integrity and reliability, SIDIAP has been subject to rigorous evaluation. Quality checks have been implemented including central identification of duplicate patient ID and visual inspection for temporal patterns in the recording of a certain variable. Furthermore, the data undergoes assessment for availability (longitudinality and reliability), plausibility (range checks and unusual values) and consistency using visualization tools. Specifically, for biochemistry data, consistency for measurements taken in different laboratories is assessed, and unit conversion is undertaken when needed.

Integrated Primary Care Information Project (IPCI), The Netherlands

The Integrated Primary Care Information (IPCI) database is a longitudinal observational database containing routinely collected data extracted from computer-based patient records of a selected group of general practitioners (GPs) across the Netherlands.(60) IPCI was started in 1992 by the department of Medical Informatics of the Erasmus University Medical Center in Rotterdam. The current database includes patient records from 2006 on, when the size of the database started to increase significantly. The demographic composition of the IPCI population mirrors that of the general Dutch population in terms of age and sex. Although the geographical spread is limited, GP practices are located in urban and non-urban areas. Patient-level data includes demographic information, patient's complaints and symptoms, diagnoses, laboratory test results, lifestyle factors and correspondence with secondary care, such as referral and discharge letters which are recorded by GPs. For complaints, symptoms and diagnoses, Dutch GPs use International Classification of Primary Care (ICPC-1) coding, an international standard developed and updated by the World Organization of Family Doctors' (WONCA) International Classification Committee. IPCI data quality has been previously documented and IPCI has proved valuable for epidemiological studies. Dates of birth and death are rounded to months. (61-65) In terms of quality control, extensive quality control steps are performed prior to each data release. These include comparison of patient characteristics between practices and checks to identify abnormal temporal data patterns in practices. Additional checks include over 200 indicators related to population characteristics (e.g. reliability of birth and mortality rates) and medical data (e.g. availability of durations of prescriptions, completeness of laboratory results, availability of hospital letters and prescriptions, proportion of patients with blood pressure measurement, etc).(60) Based on this information, two quality scores have been created. Practices with low scores have been excluded.
The IPCI database is registered on the European Medicines Agency (EMA) ENCePP resources database (http://www.encepp.eu). IPCI does not include linkage to hospitalisation data.

8.3 Study Period

The study period will be from 01/01/2017 to 31/12/2022.

8.4 Follow-up

Subjects included in the study will be followed up to one year from cancer diagnosis (index date) in order to calculate one-year hospitalisation and mortality rates. Subjects will be followed up until the earliest of the following: 1) loss to follow-up, 2) end of data availability, 3) date of death, or 4) one year after index date.

Table 4: Operational Definition of Time 0 (index date) and other primary time anchors

Study population name(s)	Time Anchor Description	Number of entries	Type of entry	Washout window	Care Setting[1]	Code Type	Diagnosis position	Incident with respect to…	Measurement characteristics /validation	Source of algorithm
All persons in each database eligible for the study	Date of selected cancer diagnosis	Single entry	Incident	Any time prior to cancer diagnosis	OP, OT	SNOMED	Any	Any cancer diagnosis except non-melanoma skin cancer	N/A	N/A

[1] OP = outpatient, OT = other, n/a = not applicable

8.5 Study Population with inclusion and exclusion criteria

The study population for patient characterisation will include all individuals aged 18 years and above with a primary diagnosis of a selected cancer (lung, breast, ovary, endometrium, prostate, pancreas, colorectal cancer, lymphoma, leukemia and myeloma) recorded between 01/01/2017 and 31/12/2022, with at least one year of prior history available before cancer diagnosis. Individuals with a diagnosis of cancer (any, excluding non-melanoma skin cancer) any time prior to the diagnosis of one of the selected cancers will be excluded. Cancer cases will be identified based on a recorded code indicating a diagnosis or observation for each specific cancer. Conditions in the OMOP CDM use the Systematized Nomenclature of Medicine (SNOMED) as the standard vocabulary for diagnosis codes. A preliminary code list for each cancer is provided in **Appendix 1 – Table 1**.

Additional eligibility of a minimum of 1 year of potential follow-up time will be imposed for the estimation of one-year hospitalisation and mortality rates, which will be 1 year prior to the data lock date as described in Table 3.

Table 5. Operational Definitions of Inclusion Criteria

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type	Diagnosis position	Applied to study populations:	Measurement characteristics /validation	Source for algorithm
Prior database history of 1 year (objective 1)	Study participants will be required to have a year of prior history observed before contributing observation time	After	1 year	OP, OT	N/A	N/A	All study participants	N/A	N/A
Minimum potential follow-up time (only for calculation of one year hospitalisation and mortality rates)	Only participants with a cancer diagnosis (index date) occurring one year prior to end of data availability in the database will be included	After	1 year	OP, OT	N/A	N/A	All study participants	N/A	N/A

[1] OP = outpatient, OT = other, n/a = not applicable

Table 6. Operational Definitions of Exclusion Criteria

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type	Diagnosis position[2]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
History of cancer diagnosis	Participants with a diagnosis of cancer (any, excluding non-melanoma skin cancer) any time prior to the recorded code of the selected cancer or prior to the start of the study period	After	Any time prior to cancer diagnosis	OP, OT	SNOMED	Any	All study participants	N/A	N/A

[1] OP = outpatient, OT = other, n/a = not applicable

8.6 Variables

8.6.1. Exposure/s

The prevalence of polypharmacy and of frailty will be calculated. The number of concomitant prescriptions will be calculated based on the overall number of prescriptions during the 90 days prior to index date, and polypharmacy will be measured as the maximum number of drug eras that overlap on any day during the 90-day period, using two definitions as the concomitant prescription of >= 5 and >= 10 medications (ingredient level) anytime during the 90 days prior to cancer diagnosis.

A frailty score will be created based on the presence of polypharmacy (defined as >=5 prescriptions of medications as mentioned above) and the following conditions included in frailty indexes previously proposed in the literature (i.e.: eFI or eFRAGICAP) (3, 67):

- Mobility and transfer problems
- Housebound
- Activity limitation
- Visual impairment
- Hearing impairment
- Requirement for care
- Social vulnerability
- Falls
- Urinary incontinence
- Weight loss and anorexia
- Memory and cognitive problems
- Dizziness
- Dyspnoea
- Sleep disturbance
- Anaemia and haematinic deficiency
- Hypertension
- Ischaemic heart disease
- Heart failure
- Cerebrovascular disease
- Peripheral vascular disease
- Atrial fibrillation
- Heart valve disease
- Hypotension/syncope
- Diabetes
- Foot problems
- Arthritis
- Respiratory disease
- Peptic ulcer
- Thyroid disease
- Chronic kidney disease
- Osteoporosis
- Fragility fracture
- Parkinsonism and tremor
- Urinary system disease
- Skin ulcer

Their definition will be based on SNOMED diagnosis codes recorded any time prior or at cancer diagnosis. A preliminary code list for condition is provided in **Appendix 1 – Table 2**.

The score will be calculated based on the number of present conditions and polypharmacy divided by 36, the total number of conditions/polypharmacy mentioned above (35 conditions and 1 for polypharmacy

prevalence) and will be further categorised accordingly into the following levels of severity according to the eFI definition: fit: 0–0.12; mild frailty: >0.12–0.24; moderate frailty= >0.24–0.36; severely frail: >=0.36.

Table 7. Operational Definitions of Exposure

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position	Applied to study populations:	Measurement characteristics /validation	Source for algorithm
Polypharmacy	The concomitant prescription of >= 5 or >= 10 medications (ingredient level, maximum number of drug eras that overlap on any day during the 90 day period),	Counts	90 days prior to ID	OP, OT	RxNorm	N/A	All study participants	N/A	N/A
Score of frailty conditions	See description in section 8.6.1 above	Counts	Any time prior or at ID	OP, OT	SNOMED	N/A	All study participants	N/A	N/A

[1] ID = index date, P = outpatient, OT = other, n/a = not applicable

8.6.2. Outcome/s

N/A

8.6.3. Other covariates, including confounders, effect modifiers and other variables

Age at cancer diagnosis will be described. The following age grouping will be used: 18-44, 45-64, 65-74, 75-84 and 85 and over. The sex (male/ female) of study participants will also be identified.

The following cancer sites will be further categorised, as follows:

- Lung cancer: NSCLC and all other lung cancers
- Leukaemia: Lymphocytic leukaemia: ALL and CLL (acute and chronic will be described individually) and myeloid leukaemia: AML and CML (acute and chronic will be described individually)
- Lymphomas: Hodgkin and non-Hodgkin lymphoma

All co-morbidities and medications will be used for large-scale patient characterisation, identified as concept/code and descendants. Other variables of interest will include number of hospitalisation and date of death during the year after date of cancer diagnosis.

Table 8. Operational Definitions of Covariates

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type	Diagnosis Position	Applied to study populations:	Measurement characteristics /validation	Source for algorithm
Co-morbidities	Large-scale patient-level characterisation with regard to underlying comorbidities	Counts	At index date (ID), for 30 to 1 day before ID, for 365 to 31 days before ID, at any time and up to 366 days before ID	OP, OT	SNOMED	N/A	All study participants	N/A	N/A
Concomitant medication	Large-scale patient-level characterisation with regard to use of concomitant drugs	Counts	At index date (ID), for 30 to 1 day before ID, for 365 to 31 days before ID, 1 to 30 post ID, 1 to 90 post ID, and 1 to 365 days post ID	OP, OT	RxNorm	N/A	All study participants	N/A	N/A
Hospitalisation and date of death	One-year hospitalization rate and mortality rate will be calculated after ID	Counts	1-year post-index	OP, OT	N/A	N/A	All study participants	N/A	N/A

[1] ID = index date, P = outpatient, OT = other, n/a = not applicable

8.7 Study size

No sample size will be calculated as this is a descriptive Disease Epidemiology Study where we are interested in the characteristics of all incident cases of selected cancers in each database.

8.8 Analysis

Table 9. Description of Study Types and Type of analysis

STUDY TYPE	STUDY CLASSIFICATION	TYPE OF ANALYSIS
Population-level descriptive epidemiology	Off-the-shelf (C1)	- Prevalence rates of the condition of interest
Patient-level characterisation	Off-the-shelf (C1)	- Large-scale characterisation - Patient-level characteristics - Prognosis / progression to a pre-specified outcome

8.8.1 Federated Network Analyses

Analyses will be conducted separately for each database. Before study initiation, test runs of the analytics are performed on a subset of the data sources or on a simulated set of patients and quality control checks are performed. Once all the tests are passed, the final package is released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally execute the analytics against the OMOP-CDM in R Studio and review and approve the by default aggregated results before returning them to the Coordination Centre. Sometimes multiple execution iterations are performed, and additional fine tuning of the code base is needed. A service desk will be available during the study execution for support.

The study results of all data sources are checked after which they are made available to the team in the Digital Research Environment (DRE) and the Study Dissemination Phase can start. All results are locked and timestamped for reproducibility and transparency.

8.8.2 Patient privacy protection

Cell suppression will be applied as required by databases to protect people's privacy. Cell counts < 5 will be masked.

8.8.3 Statistical model specification and assumptions of the analytical approach considered

Data analyses

The prevalence of the score of frailty conditions (based on the definition described in section 8.6.1) and polypharmacy (objective 1) will be estimated at the time of cancer diagnosis.

Large-scale patient-level characterisation (objective 2) will be conducted for individuals with different frailty (fit, mild, moderate, severe frailty) and polypharmacy categories (no polypharmacy, >=5 medications, >=10 medications) as follows: Age and sex at time of cancer diagnosis will be described.

Medical history will assessed for anytime –and up to 365 days before index date, for 365 to 31 days before index date, for 30 to 1 day before index date, and at index date; Medication use history will be reported for 365 to 31 days before index date, for 30 to 1 day before index date, and at index date; hospitalization and mortality rates will be calculated for up to 365 days after index date.

For all analyses absolute and relative frequencies will be reported. A minimum cell count of 5 will be used when reporting results, with any smaller counts reported as "<5" and zero counts as "0". Analyses will be done separately for each database and selected cancer. Next to overall reporting, stratification by age category (18-44, 45-64, 65-74, 75-84 and 85+) and sex will be conducted when possible (minimum cell count reached). Results for objective one on the estimation of the prevalence of frailty in cancer patients, results will be further stratified by time period (before or after 2020).

Missing data

As sates in Section 8.4, subjects will be followed-up from the date of cancer diagnosis until the earliest of the following: 1) loss to follow-up, 2) end of data availability, 3) date of death, or 4) one year after index date.

Patients in 1) and 2) will have missing part of their follow-up. They will be censored at the time of loss to follow-up or end of data availability and that reported hospitalisation and mortality rates will implicitly assume censoring occurs at random. Patients in 4) will be administratively censored.

R-packages

We will use the R packages "IncidencePrevalence" (https://github.com/darwin-eu/IncidencePrevalence) for the computation of prevalence, "PatientProfiles" (https://github.com/darwin-eu-dev/PatientProfiles) for the patient-level characterization of demographics and clinical characteristics.

8.9 Evidence synthesis

Results from analyses described in section 8.8 will be presented separately for each database and no meta-analysis of results will be conducted.

9. DATA MANAGEMENT

9.1 Data management

All databases are mapped to the OMOP common data model. This enables the use of standardised analytics and tools across the network since the structure of the data and the terminology system is harmonised. The OMOP CDM is developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org

The analytic code for this study will be written in R. Each data partner will execute the study code against their database containing patient-level data and will then return the results set which will only contain

aggregated data. The results from each of the contributing data sites will then be combined in tables and figures for the study report.

9.2 Data storage and protection

For this study, participants from various EU member states will process personal data from individuals which is collected in national/regional electronic health record databases. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All databases used in this study are already used for pharmaco-epidemiological research and have a well-developed mechanism to ensure that European and local regulations dealing with ethical use of the data and adequate privacy control are adhered to. In agreement with these regulations, rather than combining person level data and performing only a central analysis, local analyses will be run, which generate non-identifiable aggregate summary results.

The output files are stored in the DARWIN Digital Research Environment. These output files do not contain any data that allow identification of subjects included in the study. The DRE implements further security measures in order to ensure a high level of stored data protection to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/20161 in the various member states.

10. QUALITY CONTROL

General database quality control

A number of open-source quality control mechanisms for the OMOP CDM have been developed (see Chapter 15 of The Book of OHDSI http://book.ohdsi.org/DataQuality.html). In particular, data partners have run the OHDSI Data Quality Dashboard tool (https://github.com/OHDSI/DataQualityDashboard). This tool provides numerous checks relating to the conformance, completeness and plausibility of the mapped data. Conformance focuses on checks that describe the compliance of the representation of data against internal or external formatting, relational, or computational definitions, completeness in the sense of data quality is solely focused on quantifying missingness, or the absence of data, while plausibility seeks to determine the believability or truthfulness of data values. Each of these categories has one or more subcategories and are evaluated in two contexts: validation and verification. Validation relates to how well data align with external benchmarks with expectations derived from known true standards, while verification relates to how well data conform to local knowledge, metadata descriptions, and system assumptions.

Study specific quality control

When defining specific cancers and co-morbidities, a systematic search of possible codes for inclusion was previously identified using CodelistGenerator R package (https://github.com/darwin-eu/CodelistGenerator). This software allows the user to define a search strategy and using this will then query the vocabulary tables of the OMOP CDM so as to find potentially relevant codes. The codes returned will be then reviewed by two clinical epidemiologist to consider their relevance. In addition, the CohortDiagnostics R package (https://github.com/OHDSI/CohortDiagnostics) will be run to assess the use of different codes across the databases contributing to the study and identify any codes potentially omitted in error. This will allow for a consideration of the validity of the study cohort of patients with the selected

cancers and co-morbidities in each of the databases, and inform decisions around whether multiple definitions are required.

The study code will be based on three R packages currently being developed to (1) estimate prevalence of polypharmacy and co-morbidities (IncidencePrevalence), (2) characterise demographic and clinical characteristics (PatientProfiles), (3) estimate 1-year hospitalisation and mortality rates (CohortSuvival). These packages will include numerous automated unit tests to ensure the validity of the codes, alongside software peer review and user testing. The R package will be made publicly available via GitHub.

11. LIMITATIONS OF THE RESEARCH METHODS

The study will be informed by routinely collected health care data and so data quality issues must be considered. In particular, the identification of cancer patients and the recording of the co-morbidities may vary across databases and while relatively few false positives would be expected, false negatives may be more likely especially for databases without patient-level linkage to secondary care data. The impact of this may be to underestimate their prevalence. Nevertheless, the cancer diagnosis has been previously validated in the SIDIAP and CPRD databases (50, 66), and the EEB database is linked to cancer registry data which is known to contain high quality data on cancer diagnoses.

Additionally, the recording of hospitalisation and date of death varies across databases. For example, hospitalisation rates will only be estimated in SIDIAP and EBB that utilises linked data, and date of death is not available in IQVIA DA Germany and IQVIA LPD Belgium. Therefore, the one-year hospitalisation and mortality rates will not be calculated in all the participating databases.

Polypharmacy is defined as the use of concomitant medications. For the polypharmacy definition we will identify medication use by counting the maximum number of drugs by ingredient level issued or dispensed on any given day within a time window prior to index date. We have not considered use of more than one drug from the same class in the count as patients may have switched during this time period rather than receive concomitant use. We also cannot be certain that all use is concomitant during this time period with the risk of misclassification being greater as look back time periods increase. A 90-day window was considered reasonable based upon experience of repeat drug prescribing typically occurring in primary care.

Finally, it is important to note that although the frailty score created in this study is based on the validated frailty indexes eFI and eFRAGICAP, in this study we will only use conditions for the definition of the index. This might underestimate frailty in our population and therefore, might also have implications in the categorisation of patients based on the severity of frailty. Also, the recording of conditions included in the score of frailty might be incomplete in most databases. The included conditions are based on the list of conditions which are frequently used for the computation of frailty indexes (3, 67), however, some of these conditions might not always be recorded using diagnoses codes within databases; therefore, this information will not be included in this study with the impact that frailty may be underestimated.

12. MANAGEMENT AND REPORTING OF ADVERSE EVENTS/ADVERSE REACTIONS

Adverse events/adverse reactions will not be collected or analysed as part of this evaluation. The nature of this non-interventional evaluation, through the use of secondary data, does not fulfil the criteria for reporting adverse events, according to module VI, VI.C.1.2.1.2 of the Good Pharmacovigilance Practices

(https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports_en.pdf).
Only in case of prospective data collection, there is a need to describe the procedures for the collection, management and reporting of individual cases of adverse events/adverse reactions.

13. GOVERNANCE BOARD ASPECTS

All data sources require approval from their respective IRB boards, with the exception of IQVIA DA Germany and IQVIA LPD Belgium which will not require any further specific approvals to undertake this study.

14. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

14.1 Study Report

A PDF report including an executive summary, and the specified tables and/or figures will be submitted to EMA by the DARWIN EU® CC upon completion of the study.

An interactive dashboard incorporating all the results (tables and figures) will be provided alongside the pdf report. The full set of underlying aggregated data used in the dashboard will also be made available if requested.

15. OTHER ASPECTS

N/A

16. REFERENCES

1. Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med. 2011;27(1):1-15.
2. O'Caoimh R, Galluzzo L, Rodríguez-Laso Á, Van der Heyden J, Ranhoff AH, Lamprini-Koula M, et al. Prevalence of frailty at population level in European ADVANTAGE Joint Action Member States: a systematic review and meta-analysis. Ann Ist Super Sanita. 2018;54(3):226-38.
3. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age Ageing. 2016;45(3):353-60.
4. Guthrie B, Makubate B, Hernandez-Santiago V, Dreischulte T. The rising tide of polypharmacy and drug-drug interactions: population database analysis 1995–2010. BMC Medicine. 2015;13(1):74.
5. Gallagher P, Lang PO, Cherubini A, Topinková E, Cruz-Jentoft A, Montero Errasquín B, et al. Prevalence of potentially inappropriate prescribing in an acutely ill population of older patients admitted to six European hospitals. Eur J Clin Pharmacol. 2011;67(11):1175-88.
6. Palmer K, Villani ER, Vetrano DL, Cherubini A, Cruz-Jentoft AJ, Curtin D, et al. Association of polypharmacy and hyperpolypharmacy with frailty states: a systematic review and meta-analysis. Eur Geriatr Med. 2019;10(1):9-36.

7. Hsu HF, Chen KM, Belcastro F, Chen YF. Polypharmacy and pattern of medication use in community-dwelling older adults: A systematic review. J Clin Nurs. 2021;30(7-8):918-28.
8. Doody P, Lord JM, Greig CA, Whittaker AC. Frailty: Pathophysiology, Theoretical and Operational Definition(s), Impact, Prevalence, Management and Prevention, in an Increasingly Economically Developed and Ageing World. Gerontology. 2023;69(8):927-45.
9. Ribeiro AR, Howlett SE, Fernandes A. Frailty-A promising concept to evaluate disease vulnerability. Mech Ageing Dev. 2020;187:111217.
10. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet. 2019;394(10206):1376-86.
11. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381(9868):752-62.
12. Handforth C, Clegg A, Young C, Simpkins S, Seymour MT, Selby PJ, et al. The prevalence and outcomes of frailty in older cancer patients: a systematic review. Ann Oncol. 2015;26(6):1091-101.
13. Ethun CG, Bilen MA, Jani AB, Maithel SK, Ogan K, Master VA. Frailty and cancer: Implications for oncology surgery, medical oncology, and radiation oncology. CA Cancer J Clin. 2017;67(5):362-77.
14. Fizazi K, Tran N, Fein L, Matsubara N, Rodriguez-Antolin A, Alekseev BY, et al. Abiraterone acetate plus prednisone in patients with newly diagnosed high-risk metastatic castration-sensitive prostate cancer (LATITUDE): final overall survival analysis of a randomised, double-blind, phase 3 trial. Lancet Oncol. 2019;20(5):686-700.
15. Beer TM, Armstrong AJ, Rathkopf DE, Loriot Y, Sternberg CN, Higano CS, et al. Enzalutamide in metastatic prostate cancer before chemotherapy. N Engl J Med. 2014;371(5):424-33.
16. Moore K, Colombo N, Scambia G, Kim BG, Oaknin A, Friedlander M, et al. Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N Engl J Med. 2018;379(26):2495-505.
17. Coleman RL, Oza AM, Lorusso D, Aghajanian C, Oaknin A, Dean A, et al. Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2017;390(10106):1949-61.
18. Ramsdale E, Mohamed M, Yu V, Otto E, Juba K, Awad H, et al. Polypharmacy, Potentially Inappropriate Medications, and Drug-Drug Interactions in Vulnerable Older Adults With Advanced Cancer Initiating Cancer Treatment. Oncologist. 2022;27(7):e580-e8.
19. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci. 2007;62(7):722-7.
20. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):M146-56.
21. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. Cmaj. 2005;173(5):489-95.
22. Jolly TA, Deal AM, Nyrop KA, Williams GR, Pergolotti M, Wood WA, et al. Geriatric assessment-identified deficits in older cancer patients with normal performance status. Oncologist. 2015;20(4):379-85.

23. Subramaniam A, Ueno R, Tiruvoipati R, Darvall J, Srikanth V, Bailey M, et al. Defining ICD-10 surrogate variables to estimate the modified frailty index: a Delphi-based approach. BMC Geriatr. 2022;22(1):422.
24. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet. 2018;391(10132):1775-82.
25. Duchesneau ED, Shmuel S, Faurot KR, Park J, Musty A, Pate V, et al. Translation of a Claims-Based Frailty Index From the International Classification of Diseases, Ninth Revision, Clinical Modification to the Tenth Revision. Am J Epidemiol. 2023;192(12):2085-93.
26. Lichtman SM, Boparai MK. Anticancer drug therapy in the older cancer patient: pharmacology and polypharmacy. Curr Treat Options Oncol. 2008;9(2-3):191-203.
27. Riechelmann RP, Tannock IF, Wang L, Saad ED, Taback NA, Krzyzanowska MK. Potential drug interactions and duplicate prescriptions among cancer patients. J Natl Cancer Inst. 2007;99(8):592-600.
28. Riechelmann RP, Zimmermann C, Chin SN, Wang L, O'Carroll A, Zarinehbaf S, et al. Potential drug interactions in cancer patients receiving supportive care exclusively. J Pain Symptom Manage. 2008;35(5):535-43.
29. Tam-McDevitt J. Polypharmacy, aging, and cancer. Oncology (Williston Park). 2008;22(9):1052-5, discussion 5, 8, 60.
30. Maggiore RJ, Gross CP, Hurria A. Polypharmacy in older adults with cancer. Oncologist. 2010;15(5):507-22.
31. Mohamed MR, Mohile SG, Juba KM, Awad H, Wells M, Loh KP, et al. Association of polypharmacy and potential drug-drug interactions with adverse treatment outcomes in older adults with advanced cancer. Cancer. 2023;129(7):1096-104.
32. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015;44(3):827-36.
33. Carey IM, Nirmalananthan N, Harris T, DeWilde S, Chaudhry UAR, Limb E, et al. Prevalence of co-morbidity and history of recent infection in patients with neuromuscular disease: A cross-sectional analysis of United Kingdom primary care data. PLoS One. 2023;18(3):e0282513.
34. Wigglesworth S, Neligan A, Dickson JM, Pullen A, Yelland E, Anjuman T, et al. The incidence and prevalence of epilepsy in the United Kingdom 2013-2018: A retrospective cohort study of UK primary care data. Seizure. 2023;105:37-42.
35. Fahmi A, Wong D, Walker L, Buchan I, Pirmohamed M, Sharma A, et al. Combinations of medicines in patients with polypharmacy aged 65-100 in primary care: Large variability in risks of adverse drug related and emergency hospital admissions. PLoS One. 2023;18(2):e0281466.
36. Rathmann W, Bongaerts B, Carius HJ, Kruppert S, Kostev K. Basic characteristics and representativeness of the German Disease Analyzer database . Int J Clin Pharmacol Ther. 2018;56(10):459-66.
37. Zappacosta S, Cascarano A, Konrad M, Tanislav C, Kostev K. Association between COVID-19 and subsequent vascular events in primary care patients in Germany. Public Health. 2022;213:107-13.
38. Tanislav C, Rosenbauer J, Zingel R, Kostev K. No increased incidence of venous thrombosis or pulmonary embolism after SARS-CoV-2 vaccination in Germany. Public Health. 2022;207:14-8.

39. Ly NF, Flach C, Lysen TS, Markov E, van Ballegooijen H, Rijnbeek P, et al. Impact of European Union Label Changes for Fluoroquinolone-Containing Medicinal Products for Systemic and Inhalation Use: Post-Referral Prescribing Trends. Drug Saf. 2023;46(4):405-16.
40. Vandenhaute J, Tsakeu E, Chevalier P, Pawaskar M, Benčina G, Vertriest J. Assessing the use of antibiotics and the burden of varicella in Belgium using a retrospective GP database analysis. BMC Infect Dis. 2021;21(1):1150.
41. Levi M, Rosselli M, Simonetti M, Brignoli O, Cancian M, Masotti A, et al. Epidemiology of iron deficiency anaemia in four European countries: a population-based study in primary care. Eur J Haematol. 2016;97(6):583-93.
42. Levi M, Simonetti M, Marconi E, Brignoli O, Cancian M, Masotti A, et al. Gender differences in determinants of iron-deficiency anemia: a population-based study conducted in four European countries. Ann Hematol. 2019;98(7):1573-82.
43. Cai T, Palagin I, Brunelli R, Cipelli R, Pellini E, Truzzi JC, et al. Office-based approach to urinary tract infections in 50 000 patients: results from the REWIND study. Int J Antimicrob Agents. 2020;56(1):105966.
44. Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. Int J Epidemiol. 2015;44(4):1137-47.
45. Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, Posner DC, et al. Publisher Correction: Stroke genetics informs drug discovery and risk prediction across ancestries. Nature. 2022;612(7938):E7.
46. Krebs K, Bovijn J, Zheng N, Lepamets M, Censin JC, Jürgenson T, et al. Genome-wide Study Identifies Association between HLA-B(∗)55:01 and Self-Reported Penicillin Allergy. Am J Hum Genet. 2020;107(4):612-21.
47. Tasa T, Krebs K, Kals M, Mägi R, Lauschke VM, Haller T, et al. Genetic variation in the Estonian population: pharmacogenomics study of adverse drug effects using electronic health records. Eur J Hum Genet. 2019;27(3):442-54.
48. Reisberg S, Krebs K, Lepamets M, Kals M, Mägi R, Metsalu K, et al. Translating genotype data of 44,000 biobank participants into clinical pharmacogenetic recommendations: challenges and solutions. Genet Med. 2019;21(6):1345-54.
49. Kasela S, Kisand K, Tserel L, Kaleviste E, Remm A, Fischer K, et al. Pathogenic implications for autoimmune mechanisms derived by comparative eQTL analysis of CD4+ versus CD8+ T cells. PLoS Genet. 2017;13(3):e1006643.
50. Recalde M, Manzano-Salgado CB, Díaz Y, Puente D, Garcia-Gil MDM, Marcos-Gragera R, et al. Validation Of Cancer Diagnoses In Electronic Health Records: Results From The Information System For Research In Primary Care (SIDIAP) In Northeast Spain. Clin Epidemiol. 2019;11:1015-24.
51. Recalde M, Davila-Batista V, Díaz Y, Leitzmann M, Romieu I, Freisling H, et al. Body mass index and waist circumference in relation to the risk of 26 types of cancer: a prospective cohort study of 3.5 million adults in Spain. BMC Med. 2021;19(1):10.
52. Ramos R, Comas-Cufí M, Martí-Lluch R, Balló E, Ponjoan A, Alves-Cabratosa L, et al. Statins for primary prevention of cardiovascular events and mortality in old and very old adults with and without type 2 diabetes: retrospective cohort study. Bmj. 2018;362:k3359.

53. Burn E, Tebé C, Fernandez-Bertolin S, Aragon M, Recalde M, Roel E, et al. The natural history of symptomatic COVID-19 during the first wave in Catalonia. Nat Commun. 2021;12(1):777.
54. Mata-Cases M, Franch-Nadal J, Real J, Gratacòs M, López-Simarro F, Khunti K, et al. Therapeutic inertia in patients treated with two or more antidiabetics in primary care: Factors predicting intensification of treatment. Diabetes Obes Metab. 2018;20(1):103-12.
55. Lane JC, Butler KL, Poveda-Marina JL, Martinez-Laguna D, Reyes C, de Bont J, et al. Preschool Obesity Is Associated With an Increased Risk of Childhood Fracture: A Longitudinal Cohort Study of 466,997 Children and Up to 11 Years of Follow-up in Catalonia, Spain. J Bone Miner Res. 2020;35(6):1022-30.
56. Monteagudo M, Nuñez A, Solntseva I, Dhalwani N, Booth A, Barrecheguren M, et al. Treatment Pathways Before and After Triple Therapy in COPD: A Population-based Study in Primary Care in Spain. Arch Bronconeumol (Engl Ed). 2021;57(3):205-13.
57. Ortega Y, Aragonès E, Piñol JL, Basora J, Araujo A, Cabré JJ. Impact of depression and/or anxiety on the presentation of cardiovascular events in a cohort with metabolic syndrome. StreX project: Five years of follow-up. Prim Care Diabetes. 2018;12(2):163-71.
58. Troncoso-Mariño A, Roso-Llorach A, López-Jiménez T, Villen N, Amado-Guirado E, Fernández-Bertolin S, et al. Medication-Related Problems in Older People with Multimorbidity in Catalonia: A Real-World Data Study with 5 Years' Follow-Up. J Clin Med. 2021;10(4).
59. Braeye T, Emborg HD, Llorente-García A, Huerta C, Martín-Merino E, Duarte-Salles T, et al. Age-specific vaccination coverage estimates for influenza, human papillomavirus and measles containing vaccines from seven population-based healthcare databases from four EU countries - The ADVANCE project. Vaccine. 2020;38(16):3243-54.
60. de Ridder MAJ, de Wilde M, de Ben C, Leyba AR, Mosseveld BMT, Verhamme KMC, et al. Data Resource Profile: The Integrated Primary Care Information (IPCI) database, The Netherlands. Int J Epidemiol. 2022;51(6):e314-e23.
61. Ali MS, Berencsi K, Marinier K, Deltour N, Perez-Guthann S, Pedersen L, et al. Comparative cardiovascular safety of strontium ranelate and bisphosphonates: a multi-database study in 5 EU countries by the EU-ADR Alliance. Osteoporos Int. 2020;31(12):2425-38.
62. Baan EJ, van den Akker ELT, Engelkes M, de Rijke YB, de Jongste JC, Sturkenboom M, et al. Hair cortisol and inhaled corticosteroid use in asthmatic children. Pediatr Pulmonol. 2020;55(2):316-21.
63. Berencsi K, Sami A, Ali MS, Marinier K, Deltour N, Perez-Gutthann S, et al. Impact of risk minimisation measures on the use of strontium ranelate in Europe: a multi-national cohort study in 5 EU countries by the EU-ADR Alliance. Osteoporos Int. 2020;31(4):721-55.
64. Engelkes M, Baan EJ, de Ridder MAJ, Svensson E, Prieto-Alhambra D, Lapi F, et al. Incidence, risk factors and re-exacerbation rate of severe asthma exacerbations in a multinational, multidatabase pediatric cohort study. Pediatr Allergy Immunol. 2020;31(5):496-505.
65. James G, Collin E, Lawrance M, Mueller A, Podhorna J, Zaremba-Pechmann L, et al. Treatment pathway analysis of newly diagnosed dementia patients in four electronic health record databases in Europe. Soc Psychiatry Psychiatr Epidemiol. 2021;56(3):409-16.
66. Margulis AV, Fortuny J, Kaye JA, Calingaert B, Reynolds M, Plana E, et al. Validation of Cancer Cases Using Primary Care, Cancer Registry, and Hospitalization Data in the United Kingdom. Epidemiology. 2018;29(2):308-13.

67. Orfila F, Carrasco-Ribelles LA, Abellana R, Roso-Llorach A, Cegri F, Reyes C, et al. Validation of an electronic frailty index with electronic health records: eFRAGICAP index. BMC Geriatrics. 2022;22(1):404.

17. ANNEXES

Appendix I: Preliminary list of codes

Table 1. Preliminary code list for primary cancers

Cancer	concept_id	concept_name
Lung; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788638/conceptsets/	3654352	Anaplastic lymphoma kinase fusion oncogene positive non-small cell lung cancer
	3654297	Anaplastic lymphoma kinase fusion oncogene negative non-small cell lung cancer
	4110705	Squamous cell carcinoma of lung
	3654301	Reactive oxygen species 1 negative non-small cell lung cancer
	4115276	Non-small cell lung cancer
	4140471	Epidermal growth factor receptor negative non-small cell lung cancer
	4143825	Epidermal growth factor receptor positive non-small cell lung cancer
	36716426	Reactive oxygen species 1 positive non-small cell lung cancer
	37109576	Squamous non-small cell lung cancer
	45766129	Non-small cell lung cancer with mutation in epidermal growth factor receptor
	45766131	Non-small cell lung cancer without mutation in epidermal growth factor receptor
	602150	Primary adenocarcinoma of middle lobe of right lung
	602159	Primary small cell carcinoma of lower lobe of left lung
	602161	Primary small cell carcinoma of upper lobe of left lung
	602163	Primary squamous cell carcinoma of upper lobe of left lung
	602167	Primary squamous cell carcinoma of lower lobe of right lung
	602169	Primary squamous cell carcinoma of middle lobe of right lung
	602171	Primary squamous cell carcinoma of upper lobe of right lung
	602696	Primary small cell carcinoma of upper lobe of right lung
	602697	Primary squamous cell carcinoma of lower lobe of left lung
	602682	Primary large cell carcinoma of lower lobe of right lung
	602695	Primary small cell carcinoma of middle lobe of right lung
	605476	Primary small cell carcinoma of lower lobe of right lung
	605821	Non-small cell lung carcinoma with NRG1 fusion
	608925	Primary small cell carcinoma of right lung
	608926	Primary small cell carcinoma of left lung
	609075	Primary non-small cell carcinoma of lower lobe of left lung
	609076	Primary non-small cell carcinoma of middle lobe of right lung

Cancer	concept_id	concept_name
	609077	Primary non-small cell carcinoma of upper lobe of left lung
	609078	Primary non-small cell carcinoma of lower lobe of right lung
	609079	Primary non-small cell carcinoma of right lung
	609080	Primary non-small cell carcinoma of left lung
	609081	Primary non-small cell carcinoma of upper lobe of right lung
	605445	Primary large cell carcinoma of lower lobe of left lung
	605446	Primary large cell carcinoma of upper lobe of left lung
	605447	Primary large cell carcinoma of upper lobe of right lung
	4110589	Large cell carcinoma of lung
	4110590	Giant cell carcinoma of lung
	4110591	Small cell carcinoma of lung
	4110705	Squamous cell carcinoma of lung
	4112738	Adenocarcinoma of lung
	4112739	Oat cell carcinoma of lung
	4155293	Carcinoma of lower lobe, bronchus or lung
	4308784	Adenocarcinoma of lung, stage I
	4312274	Squamous cell carcinoma of lung, TNM stage 4
	4314040	Large cell carcinoma of lung, TNM stage 1
	4314156	Adenocarcinoma of lung, stage IV
	4314172	Non-small cell carcinoma of lung, TNM stage 2
	4314832	Adenocarcinoma of lung, stage II
	4307118	Large cell carcinoma of lung, TNM stage 2
	4308479	Non-small cell carcinoma of lung, TNM stage 4
	4310703	Non-small cell carcinoma of lung, TNM stage 1
	4311452	Adenocarcinoma of lung, stage III
	4311997	Non-small cell carcinoma of lung, TNM stage 3
	4312768	Large cell carcinoma of lung, TNM stage 3
	4310448	Squamous cell carcinoma of lung, TNM stage 1
	4313200	Squamous cell carcinoma of lung, TNM stage 2
	4313751	Large cell carcinoma of lung, TNM stage 4
	4322387	Squamous cell carcinoma of lung, TNM stage 3
	36686537	Large cell carcinoma of left lung
	36686538	Large cell carcinoma of right lung
	36712815	Squamous cell carcinoma of left lung
	36712816	Squamous cell carcinoma of right lung
	36712981	Adenocarcinoma of right lung
	36717017	Primary adenocarcinoma of upper lobe of right lung
	36712707	Primary adenocarcinoma of lower lobe of left lung
	36712708	Primary adenocarcinoma of upper lobe of left lung
	36712709	Primary adenocarcinoma of lower lobe of right lung
	36713366	Extensive stage primary small cell carcinoma of lung

Cancer	concept_id	concept_name
	37311684	Acinar cell cystadenocarcinoma of lung
	40492938	Carcinoma of lung
	42539251	Adenocarcinoma of left lung
	45768916	Primary adenocarcinoma of lung
	45768917	Primary mucinous adenocarcinoma of lung
	45768918	Primary clear cell squamous cell carcinoma of lung
	45768919	Primary basaloid squamous cell carcinoma of lung
	45768920	Primary undifferentiated carcinoma of lung
	45768921	Primary spindle cell carcinoma of lung
	45768922	Primary pleomorphic carcinoma of lung
	45768923	Primary pseudosarcomatous carcinoma of lung
	45768927	Primary myoepithelial carcinoma of lung
	45768928	Primary adenoid cystic carcinoma of lung
	45768929	Primary salivary gland type carcinoma of lung
	45768930	Primary mixed mucinous and non-mucinous bronchiolo-alveolar carcinoma of lung
	45768931	Primary non-mucinous bronchiolo-alveolar carcinoma of lung
	45768932	Primary mucinous bronchiolo-alveolar carcinoma of lung
	45769034	Primary mucinous cystadenocarcinoma of lung
	45768879	Primary fetal adenocarcinoma of lung
	45768880	Primary mixed subtype adenocarcinoma of lung
	45768881	Primary adenosquamous carcinoma of lung
	45768883	Primary small cell non-keratinizing squamous cell carcinoma of lung
	45768884	Primary acinar cell carcinoma of lung
	45768885	Primary solid carcinoma of lung
	45768886	Primary papillary adenocarcinoma of lung
	45772933	Primary signet ring cell carcinoma of lung
	45772938	Primary papillary squamous cell carcinoma of lung
	45772939	Primary mucoepidermoid carcinoma of lung
	46272955	Primary clear cell adenocarcinoma of lung
	4111804	Squamous cell carcinoma of trachea
	4112736	Adenoid cystic carcinoma of trachea
	45768933	Primary clear cell adenocarcinoma of trachea
	45768934	Primary papillary adenocarcinoma of trachea
	45768935	Primary mucinous adenocarcinoma of trachea
	45768936	Primary adenocarcinoma of trachea
	45768855	Primary signet ring cell carcinoma of trachea
	45768856	Primary myoepithelial carcinoma of trachea
	45768857	Primary mucoepidermoid carcinoma of trachea
	45768858	Primary salivary gland type carcinoma of trachea

Cancer	concept_id	concept_name
	45768860	Primary mucinous cystadenocarcinoma of trachea
	45768861	Primary solid carcinoma of trachea
	45768862	Primary undifferentiated carcinoma of trachea
	45768863	Primary acinar cell carcinoma of trachea
	45768865	Primary squamous cell carcinoma of trachea
	45768866	Primary clear cell squamous cell carcinoma of trachea
	45768867	Primary papillary squamous cell carcinoma of trachea
	45768868	Primary giant cell carcinoma of trachea
	45768869	Primary adenosquamous carcinoma of trachea
	45768870	Primary spindle cell carcinoma of trachea
	45768952	Primary lymphoepithelial carcinoma of trachea
	45768953	Primary squamous cell adenoid carcinoma of trachea
	45772932	Primary basaloid squamous cell carcinoma of trachea
	45772942	Primary verrucous carcinoma of trachea
	4177112	Malignant tumor of trachea
	4241676	T4: Lung tumor of any size that invades any of the following: mediastinum; heart; great vessels; trachea; esophagus; vertebral body; carina
	258369	Primary malignant neoplasm of lung
	443388	Malignant tumor of lung
	619299	Primary malignant neoplasm of left lung
	619300	Primary malignant neoplasm of right lung
	4293156	pT2: Tumor of lung as per AJCC 6th Edition definition (lung)
	4298502	pT4: Tumor of lung as per AJCC 6th Edition definition (lung)
	45769098	Primary malignant epithelial neoplasm of trachea
	443399	Malignant tumour of bronchus
	4208307	Nonsquamous nonsmall cell neoplasm of lung
	4197581	Squamous cell carcinoma of bronchus in left upper lobe
	4196725	Squamous cell carcinoma of bronchus in right lower lobe
	4197582	Squamous cell carcinoma of bronchus in right middle lobe
	4197583	Squamous cell carcinoma of bronchus in right upper lobe
	4196724	Squamous cell carcinoma of bronchus in left lower lobe
	37395650	Primary small cell malignant neoplasm of lung, TNM stage 3
	37395648	Primary small cell malignant neoplasm of lung, TNM stage 1
	37395649	Primary small cell malignant neoplasm of lung, TNM stage 2
	37395651	Primary small cell malignant neoplasm of lung, TNM stage 4
	4094874	Malignant neoplasm of cartilage of trachea
	4247821	Primary malignant neoplasm of carina
	4092215	Malignant neoplasm of carina of bronchus
	4095449	Malignant neoplasm of mucosa of trachea
	432262	Primary malignant neoplasm of trachea
	433973	Primary malignant neoplasm of bronchus of left lower lobe

Cancer	concept_id	concept_name
	4246121	Primary malignant neoplasm of hilus of lung
	4246804	Primary malignant neoplasm of bronchus of right middle lobe
	4247832	Primary malignant neoplasm of lower lobe of left lung
	45769035	Primary carcinosarcoma of lung
	45769097	Pleuropulmonary blastoma type III
	4312567	Primary malignant neoplasm of upper lobe of right lung
	45769095	Pleuropulmonary blastoma type I
	36716500	Primary malignant neuroendocrine neoplasm of lung
	4157454	Primary malignant neoplasm of lower lobe, bronchus or lung
	619301	Bilateral primary malignant neoplasm of lungs
	4246027	Primary malignant neoplasm of bronchus of left upper lobe
	4246148	Primary malignant neoplasm of right lower lobe of lung
	45769096	Pleuropulmonary blastoma type II
	45769094	Pleuropulmonary blastoma
	4110706	Pancoast tumor
	4311501	Primary malignant neoplasm of right middle lobe of lung
	40391740	Pulmonary blastoma
	4247727	Primary malignant neoplasm of bronchus of right lower lobe
	258375	Overlapping malignant neoplasm of bronchus and lung
	4246805	Primary malignant neoplasm of bronchus of right upper lobe
	4246126	Primary malignant neoplasm of left upper lobe of lung
	4111807	Epithelioid hemangioendothelioma of lung
	4151250	Malignant neoplasm of upper lobe, bronchus or lung
	4089756	Malignant neoplasm of lower lobe of lung
	37110034	Malignant neoplasm of right upper lobe of lung
	37110032	Malignant neoplasm of upper lobe of left lung
	37110033	Malignant neoplasm of lower lobe of left lung
	4092216	Malignant neoplasm of upper lobe of lung
	765056	Malignant carcinoid tumor of lung
	4110587	Malignant tumor of lung parenchyma
	37110031	Malignant neoplasm of lower lobe of right lung
	4092218	Malignant neoplasm of middle lobe bronchus
	4089754	Malignant neoplasm of middle lobe of lung
	602698	Primary squamous cell carcinoma of left main bronchus
	37111619	Malignant carcinoid tumor of bronchus
	257503	Primary malignant neoplasm of main bronchus
	4089755	Malignant neoplasm of lower lobe bronchus
	4162250	Carcinoma of main bronchus
	37116436	Malignant epithelial neoplasm of bronchus

Cancer	concept_id	concept_name
	602160	Primary small cell carcinoma of left main bronchus
	602168	Primary squamous cell carcinoma of right main bronchus
	605477	Primary small cell carcinoma of right main bronchus
	4157333	Malignant neoplasm of main bronchus
	442139	Primary malignant neoplasm of bronchus
	36716499	Primary malignant neuroendocrine neoplasm of bronchus
	4111805	Squamous cell carcinoma of bronchus
	601155	Primary malignant neoplasm of right main bronchus
	4095450	Malignant neoplasm of upper lobe bronchus
	605430	Primary adenocarcinoma of left main bronchus
	601154	Primary malignant neoplasm of left main bronchus
	602149	Primary adenocarcinoma of right main bronchus
NSCLC	36684857	Metastatic non-small cell lung cancer
	4308479	Non-small cell carcinoma of lung, TNM stage 4
	4110705	Squamous cell carcinoma of lung
	4115276	Non-small cell lung cancer
	37109576	Squamous non-small cell lung cancer
	4110589	Large cell carcinoma of lung
	4112738	Adenocarcinoma of lung
	44501471	Non-small cell carcinoma of middle lobe, lung
	44500188	Non-small cell carcinoma of lower lobe, lung
	44499422	Non-small cell carcinoma of upper lobe, lung
All other Lung cancers (Lung cancer excluding NSCLC codes)		
Colorectal; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788623/conceptsets/	4180780	Malignant tumor of anal canal
	4198567	HNPCC - hereditary nonpolyposis colon cancer
	79740	Overlapping malignant neoplasm of colon
	197500	Primary malignant neoplasm of colon
	432257	Primary malignant neoplasm of transverse colon
	436635	Primary malignant neoplasm of sigmoid colon
	437798	Primary malignant neoplasm of splenic flexure of colon
	438979	Primary malignant neoplasm of hepatic flexure of colon
	441800	Primary malignant neoplasm of descending colon
	603310	Primary malignant gastrointestinal stromal neoplasm of colon
	761001	Primary malignant neuroendocrine neoplasm of ascending colon
	4089661	Malignant neoplasm, overlapping lesion of colon
	4092078	Malignant neoplasm of mesocolon
	4247360	Primary malignant neoplasm of the mesocolon
	4247719	Primary malignant neoplasm of ascending colon

Cancer	concept_id	concept_name
	36683531	Malignant neoplasm of colon and/or rectum
	36715913	Primary malignant neuroendocrine neoplasm of colon
	37018659	Overlapping malignant neoplasm of colon and rectum
	435754	Malignant tumor of ascending colon
	443381	Malignant tumor of sigmoid colon
	443382	Malignant tumor of descending colon
	443384	Malignant tumor of transverse colon
	760957	Malignant carcinoid tumor of descending colon
	760958	Malignant carcinoid tumor of ascending colon
	4180790	Malignant tumor of colon
	37111620	Malignant carcinoid tumor of colon
	609194	Primary squamous cell carcinoma of skin of anus
	3655878	Neuroendocrine carcinoma of anus
	4309574	Adenocarcinoma of anus
	36716972	Primary squamous cell carcinoma of anus
	37016237	Primary adenocarcinoma of anus
	40487139	Carcinoma of skin of anus
	40489405	Carcinoma of anus
	4115028	Carcinoma of sigmoid colon
	4149847	Carcinoma of colon
	4193165	Carcinoma of descending colon
	4193871	Carcinoma of transverse colon
	4200514	Adenocarcinoma of sigmoid colon
	4207182	Carcinoma of ascending colon
	4307687	Carcinoma of colon, stage II
	4310858	Carcinoma of colon, stage III
	4312001	Carcinoma of colon, stage IV
	4312240	Carcinoma of colon, stage I
	35624316	Squamous cell carcinoma of colon
	36713361	Primary adenocarcinoma of ascending colon
	36715911	Primary adenocarcinoma of ascending colon and right flexure
	36715912	Primary adenocarcinoma of transverse colon
	36717181	Primary neuroendocrine carcinoma of colon
	36717495	Primary adenocarcinoma of descending colon and splenic flexure
	37208245	Primary adenocarcinoma of descending colon
	42872396	Primary adenocarcinoma of colon
	4110575	Adenocarcinoma of rectum
	35624314	Squamous cell carcinoma of rectum
	36717182	Primary neuroendocrine carcinoma of rectum
	40492939	Carcinoma of upper rectum

Cancer	concept_id	concept_name
	4180791	Malignant tumor of hepatic flexure
	4180792	Malignant tumor of rectosigmoid junction
	4181344	Malignant tumor of splenic flexure
	443390	Malignant tumor of rectum
	40481902	Malignant neoplasm of anorectum
	40487050	Anorectal adenocarcinoma
	436348	Primary malignant neoplasm of anal canal
	4116241	Squamous cell carcinoma of anal margin
	35622690	Adenocarcinoma of anal canal
	4112143	Carcinoma of anal canal
	36715921	Primary cloacogenic carcinoma of anal canal
	36716511	Primary squamous cell carcinoma of anal canal
	37116449	Primary malignant neuroendocrine neoplasm of anal canal
	438699	Primary malignant neoplasm of rectosigmoid junction
	4151260	Carcinoma of the rectosigmoid junction
	37016239	Primary adenocarcinoma of rectosigmoid junction
	4095430	Malignant neoplasm of rectum, rectosigmoid junction and anus
	608050	Leiomyosarcoma of colon
	36715914	Primary malignant neuroendocrine neoplasm of rectum
	4112730	Malignant tumor of anorectal junction
	4322376	Adenocarcinoma of rectosigmoid junction
	603311	Primary malignant gastrointestinal stromal neoplasm of rectum
	37018934	Malignant carcinoid tumor of rectum
	74582	Primary malignant neoplasm of rectum
	438090	Overlapping malignant neoplasm of rectum, anus and anal canal
	608046	Leiomyosarcoma of rectum
	37116448	Primary malignant neuroendocrine neoplasm of anus
	80045	Primary malignant neoplasm of anus
	4180911	Malignant tumor of anus
	608069	Malignant mesenchymal neoplasm of anus
Prostate; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788649/conceptsets/	4163261	Malignant tumor of prostate
	37311236	Infiltrating duct carcinoma of prostate
	36716186	Hormone sensitive prostate cancer
	4116087	Carcinoma of prostate
	37395835	Familial prostate cancer
	4164017	Squamous cell carcinoma of prostate
	4082919	Endometrioid carcinoma of prostate
	37311683	Acinar cell cystadenocarcinoma of prostate

Cancer	concept_id	concept_name
	4288534	Small cell carcinoma of prostate
	200962	Primary malignant neoplasm of prostate
	4141960	Hormone refractory prostate cancer
	4161028	Adenocarcinoma of prostate
Pancreas; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788647/conceptsets/	602008	Primary adenocarcinoma of neck of pancreas
	602011	Primary adenocarcinoma of tail of pancreas
	605820	Adenocarcinoma of pancreas with NRG1 fusion
	606747	Infiltrating duct carcinoma of pancreas
	3655584	Mucinous cystic neoplasm with invasive carcinoma of pancreas
	3655588	Mixed ductal-neuroendocrine carcinoma of pancreas
	4110585	Carcinoma of endocrine pancreas
	4157459	Carcinoma of pancreas
	4178960	Carcinoma of tail of pancreas
	4181331	Carcinoma of body of pancreas
	4209933	Carcinoma of head of pancreas
	4340498	Cystadenocarcinoma of pancreas
	36674768	Squamous cell carcinoma of exocrine pancreas
	36683250	Invasive intraductal papillary-mucinous carcinoma of pancreas
	36713362	Primary adenocarcinoma of body of pancreas
	36713363	Primary adenocarcinoma of head of pancreas
	37204187	Solid pseudopapillary carcinoma of pancreas
	37204808	Serous cystadenocarcinoma of pancreas
	37204852	Acinar cell carcinoma of pancreas
	37206235	Mucinous cystadenocarcinoma of pancreas
	42872399	Primary adenocarcinoma of pancreas
	45763891	Adenocarcinoma of pancreas
	765387	Malignant carcinoid tumour of pancreas
	4092072	Malignant tumour of body of pancreas
	4095436	Malignant tumour of tail of pancreas
	4112734	Malignant tumour of endocrine pancreas
	4111024	Malignant tumour of exocrine pancreas
	4180793	Malignant tumour of pancreas
	4178967	Malignant tumour of head of pancreas
	37311469	Pancreatic ductal adenocarcinoma
	192261	Overlapping malignant neoplasm of pancreas
	609180	Primary malignant gastrointestinal stromal neoplasm of pancreas
	37017016	Malignant insulinoma
	602009	Primary adenocarcinoma of pancreatic duct

Cancer	concept_id	concept_name
	36715928	Primary adenocarcinoma of ampulla of Vater
	434293	Primary malignant neoplasm of body of pancreas
	432843	Primary malignant neoplasm of tail of pancreas
	4095437	Malignant tumour of Islets of Langerhans
	25486	Primary malignant neoplasm of islets of Langerhans
	4094866	Malignant tumor of pancreatic duct
	433423	Primary malignant neoplasm of pancreatic duct
	199754	Primary malignant neoplasm of pancreas
	601133	Primary malignant neoplasm of neck of pancreas
	40391739	Pancreatoblastoma
	37395837	Familial malignant neoplasm of pancreas
	440649	Primary malignant neoplasm of head of pancreas
	4092074	Malignant neoplasm of ectopic pancreatic tissue
	42536743	Primary malignant neuroendocrine neoplasm of pancreas
Ovary; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788646/conceptsets/	4307838	Ovarian cancer, disseminated
	36674976	Hereditary site-specific ovarian cancer syndrome
	37397555	Hereditary breast and ovarian cancer syndrome
	4112860	Brenner tumor of ovary
	4300688	Yolk sac tumor
	4110865	Theca cell tumor of ovary
	4112079	Sertoli-Leydig cell tumor of ovary
	4112864	Malignant germ cell tumor of ovary
	4116073	Malignant epithelial tumor of ovary
	4116074	Malignant sex cord tumor of ovary
	4181351	Malignant tumor of ovary
	4240443	T1c (IC): Tumor limited to one or both ovaries
	4265308	pT1c (IC): Tumor limited to one or both ovaries
	4110863	Undifferentiated carcinoma of ovary
	4310444	Carcinoma of ovary, stage 4
	4311462	Carcinoma of ovary, stage 2
	4311576	Carcinoma of ovary, stage 1
	4313202	Carcinoma of ovary, stage 3
	4112862	Granulosa cell tumor of ovary
	4116077	Dysgerminoma of ovary
	36686093	Malignant germ cell neoplasm of right ovary
	36686094	Malignant germ cell neoplasm of left ovary
	37396882	Theca steroid producing cell malignant neoplasm of ovary
	4116076	Embryonal carcinoma of ovary
	4112865	Choriocarcinoma of ovary
	4110870	Endodermal sinus tumor of ovary

Cancer	concept_id	concept_name
	37396690	Primary non-gestational choriocarcinoma of ovary
	37311080	Malignant immature teratoma of ovary
	37116598	Primary mucinous adenocarcinoma of ovary
	4112857	Endometrioid carcinoma ovary
	602307	Left ovarian primary mucinous cystadenocarcinoma
	602311	Left ovarian primary endometrioid carcinoma
	602310	Right ovarian primary endometrioid carcinoma
	4230544	Krukenberg tumor
	4112856	Mucinous cystadenocarcinoma of ovary
	4198275	Cystadenocarcinoma of ovary
	36674767	Small cell carcinoma of ovary
	35621826	Clear cell adenocarcinoma of ovary
	602306	Right ovarian primary mucinous cystadenocarcinoma
	36716618	Primary high grade serous adenocarcinoma of ovary
	200051	Primary malignant neoplasm of ovary
	36687125	Carcinosarcoma of bilateral ovaries
	4283749	Malignant tumor involving right ovary by direct extension from left ovary
	4289682	Malignant tumor involving left ovary by direct extension from right ovary
	4112855	Serous papillary cystadenocarcinoma ovary
	4289681	Primary malignant neoplasm of left ovary
	36687124	Carcinosarcoma of right ovary
	4289392	Primary malignant neoplasm of right ovary
	45765433	Carcinosarcoma of ovary
	36717228	Primary low grade serous adenocarcinoma of ovary
	4307986	Sarcoma of ovary
	602308	Primary serous papillary cystadenocarcinoma of right ovary
	36687123	Carcinosarcoma of left ovary
	602309	Primary serous papillary cystadenocarcinoma of left ovary
	608868	Right ovarian primary sarcoma
	608867	Left ovarian primary sarcoma
Leukemia; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788635/conceptsets/	132853	Lymphoid leukemia
	134305	Acute lymphoid leukemia
	134603	Chronic myeloid leukemia
	135499	Subacute myeloid leukemia
	135768	Acute monocytic leukemia
	136056	Chronic monocytic leukemia
	136656	Subacute lymphoid leukemia
	136930	Megakaryocytic leukemia in remission
	138099	Erythroleukemia, FAB M6

Cancer	concept_id	concept_name
	138379	Chronic lymphoid leukemia, disease
	138708	Acute leukemia
	140057	Chronic leukemia
	140352	Acute myeloid leukemia, disease
	140666	Myeloid leukemia
	313159	Megakaryocytic leukemia
	313430	Subacute monocytic leukemia
	315497	Subacute leukemia
	317510	Leukemia
	321526	Monocytic leukemia
	600661	Mixed phenotype acute leukemia with T-cell and myeloid lineage
	600663	Mixed phenotype acute leukemia with myeloid and B-cell lymphoid phenotypes
	606958	Acute myeloid leukemia with t(8;21)(q22;q22) RUNX1-RUNX1T1
	607405	Mixed phenotype acute leukemia with t(9;22) (q34;q11.2); BCR-ABL1
	608277	Acute myeloid leukemia with 11q23 abnormality
	3572256	Refractory acute myeloid leukemia
	3654662	Acute myeloid leukemia with inv(16)(p13.1q22) or t(16;16)(p13.1;q22) CBFB-MYH11
	3654647	B lymphoblastic leukemia lymphoma with t(5;14)(q31;q32); IL3-IGH
	3654648	B lymphoblastic leukemia lymphoma with t(v;11q23); MLL rearranged
	3654649	B lymphoblastic leukemia lymphoma with t(12;21) (p13;q22); TEL/AML1 (ETV6-RUNX1)
	3654650	B lymphoblastic leukemia lymphoma with t(1;19)(Q23;P13.3); E2A-PBX1 (TCF3/PBX1)
	3654651	B lymphoblastic leukemia lymphoma with hypodiploidy
	3654653	B lymphoblastic leukemia lymphoma with hyperdiploidy
	4001331	Prolymphocytic leukemia (clinical)
	4002496	Mast cell leukemia (clinical)
	4002497	Acute promyelocytic leukemia, FAB M3
	4003187	Acute myelomonocytic leukemia, FAB M4
	4003188	Adult T-cell leukemia/lymphoma
	4038845	Hairy cell leukemia (clinical)
	4079281	Null cell acute lymphoblastic leukemia
	4079280	Common acute lymphoblastic leukemia
	4079282	Atypical chronic myeloid leukemia
	4079683	T-cell prolymphocytic leukemia
	4079686	Acute megakaryoblastic leukemia
	4079690	Prethymic and thymic T-cell lymphoma/leukemia

Cancer	concept_id	concept_name
	4081867	Acute biphenotypic leukemia
	4082311	B-cell chronic lymphocytic leukemia
	4082459	Hairy cell leukemia variant
	4082460	Large granular lymphocytic leukemia
	4082461	Precursor B-cell acute lymphoblastic leukemia
	4082462	T-cell acute lymphoblastic leukemia
	4082338	Chronic lymphocytic prolymphocytic leukemia syndrome
	4082481	Juvenile chronic myeloid leukemia
	4082485	Acute monoblastic leukemia
	4091925	Chronic eosinophilic leukemia
	4094549	Aleukemic lymphoid leukemia
	4094550	Adult T-cell leukemia
	4094553	Aleukemic monocytic leukemia
	4095897	Compound leukemias
	4097581	Chronic neutrophilic leukemia
	4097585	Aleukemic myeloid leukemia
	4097706	Myelomonocytic leukemia
	4097707	Subacute myelomonocytic leukemia
	4112803	Acute promyelocytic leukemia - hypogranular variant
	4116880	Acute myelomonocytic leukemia - eosinophilic variant
	4121332	Aleukemic leukemia
	4133599	Chronic myelomonocytic leukemia
	4138008	Philadelphia chromosome-positive acute lymphoblastic leukemia
	4139554	Atypical hairy cell leukemia
	4144191	Basophilic leukemia
	4146022	Neutrophilic leukemia
	4153344	Acute lymphoblastic leukemia, transitional pre-B-cell
	4173824	B-cell chronic lymphocytic leukemia variant
	4173955	T-cell chronic lymphocytic leukemia
	4173963	B-cell acute lymphoblastic leukemia
	4173970	Acute eosinophilic leukemia
	4173974	B-cell prolymphocytic leukemia
	4175688	Hypergranular promyelocytic leukemia
	4185301	Chronic myeloid leukemia in myeloid blast crisis
	4187635	Accelerated phase chronic myeloid leukemia
	4188973	Chronic myeloid leukemia in lymphoid blast crisis
	4188975	Chronic phase chronic myeloid leukemia
	4189938	Acute monocytic/monoblastic leukemia
	4211481	Blastic phase chronic myeloid leukemia
	4221907	Precursor T cell lymphoblastic leukemia/lymphoblastic lymphoma

Cancer	concept_id	concept_name
	4230253	Acute myeloid leukemia without maturation, FAB M1
	4233531	Acute myeloid leukemia, minimal differentiation, FAB M0
	4234749	Acute myeloid leukemia with maturation, FAB M2
	4245460	Hairy cell leukemia of spleen
	4297353	Leukemic infiltration of skin (chronic T-cell lymphocytic leukemia)
	4297355	Aggressive NK-cell leukemia involving skin
	4299143	Leukemic infiltration of skin (T-cell lymphoblastic leukemia)
	4299273	Mast cell leukemia affecting skin
	4299151	Leukemic infiltration of skin in hairy-cell leukemia
	4299153	Leukemic infiltration of skin in myeloid leukemia
	4299154	Leukemic infiltration of skin in chronic myeloid leukemia
	4299155	Leukemic infiltration of skin in monocytic leukemia
	4300784	Leukemic infiltration of skin in acute myeloid leukemia
	4301665	Leukemic infiltration of skin (T-cell prolymphocytic leukemia)
	4326339	Smoldering chronic lymphocytic leukemia
	35607963	Megakaryoblastic acute myeloid leukemia with t(1;22)(p13;q13)
	35622003	Acute myeloid leukemia with NPM1 somatic mutation
	35622696	Acute myeloid leukemia with CEBPA somatic mutations
	35622760	Inherited acute myeloid leukemia
	35623630	Acute myeloid leukemia and myelodysplastic syndrome related to alkylating agent
	35623631	Acute myeloid leukemia and myelodysplastic syndrome related to topoisomerase type 2 inhibitor
	35623633	Acute myeloid leukemia and myelodysplastic syndrome related to radiation
	36674687	Aleukemic mast cell leukemia
	36676614	Differentiation syndrome due to and following chemotherapy co-occurrent with acute promyelocytic leukemia
	36683269	Acute myeloid leukemia with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1
	36715966	Adult T-cell leukemia/lymphoma of skin
	36715587	Acute myeloid leukemia due to recurrent genetic abnormality
	36715589	Acute leukemia of ambiguous lineage
	36717231	Meningeal leukemia
	36717461	Therapy related acute myeloid leukemia and myelodysplastic syndrome
	36717161	Aggressive natural killer-cell leukemia
	36712835	Refractory acute lymphoid leukemia
	37017893	Philadelphia chromosome-negative precursor B-cell acute lymphoblastic leukemia

Cancer	concept_id	concept_name
	37018869	Disorder of central nervous system co-occurrent and due to acute lymphoid leukemia
	37109936	B lymphoblastic leukemia lymphoma with t(9:22)(q34;q11.2); BCR-ABL 1
	37110870	Acute myeloid leukemia with t(8;16)(p11;p13) translocation
	37110902	Chronic lymphocytic leukemia genetic mutation variant
	37116722	Acute myeloid leukemia with t(6;9)(p23;q34) translocation
	37119145	Myeloid leukemia co-occurrent with Down syndrome
	37204530	Non-chronic lymphocytic leukemia monoclonal B-cell lymphocytosis
	37204479	Noonan syndrome-like disorder with juvenile myelomonocytic leukemia
	37206728	Monoclonal B-cell lymphocytosis chronic lymphocytic leukemia-type
	40481524	Acute myeloid leukemia with t(9:11)(p22;q23); MLLT3-MLL
	40482847	Juvenile myelomonocytic leukemia
	40483761	Acute myeloid leukemia with myelodysplasia-related changes
	40486741	Philadelphia chromosome negative chronic myelogenous leukemia
	40493442	Philadelphia chromosome positive chronic myelogenous leukemia
	42538579	Therapy related acute myeloid leukemia due to and following administration of antineoplastic agent
	42539431	Acute myeloid leukemia with FMS-like tyrosine kinase-3 mutation
	44783718	T-cell large granular lymphocytic leukemia
	45766268	Cytogenetically normal acute myeloid leukemia
	45767656	Gingivitis due to leukemia
	45765495	Core binding factor acute myeloid leukemia
	46271363	Periodontitis co-occurrent with leukemia
	4301780	Leukemic infiltration of skin
	44784140	Emberger syndrome
	132570	Leukemic reticuloendotheliosis of lymph nodes of head, face and neck
	193429	Leukemic reticuloendotheliosis of intra-abdominal lymph nodes
	44814026	Clinical stage C chronic lymphocytic leukaemia
	132852	Leukemic reticuloendotheliosis of extranodal AND/OR solid organ site
	44811228	Clinical stage B chronic lymphocytic leukaemia
	439269	Leukemic reticuloendotheliosis of lymph nodes of axilla and upper limb
	318989	Leukemic reticuloendotheliosis of lymph nodes of multiple sites
	442095	Leukemic reticuloendotheliosis of intrathoracic lymph nodes

Cancer	concept_id	concept_name
	439268	Leukemic reticuloendotheliosis of lymph nodes of inguinal region and lower limb
	196650	Leukemic reticuloendotheliosis of intrapelvic lymph nodes
	44811227	Clinical stage A chronic lymphocytic leukaemia
	44807009	Acute myeloid leukaemia with 11q23 abnormality
	4097582	Chloroma
	4227963	Precursor T-cell lymphoblastic lymphoma
	4298848	T-cell leukemic infiltration of skin
Acute Myeloid Leukemia	140352	Acute myeloid leukemia, disease
	44807009	Acute myeloid leukaemia with 11q23 abnormality
	135499	Subacute myeloid leukemia
	4233531	Acute myeloid leukemia, minimal differentiation, FAB M0
	4230253	Acute myeloid leukemia without maturation, FAB M1
	4234749	Acute myeloid leukemia with maturation, FAB M2
	4300784	Leukemic infiltration of skin in acute myeloid leukemia
	4138903	Acute myeloid leukemia with maturation, FAB M2, in remission
	40481524	Acute myeloid leukemia with t(9:11)(p22;q23); MLLT3-MLL
	40483761	Acute myeloid leukemia with myelodysplasia-related changes
	45765495	Core binding factor acute myeloid leukemia
	45766268	Cytogenetically normal acute myeloid leukemia
	36715587	Acute myeloid leukemia due to recurrent genetic abnormality
	36717461	Therapy related acute myeloid leukemia and myelodysplastic syndrome
	37110870	Acute myeloid leukemia with t(8;16)(p11;p13) translocation
	37116722	Acute myeloid leukemia with t(6;9)(p23;q34) translocation
	42539431	Acute myeloid leukemia with FMS-like tyrosine kinase-3 mutation
	35622003	Acute myeloid leukemia with NPM1 somatic mutation
	35607963	Megakaryoblastic acute myeloid leukemia with t(1;22)(p13;q13)
	35622696	Acute myeloid leukemia with CEBPA somatic mutations
	35622760	Inherited acute myeloid leukemia
	36683269	Acute myeloid leukemia with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1
	3654662	Acute myeloid leukemia with inv(16)(p13.1q22) or t(16;16)(p13.1;q22) CBFB-MYH11
	4002497	Acute promyelocytic leukemia, FAB M3
	4112803	Acute promyelocytic leukemia - hypogranular variant
	4137687	Acute promyelocytic leukemia, FAB M3, in remission
	4175688	Hypergranular promyelocytic leukemia
	4116880	Acute myelomonocytic leukemia - eosinophilic variant

Cancer	concept_id	concept_name
	4079686	Acute megakaryoblastic leukemia
	138099	Erythroleukemia, FAB M6
	4144191	Basophilic leukemia
	4003187	Acute myelomonocytic leukemia, FAB M4
	135768	Acute monocytic leukemia
	4003184	Acute panmyelosis with myelofibrosis
	4082485	Acute monoblastic leukemia
	4189938	Acute monocytic/monoblastic leukemia
	4173970	Acute eosinophilic leukemia
	4304355	Acute myeloid leukemia with abnormal marrow eosinophils
	4304199	Acute myeloid leukemia, minimal differentiation
	4304356	Acute myeloid leukemia without maturation
	4304051	Acute myeloid leukemia with maturation
	4029177	Acute myeloid leukemia with myelodysplasia-related changes
	4028713	Acute myeloid leukemia, t(8;21) (q22;q22)
	4029663	Acute myeloid leukemia, 11q23 abnormalities
	4030263	Therapy-related acute myeloid leukemia and myelodysplastic syndrome
	4031360	Acute myeloid leukemia, M6 type
	4073533	Acute myeloid leukemia, no ICD-O subtype
	4265011	Acute myeloid leukemia with recurrent genetic abnormality
	4264447	Acute myeloid leukemia with multilineage dysplasia following a myelodysplastic syndrome or myelodysplastic syndrome/myeloproliferative disorder
	4288090	Acute myeloid leukemia with multilineage dysplasia without antecedent myelodysplastic syndrome
	4184848	Acute myeloid leukemia
	42872921	Mixed phenotype acute leukemia B/myeloid
	42872922	Mixed phenotype acute leukemia T/myeloid
	42872933	Acute myeloid leukemia with t(6;9)(p23;q34); DEK-NUP214
	42872934	Acute myeloid leukemia with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); RPN1-EVI1
	42872942	Acute myeloid leukemia (megakaryoblastic) with t(1;22)(p13;q13); RBM15-MKL1
	45771384	Acute myeloid leukemia with mutation of CEBPA (CCAAT enhancer binding protein alpha) gene
	45766616	Acute myeloid leukemia with mutated NPM1
	37110871	Acute myeloid leukemia with t(8;16)(p11;p13) translocation
	42535969	Acute myeloid leukemia with FMS-like tyrosine kinase-3 mutation
	37204375	Acute myeloid leukemia with BCR-ABL1
	37204557	Acute myeloid leukemia with mutated RUNX1
	37312067	Acute myeloid leukemia with biallelic mutation of CEBPA (CCAAT enhancer binding protein alpha) gene

Cancer	concept_id	concept_name
Chronic Myeloid Leukemia	134603	Chronic myelocytic leukaemia
	4187635	Accelerated phase chronic myeloid leukemia
	4079282	Atypical chronic myeloid leukemia
	4211481	Blastic phase chronic myeloid leukemia
	36402753	Chronic myelogenous leukemia, BCR/ABL positive of blood
	44501329	Chronic myelogenous leukemia, BCR/ABL positive of bone marrow
	36563942	Chronic myelogenous leukemia, BCR/ABL positive of hematopoietic system, NOS
	36547897	Chronic myelogenous leukemia, BCR/ABL positive of reticuloendothelial system, NOS
	134603	Chronic myeloid leukemia
	4188973	Chronic myeloid leukemia in lymphoid blast crisis
	4185301	Chronic myeloid leukemia in myeloid blast crisis
	4097581	Chronic neutrophilic leukemia
	4188975	Chronic phase chronic myeloid leukemia
	4082481	Juvenile chronic myeloid leukemia
	40482847	Juvenile myelomonocytic leukemia
	37311926	Juvenile myelomonocytic leukemia in remission
	4299154	Leukemic infiltration of skin in chronic myeloid leukemia
	40486741	Philadelphia chromosome negative chronic myelogenous leukemia
	40493442	Philadelphia chromosome positive chronic myelogenous leukemia
Acute lymphoblastic leukemia	134305	Acute lymphoid leukaemia
	136656	Subacute lymphoid leukemia
	37018869	Disorder of central nervous system co-occurrent and due to acute lymphoid leukemia
	4173963	B-cell acute lymphoblastic leukemia
	4082461	Precursor B-cell acute lymphoblastic leukemia
	4079280	Common acute lymphoblastic leukemia
	4079281	Null cell acute lymphoblastic leukemia
	4082462	T-cell acute lymphoblastic leukemia
	4153344	Acute lymphoblastic leukemia, transitional pre-B-cell
	4299143	Leukemic infiltration of skin (T-cell lymphoblastic leukemia)
	4221907	Precursor T cell lymphoblastic leukemia/lymphoblastic lymphoma
	4138008	Philadelphia chromosome-positive acute lymphoblastic leukemia
	37017893	Philadelphia chromosome-negative precursor B-cell acute lymphoblastic leukemia
	37109936	B lymphoblastic leukemia lymphoma with t(9:22)(q34;q11.2); BCR-ABL 1
	3654647	B lymphoblastic leukemia lymphoma with t(5;14)(q31;q32); IL3-IGH

Cancer	concept_id	concept_name
	3654648	B lymphoblastic leukemia lymphoma with t(v;11q23); MLL rearranged
	3654649	B lymphoblastic leukemia lymphoma with t(12;21)(p13;q22); TEL/AML1 (ETV6-RUNX1)
	3654650	B lymphoblastic leukemia lymphoma with t(1;19)(Q23;P13.3); E2A-PBX1 (TCF3/PBX1)
	3654651	B lymphoblastic leukemia lymphoma with hypodiploidy
	3654653	B lymphoblastic leukemia lymphoma with hyperdiploidy
	4003188	Adult T-cell leukemia/lymphoma
	4081867	Acute biphenotypic leukemia
	4227963	Precursor T-cell lymphoblastic lymphoma
	4030260	Precursor cell lymphoblastic leukemia
	4029662	Precursor B-cell lymphoblastic leukemia
	4030261	Precursor T-cell lymphoblastic leukemia
	4264448	Precursor B-lymphoblastic leukemia/lymphoblastic lymphoma
	4288091	Precursor T cell lymphoblastic leukemia/lymphoblastic lymphoma
	4189936	Acute lymphoblastic leukemia - category
	4143821	Philadelphia chromosome-positive acute lymphoblastic leukemia
	42872925	B lymphoblastic leukemia / lymphoma - category
	42872954	B lymphoblastic leukemia lymphoma, no ICD-O subtype
	42872955	B lymphoblastic leukemia lymphoma with t(9;22)(q34;q11.2); BCR-ABL1
	42872956	B lymphoblastic leukemia lymphoma with t(v;11q23); MLL rearranged
	42872957	B lymphoblastic leukemia lymphoma with t(12;21)(p13;q22); TEL-AML1 (ETV6-RUNX1)
	42872958	B lymphoblastic leukemia lymphoma with hyperdiploidy
	42872959	B lymphoblastic leukemia lymphoma with hypodiploidy (Hypodiploid ALL)
	42872960	B lymphoblastic leukemia lymphoma with t(5;14)(q31;q32); IL3-IGH
	42872961	B lymphoblastic leukemia lymphoma with t(1;19)(q23;p13.3); E2A-PBX1 (TCF3-PBX1)
	45766617	T lymphoblastic leukemia/lymphoma
	37204662	NK-lymphoblastic leukemia/lymphoma
	37204838	B-lymphoblastic leukemia lymphoma BCR-ABL1-like
	37206196	B lymphoblastic leukemia lymphoma with iAMP21
Chronic lymphoblastic leukemia	138379	Chronic lymphoid leukaemia
	44811228	Clinical stage B chronic lymphocytic leukaemia
	44814026	Clinical stage C chronic lymphocytic leukaemia
	4082311	B-cell chronic lymphocytic leukemia
	4173824	B-cell chronic lymphocytic leukemia variant

Cancer	concept_id	concept_name
	4173955	T-cell chronic lymphocytic leukemia
	4082338	Chronic lymphocytic prolymphocytic leukemia syndrome
	4297353	Leukemic infiltration of skin (chronic T-cell lymphocytic leukemia)
	4326339	Smoldering chronic lymphocytic leukemia
	37110902	Chronic lymphocytic leukemia genetic mutation variant
	37206728	Monoclonal B-cell lymphocytosis chronic lymphocytic leukemia-type
	138379	Chronic lymphoid leukemia, disease
	44783718	T-cell large granular lymphocytic leukemia
	132570	Leukemic reticuloendotheliosis of lymph nodes of head, face and neck
	4082460	Large granular lymphocytic leukemia
	193429	Leukemic reticuloendotheliosis of intra-abdominal lymph nodes
	4173974	B-cell prolymphocytic leukemia
	132852	Leukemic reticuloendotheliosis of extranodal AND/OR solid organ site
	4082459	Hairy cell leukemia variant
	4299151	Leukemic infiltration of skin in hairy-cell leukemia
	4001331	Prolymphocytic leukemia (clinical)
	4245460	Hairy cell leukemia of spleen
	4139554	Atypical hairy cell leukemia
	439269	Leukemic reticuloendotheliosis of lymph nodes of axilla and upper limb
	4038845	Hairy cell leukemia (clinical)
	318989	Leukemic reticuloendotheliosis of lymph nodes of multiple sites
	4079683	T-cell prolymphocytic leukemia
	439268	Leukemic reticuloendotheliosis of lymph nodes of inguinal region and lower limb
	196650	Leukemic reticuloendotheliosis of intrapelvic lymph nodes
	442095	Leukemic reticuloendotheliosis of intrathoracic lymph nodes
	4180093	Chronic lymphocytic leukemia
	37312112	Monoclonal B-cell lymphocytosis chronic lymphocytic leukemia-type
	37312109	Monoclonal B-cell lymphocytosis non-chronic lymphocytic leukemia type
	4186899	Chronic lymphoid leukemia - category
Multiple Myeloma	4224628	Amyloid light chain amyloidosis due to multiple myeloma
	4258135	Asymptomatic multiple myeloma
	4043447	Bone marrow: myeloma cells
	4094548	Extramedullary plasmacytoma
	46270015	History of multiple myeloma

Cancer	concept_id	concept_name
	37209514	Hypogammaglobulinemia due to multiple myeloma
	4111355	IgA myeloma
	4112310	IgD myeloma
	4111356	IgG myeloma
	4259972	Indolent multiple myeloma
	4188299	Kappa light chain myeloma
	4197600	Lambda light chain myeloma
	4082464	Light chain myeloma
	37016161	Light chain nephropathy due to multiple myeloma
	437233	Multiple myeloma
	4210177	Multiple myeloma
	4214660	Multiple solitary plasmacytomas
	4019477	Myeloma-associated amyloidosis
	4137433	Myeloma kidney
	4043713	Neuropathy due to multiple myeloma
	4079684	Non-secretory myeloma
	42538151	Osteoporosis co-occurrent and due to multiple myeloma
	4137510	Osteosclerotic myeloma
	133154	Plasma cell leukemia
	4028859	Plasma cell leukemia
	4190641	Plasma cell myeloma - category
	4190642	Plasma cell myeloma/plasmacytoma
	4163558	Plasma cell myeloma/plasmacytoma
	4216139	Plasmacytoma
	4024874	Plasmacytoma
	4300702	Primary cutaneous plasmacytoma
	4184985	Smoldering myeloma
	4145040	Solitary osseous myeloma
Breast; https://atlas-demo.ohdsi.org/#/cohortdefinition/1788621/conceptsets/	759932	Infiltrating duct carcinoma of left female breast
	759933	Infiltrating duct carcinoma of right female breast
	761170	Infiltrating duct carcinoma of bilateral female breasts
	4237178	Infiltrating duct carcinoma of breast
	37017351	Invasive carcinoma of breast
	3179883	Invasive carcinoma of breast without extensive intraductal component
	3184724	Invasive carcinoma of breast with extensive intraductal component
	4116071	Carcinoma of breast (disorder)
	36712719	Infiltrating ductal carcinoma of upper inner quadrant of left female breast

Cancer	concept_id	concept_name
	36712720	Infiltrating ductal carcinoma of upper outer quadrant of left female breast
	36712721	Infiltrating ductal carcinoma of central portion of right female breast
	36712722	Infiltrating ductal carcinoma of upper inner quadrant of right female breast
	36712723	Infiltrating ductal carcinoma of upper outer quadrant of right female breast
	36712724	Infiltrating lobular carcinoma of left female breast
	36712725	Infiltrating lobular carcinoma of right female breast
	36717260	Infiltrating ductal carcinoma of central portion of left female breast
	37208322	Infiltrating ductal carcinoma of axillary tail of left female breast
	37208324	Infiltrating ductal carcinoma of lower inner quadrant of left female breast
	37208325	Infiltrating ductal carcinoma of lower outer quadrant of left female breast
	37208326	Infiltrating ductal carcinoma of axillary tail of right female breast
	37208328	Infiltrating ductal carcinoma of lower inner quadrant of right female breast
	37208329	Infiltrating ductal carcinoma of lower outer quadrant of right female breast
	40486563	Carcinoma of female breast
	40492507	Infiltrating duct carcinoma of female breast
	37018660	Primary malignant inflammatory neoplasm of female breast
Endometrium	4048226	Adenocarcinoma of endometrium
	40491873	Sarcoma of endometrium
	4247238	Primary malignant neoplasm of endometrium
	37115735	Endometrial carcinosarcoma
	37116597	Primary undifferentiated carcinoma of endometrium
	37116596	Primary squamous cell carcinoma of endometrium
	36716616	Primary small cell carcinoma of endometrium
	4110871	Endometrial carcinoma
Hodgkin Lymphoma	4038835	Hodgkin's disease (clinical)
	42538850	Classical Hodgkin lymphoma
Non-Hodgkin Lymphoma (excludes Hodgkin concepts)	147411	Follicular non-Hodgkin's lymphoma
	4038838	Non-Hodgkin's lymphoma (clinical)
	432571	Malignant lymphoma
	35624275	Primary bone lymphoma
	373152	Primary central nervous system lymphoma
	4002358	Lymphoepithelioid lymphoma (clinical)

Concept IDs include descendants unless highlighted as being excluded. By OMOP standards descendants automatically include the ancestor.
Before finalising the concept sets, CohortDiagnostics will run on cohorts created using the initial concept sets to check code counts and patient characteristics which might give indications to adjust the concept sets.

Table 2. Preliminary code list for conditions for the calculation of the approximate score of frailty

Condition	Concept ID	Concept Name
Mobility and transfer problems	4053076	Mobility poor
	4306934	Impaired mobility
	4052049	Mobility fair
	4310235	Reduced mobility
	4031883	Impaired bed mobility
	4032531	Impaired wheelchair mobility
	4052047	Mobility very poor
	4119464	Does not transfer between wheelchair and toilet
	4118805	Unable to transfer between wheelchair and toilet
	4199114	Difficulty mobilizing using mobility aids
	4199113	Does not mobilize using mobility aids
	4199111	Unable to mobilize using mobility aids
	4023190	Wheelchair bound
	4136754	Dependent on helper pushing wheelchair
	4200353	Able to mobilize using mobility aids
	4199115	Able to mobilize using wheelchair
	4199721	Able to move around supporting self on furniture
	44790310	Able to walk short distances
	45878557	Completely immobile
	36716239	Dependent for sitting
	36716240	Dependent for standing
	4146424	Dependent for walking
	46272933	Deterioration in ability to walk
	4200194	Difficulty mobilizing
	4199094	Difficulty mobilizing indoors
	4199114	Difficulty mobilizing using mobility aids
	4199116	Difficulty mobilizing using wheelchair
	4199431	Difficulty moving
	4200817	Difficulty moving around supporting self on furniture
	4107789	Difficulty shuffling
	4107851	Difficulty sitting
	4199552	Difficulty sitting unsupported
	4093668	Difficulty standing
	4154006	Difficulty transferring weight

Condition	Concept ID	Concept Name
	36714126	Difficulty walking
	4086550	Difficulty walking up a slope
	4112788	Difficulty weight-bearing
	4199112	Does mobilize using aids
	4199725	Does mobilize using wheelchair
	4200815	Does move around supporting self on furniture
	4200193	Does not mobilize
	4200798	Does not mobilize indoors
	4200183	Does not move
	4084746	Does not shuffle
	4106333	Does not sit
	4199551	Does not sit unsupported
	4106335	Does not stand
	4154005	Does not transfer weight
	4086871	Does not walk
	4086549	Does not walk up a slope
	4112787	Does not weight-bear
	4295037	Get up and go test - abnormal
	4009877	Immobile
	1621081	Immobile or
	4031883	Impaired bed mobility
	4306934	Impaired mobility
	3198828	Increased weakness when ambulating
	4010359	Loss of control of walking
	1314392	Patient is not ambulatory, bed ridden, immobile, confined to chair, wheelchair bound, dependent on helper pushing wheelchair, independent in wheelchair or minimal help in wheelchair
	1314394	Patient not ambulatory, bed ridden, immobile, confined to chair, wheelchair bound, dependent on helper pushing wheelchair, independent in wheelchair or minimal help in wheelchair
	44790681	Patient unable to get up unaided
	4012646	Stick only for walking
	4012944	Tripod/quadrupod: walking
	4199550	Unable to mobilize
	4199093	Unable to mobilize indoors
	4199111	Unable to mobilize using mobility aids
	4200355	Unable to mobilize using wheelchair

Condition	Concept ID	Concept Name
	4200350	Unable to move around supporting self on furniture
	4105451	Unable to shuffle
	4106332	Unable to sit
	4023187	Unable to sit unsupported
	4060223	Unable to stand
	4151066	Unable to transfer weight
	4086548	Unable to walk
	44792042	Unable to walk long distances
	4086874	Unable to walk up a slope
	4116707	Unable to weight-bear
	44789400	Uses wheelchair outdoors
	4012945	Uses zimmer frame
	45878235	Very limited, Immobile
	4266144	Walking aid use - finding
	439405	Walking disability
Housebound	40299189	Housebound
	4052962	Housebound
	45877743	Bedridden
Activity limitation	44811145	Unfit for activity
	5767124	Difficulty performing personal grooming activity
	36716238	Dependent for personal hygiene activity
	4110470	Difficulty performing personal hygiene activity
	4109859	Unable to perform bathing activity
	4032520	Activity of daily living (ADL) alteration
	4031882	Activity alteration
Visual impairment	4265433	Visual impairment
	4023310	Blindness AND/OR vision impairment level (disorder)
	40305578	Blindness
	44797518	Visual disturbances and blindness
	375545	cataract
	37541	Glaucoma
	374034	Visual disturbance
	42872584	Registration of visual impairment
Hearing impairment	36715579	Acquired hearing loss
	377889	Hearing loss

Condition	Concept ID	Concept Name
	439378	Ear anomalies with hearing impairment
	444291	Sensory hearing loss
	44805060	Wears bone anchored hearing aid
	379832	Mixed conductive AND sensorineural hearing loss
	378444	Hearing disroder
	42539697	External hearing aid in situ
	4246497	Hearing aid
Requirement for care	4192880	Lives in a residential home
	35609081	Home visit requested by care home staff
	4052486	Lives in a nursing home
	4192880	Lives in a residential home
	44791364	Lives in care home
	4074789	Lives in supported home
	4022081	Living in residential institution
	3661927	Living temporarily in care home
	44790305	Local authority care home
	40486978	Nursing home acquired pressure ulcer
	44790706	Pain commenced - residential home
	44791204	Place of occurrence of injury: residential home environment
	44802299	Previously lived in care home
	44788859	Private or voluntary care home
	4147552	Private residential home
	765265	Problem related to living in residential institution
	36713971	Referred by care home
	44814152	Referred by nursing home
	44814153	Referred by residential home
	4119866	Residential home
	44804659	Residential home acquired pressure ulcer
	4305680	Residential institution
	37310422	Seen by clinical pharmacist in care home
	4088536	Seen in nursing home
Social vulnerability	4019836	Social exclusion
	4228687	Impaired social interaction
	4309238	Social isolation
	4019835	Social withdrawal
	4297462	Social isolation (rejection)

Condition	Concept ID	Concept Name
	4172829	Limited social contact
	4317527	Family-related social factor
	42690410	Carer behavior is cause for safeguarding concern
	44792191	Extensive support provided by carer
	4147192	Feeling lonely
	44805255	Has an informal carer
	44805672	Has an older carer
	44807727	Has a paid carer
	44805674	Has a parent carer
	37394063	Has kinship carer
	44813864	Has socially isolated carer
	44806914	Has voluntary carer
	44792192	Inadequate support provided by carer
	4023168	Lives alone
	4052158	Lives alone needs housekeeper
	4053087	Lives alone no help available
	45879223	Lonely
	44789099	Parent is informal carer
	44789487	Partner is informal carer
	44810043	Referred by Social Services
	44790469	Relative is informal carer
	44789986	Report received from social services
	37208707	Requires carer to be present at encounters
	4052789	Social problem
	4209159	Social problem not due to a mental disorder
	44788883	Under care of social services
	4221049	Vulnerable adult
	44791055	Vulnerable elderly person
	44791931	Vulnerable family
	4151777	Vulnerable family support
	44803964	Vulnerable group
	4116985	Vulnerable personality
Falls	4087528	Recurrent falls
	4256754	Falls caused by medication
	4224116	Unexplained recurrent falls
	4329906	At risk for injury due to fall
Urinary incontinence	193326	Urge incontinence of urine
	193598	Extravasation of urine

Condition	Concept ID	Concept Name
	193874	Nocturnal enuresis
	195007	Female stress incontinence
	195079	Functional urinary incontinence
	197102	Unaware of passing urine
	197378	Overflow incontinence of urine
	197672	Urinary incontinence
	443524	Mixed urinary incontinence
	444035	Incontinence
	606405	Extra urethral urinary incontinence
	606955	Stress incontinence following surgical procedure
	4012368	Increased frequency of urination
	4030763	At risk for urge incontinence
	4032498	Abnormal bladder continence
	4032530	Total urinary incontinence
	4092642	Urinary loss
	4096552	Unaware of need to urinate
	4126278	Postural urinary incontinence
	4153667	Urinary incontinence due to urethral sphincter incompetence
	4172646	Urinary incontinence of non-organic origin
	4302457	Double incontinence
	4314023	Incontinence due to detrusor instability
	37119132	Urinary incontinence co-occurrent and due to prolapse of female genital organ
	37208161	Daily urinary incontinence
	40480232	Male urinary stress incontinence
	40481801	Stress incontinence after prostatectomy
	40490423	Incontinence without sensory awareness
	42536555	Stress incontinence co-occurrent and due to pelvic organ prolapse
	42538537	Overflow incontinence of urine due to prolapse of female genital organ
	42538538	Urge incontinence due to prolapse of female genital organ
	42538539	Mixed incontinence due to prolapse of female genital organ
	42872846	Intermittent urinary incontinence
	44808460	Sneezing incontinence of urine

Condition	Concept ID	Concept Name
	45757352	Urinary incontinence due to benign prostatic hypertrophy
	45770268	Functional urinary and faecal incontinence
Weight loss and anorexia	436675	Anorexia nervosa
	4091029	Anorexia symptom
	435928	Abnormal weight loss
	4216971	Mood anorexia
	4333683	Atypical anorexia nervosa
	4300305	Anorexia nervosa, restricting type
	44784528	Anorexia nervosa in remission
	4269485	Anorexia nervosa, binge-eating purging type
	37204325	Facial dysmorphism, anorexia, cachexia, eye and skin anomalies syndrome
	442165	Loss of appetite
	44788734	Complaining of weight loss
	4156515	Malnutrition (calorie)
	763515	Chronic disease-related malnutrition
	45773690	Hypoalbuminemia due to protein calorie malnutrition
	134765	Cachexia
	4109384	Cardiac cachexia
	4347292	Muscle cachexia
	37312021	Malignant cachexia
	36676905	Severe dermatitis, multiple allergies, metabolic wasting syndrome
	4078430	Severe systemic illness-induced respiratory muscle wasting
	4031171	Severe systemic illness-induced skeletal muscle wasting
	4031170	Severe systemic illness tissue wasting
	4123542	Wasting disease
	4229881	Weight loss
Memory and cognitive problems	439795	Minimal cognitive impairment
	443432	Impaired cognition
	761978	Cognitive impairment due to multiple sclerosis
	3654469	Amnestic mild cognitive disorder

Condition	Concept ID	Concept Name
	3654907	Cognitive impairment caused by ingestible alcohol
	4009705	Age-related cognitive decline
	4022572	Disturbance of cognitive learning
	4023989	Cognitive perceptual pattern
	4047110	Language-related cognitive disorder
	4297400	Mild cognitive disorder
	4333671	Age-associated memory impairment
	40480615	Cognitive disorder
	40482301	Residual cognitive deficit as late effect of cerebrovascular accident
	42535016	Cognitive deficit in communication skills
	42535017	Cognitive deficit in visuospatial function
	42535018	Cognitive deficit in psychomotor function
	42535681	Cognitive deficit due to and following ischemic cerebrovascular accident
	42535682	Cognitive deficit due to and following hemorrhagic cerebrovascular accident
	42535706	Cognitive deficit due to and following embolic cerebrovascular accident
	42537139	Dissociative neurological symptom disorder co-occurrent with cognitive symptoms
	42539256	Cognitive deficit due to and following cerebrovascular disease
	42539270	Cognitive deficit due to and following nontraumatic subarachnoid hemorrhage
	42539271	Cognitive deficit due to and following nontraumatic intracerebral hemorrhage
	45765899	Moderate cognitive impairment
	45765900	Severe cognitive impairment
	46271045	Neurocognitive disorder
	40480615	Cognitive disorder
	4182210	Dementia
	42690615	Difficulty remembering past events
	42689981	Difficulty remembering people
	42689830	Difficulty remembering places
	42689982	Difficulty remembering routines
	42690112	Does not remember past events
	42690742	Does not remember people
	42690113	Does not remember places

Condition	Concept ID	Concept Name
	42689849	Does not remember routines
	443432	Impaired cognition
	4103572	Organic memory impairment
	4135668	Poor auditory sequential memory
	4085496	Poor long-term memory
	4084412	Poor short-term memory
	4131380	Poor visual sequential memory
	42690368	Unable to remember past events
	42690369	Unable to remember people
	42690647	Unable to remember places
	42690370	Unable to remember routines
	4141586	Uncompensated short term memory deficit
	4043378	Frontotemporal dementia
Dyspnea	4144682	Expiratory dyspnea
	4192279	Medical Research Council Dyspnoea scale grade 5
	4193263	Medical Research Council Dyspnoea scale grade 2
	4206307	Paroxysmal nocturnal dyspnea
	4212233	Dyspnea after eating
	4217021	Dyspnea, class II
	4219335	Dyspnea, class III
	4219740	Borg Breathlessness Score finding
	4228754	Dyspnea associated with AIDS
	4244276	Paroxysmal dyspnea
	4248284	Dyspnea, class IV
	4263848	Dyspnea on exertion
	4307188	Medical Research Council Dyspnoea scale grade 3
	4310059	Medical Research Council Dyspnoea scale grade 1
	4310172	Medical Research Council Dyspnoea scale grade 4
	35610139	eMRC (extended Medical Research Council) dyspnoea scale grade 1
	35610140	eMRC (extended Medical Research Council) dyspnoea scale grade 2
	35610141	eMRC (extended Medical Research Council) dyspnoea scale grade 3
	35610142	eMRC (extended Medical Research Council) dyspnoea scale grade 5a

Condition	Concept ID	Concept Name
	35610143	eMRC (extended Medical Research Council) dyspnoea scale grade 4
	35610144	eMRC (extended Medical Research Council) dyspnoea scale grade 5b
	36685567	mMRC (modified Medical Research Council) dyspnoea scale grade 0
	36685568	mMRC (modified Medical Research Council) dyspnoea scale grade 1
	36685569	mMRC (modified Medical Research Council) dyspnoea scale grade 2
	36685570	mMRC (modified Medical Research Council) dyspnoea scale grade 3
	36685571	mMRC (modified Medical Research Council) dyspnoea scale grade 4
Sleep disturbance	435657	Dyssomnia
	4204989	Disturbance in sleep behavior
	40480927	Sleep dysfunction with sleep stage disturbance
	40482260	Sleep dysfunction with arousal disturbance
	4132137	Sleep pattern disturbance
	435524	Sleep disorder
	42690122	Does not sleep
	374905	Non-organic sleep disorder
	443544	Organic sleep disorder
	4200883	Poor sleep pattern
	4215402	Primary insomnia
	37110488	Chronic insomnia
	4102985	Nonorganic insomnia
	4086851	Cannot sleep at all
	4115402	Difficulty sleeping
	42689992	Difficulty sleeping without sedation
	42689991	Difficulty sleeping with sedation
	43530738	Disruptions of 24 hour sleep-wake cycle
	4204989	Disturbance in sleep behavior
	42690122	Does not sleep
	42690715	Does not sleep without sedation
	42690123	Does not sleep with sedation
	434172	Insomnia with sleep apnea

Condition	Concept ID	Concept Name
	436522	Irregular sleep-wake pattern
	3173994	Poor sleep hygiene
	4305303	Sleep deprivation
	42690380	Unable to sleep without sedation
	42690379	Unable to sleep with sedation
Anaemia and haematinic deficiency	434622	deficiency anemias
	4306430	Ferritin level low
	37016158	Low serum ferritin
	136949	Refractory anemia with excess blasts (clinical)
	137829	Aplastic anemia
	140065	Pure red cell aplasia
	140681	Constitutional aplastic anemia
	432282	Sideroblastic anemia
	432295	Pernicious anemia
	432588	Megaloblastic anemia due to vitamin B-12 deficiency
	432875	Anemia due to chronic blood loss
	433168	Iron deficiency anemia secondary to inadequate dietary iron intake
	434894	Acute posthemorrhagic anemia
	435503	Hemolytic anemia
	435789	Megaloblastic anemia
	436659	Iron deficiency anemia
	437247	Anemia of chronic disease
	437834	Non-autoimmune hemolytic anemia
	438722	Non megaloblastic anemia associated with nutritional deficiency
	439777	Anemia
	440977	Megaloblastic anemia due to folate deficiency
	440979	Acquired hemolytic anemia
	441258	Anemia in neoplastic disease
	441269	Autoimmune hemolytic anemia
	443961	Anemia of chronic renal failure
	602143	Anemia due to chronic kidney disease stage 1
	603007	Pernicious anemia due to autoimmune disorder
	603011	Vitamin B12 deficiency anemia following total gastrectomy

Condition	Concept ID	Concept Name
	605330	Restless leg syndrome due to iron deficiency anemia
	606065	Congenital megaloblastic anemia due to transcobalamin II deficiency
	606921	Post gastrectomy iron deficiency anemia
	606928	Iron deficiency anemia due to celiac disease
	606940	Drug-induced non autoimmune hemolytic anemia
	607426	Vitamin B12 deficiency anemia due to chronic atrophic gastritis
	608596	Anemia caused by antineoplastic agent
	3661626	Megaloblastic anemia due to dihydrofolate reductase deficiency
	4002495	Refractory anemia with excess blasts in transformation (clinical)
	4003185	Refractory anemia (clinical)
	4003186	Refractory anemia with ringed sideroblasts (clinical)
	4006467	Anemia due to infection
	4006468	Anemia due to physical agent
	4008273	Coombs negative hemolytic anemia
	4008663	Megaloblastic anemia due to exfoliative dermatitis
	4009306	Anemia due to copper deficiency
	4009785	Anemia due to membrane defect
	4015896	Anemia due to niacin deficiency
	4019001	Regenerative anemia
	4021911	Megaloblastic anemia due to poor nutrition
	4031699	Humoral immunologic aplastic anemia
	4032006	Dimorphic anemia
	4032352	Hemolytic anemia due to hyperbaric oxygen
	4034963	Megaloblastic anemia, thiamine-responsive, with diabetes mellitus and sensorineural deafness
	4035974	Drug-induced enzyme deficiency anemia
	4039536	Autoimmune hemolytic anemia due to complement
	4044728	Thiamine-responsive megaloblastic anemia
	4045142	Megaloblastic anemia due to gastrectomy
	4046563	G-6-PD class I variant anemia
	4071444	Anemia of thyroid dysfunction
	4079181	G-6-PD class III variant anemia
	4082253	Unstable hemoglobin disease

Condition	Concept ID	Concept Name
	4082917	Drug-induced immune hemolytic anemia, immune complex type
	4085853	Hemolytic anemia due to babesiosis
	4092893	Anemia due to vitamin E deficiency
	4096927	Megaloblastic anemia due to vitamin B-12 malabsorption with proteinuria
	4097961	Acute megaloblastic anemia due to dialysis
	4098008	Folate deficiency anemia due to malabsorption
	4098009	Folate deficiency anemia due to liver disorders
	4098013	Hemolytic anemia due to hexokinase deficiency
	4098017	Primary cold-type hemolytic anemia
	4098018	Mechanical hemolytic anemia
	4098019	Toxic hemolytic anemia
	4098027	Aplastic anemia due to radiation
	4098131	Myelophthisic anemia
	4098145	Idiopathic aplastic anemia
	4098627	Idiopathic hypochromic anemia
	4098740	Vitamin B12 deficiency anemia due to malabsorption with proteinuria
	4098746	Hemolytic anemia due to pyruvate kinase deficiency
	4098747	Anemia due to disorders of nucleotide metabolism
	4098760	Transient hypoplastic anemia
	4098762	Pyridoxine-responsive sideroblastic anemia
	4099508	Refractory anemia without sideroblasts, so stated
	4099603	Megaloblastic anemia due to hemodialysis
	4100962	Megaloblastic anemia due to impaired absorption of folate
	4100985	Iron deficiency anemia due to dietary causes
	4100987	Folate deficiency anemia, drug-induced
	4100991	Hemolytic anemia due to triose phosphate isomerase deficiency
	4100998	Aplastic anemia due to toxic cause
	4101000	Secondary sideroblastic anemia due to drugs and toxins
	4101001	Chronic anemia
	4101458	Combined B12 and folate deficiency anemia
	4101573	Vitamin C deficiency anemia
	4101582	Aplastic anemia due to chronic disease

Condition	Concept ID	Concept Name
	4101583	Aplastic anemia due to infection
	4101584	Secondary sideroblastic anemia due to disease
	4104541	Anemia due to pentose phosphate pathway defect
	4105643	Myasthenic syndrome due to pernicious anemia
	4114026	Normocytic anemia
	4116343	Normocytic anemia following acute bleed
	4120446	Iron deficiency without anemia
	4120448	Normocytic anemia due to aplasia
	4120450	Normocytic anemia due to chronic blood loss
	4121106	Microcytic anemia
	4121110	Selective malabsorption of cyanocobalamin
	4121115	Sickle cell anemia with high hemoglobin F
	4122079	Deficiency anemias, excluding iron
	4122923	Dilutional anemia
	4122924	Anemia of renal disease
	4122927	Combined deficiency anemia
	4125491	Megaloblastic anemia due to dietary causes
	4125493	Vegan's anemia
	4125630	Chronic non-spherocytic hemolytic anemia
	4130191	Secondary warm autoimmune hemolytic anemia
	4130680	Autoimmune hemolytic anemia, categorized by antibody class AND/OR complement
	4131127	Secondary autoimmune hemolytic anemia
	4131914	Sickle cell anemia with coexistent alpha-thalassemia
	4131915	Primary (idiopathic) autoimmune hemolytic anemia
	4131917	Paroxysmal cold hemoglobinuria
	4131919	Drug-induced immune hemolytic anemia, hapten type
	4132085	Anemia due to alloimmune destruction of transfused red cells
	4135931	Anemia of endocrine disorder
	4143167	Autoimmune hemolytic anemia due to IgA plus complement
	4143351	Folate deficiency anemia due to dietary causes
	4143629	Anemia due to mechanical damage
	4144077	G-6-PD class II variant anemia
	4144811	Anemia due to zinc deficiency

Condition	Concept ID	Concept Name
	4145277	HNSHA due to pyrimidine-5'-nucleotidase deficiency
	4146086	Constitutional aplastic anemia with malformation
	4146088	Aplastic anemia due to drugs
	4146771	Anemia in ovarian carcinoma
	4146936	Drug-induced autoimmune hemolytic anemia
	4147365	Anemia of adrenal dysfunction
	4147491	Vitamin B12 deficiency anemia due to dietary causes
	4147600	Megaloblastic anemia due to pancreatic insufficiency
	4147911	Megaloblastic anemia due to inborn errors of metabolism
	4148471	Fanconi's anemia
	4149183	Megaloblastic anemia due to error of folate metabolism
	4150499	Refractory megaloblastic anemia
	4150547	Anemia secondary to renal failure
	4151502	Hemolytic anemia due to Clostridium welchii
	4155187	Traumatic cardiac hemolytic anemia
	4156842	Intracorpuscular hemolytic anemia
	4157495	Sideropenic anemia with reticuloendothelial siderosis
	4158891	Anemia due to isoimmunization
	4159651	Traumatic hemolytic anemia
	4159748	Hand-foot syndrome in sickle cell anemia
	4160238	Simple chronic anemia
	4160887	Cold autoimmune hemolytic anemia
	4168286	Anemia due to starvation
	4168772	Megaloblastic anemia due to chronic hemolytic anemia
	4171026	HNSHA due to NADH diaphorase deficiency
	4172446	Microcytic normochromic anemia
	4173028	Idiopathic sideroblastic anemia
	4174412	Anemia due to copper
	4176884	Hemolytic anemia due to nonlymphoid neoplasm
	4177177	Cellular immunologic aplastic anemia
	4178677	Congenital nonspherocytic hemolytic anemia due to inborn error of metabolism

Condition	Concept ID	Concept Name
	4184200	Secondary aplastic anemia
	4184603	Megaloblastic anemia due to Zollinger-Ellison syndrome
	4184758	Acquired aplastic anemia
	4195171	Normocytic hypochromic anemia
	4195271	Megaloblastic anemia due to vegetarianism
	4201444	Anemia due to riboflavin deficiency
	4203291	HNSHA due to increased adenosine deaminase activity
	4207240	Anemia due to intrinsic red cell abnormality
	4211348	Aplastic anemia associated with pancreatitis
	4211695	Acute pure red cell aplasia
	4213893	Achlorhydric anemia
	4214023	G-6-PD class V variant anemia
	4215784	Autoimmune hemolytic anemia due to IgM
	4215791	Acute megaloblastic anemia secondary to total parenteral nutrition
	4216915	Hemoglobin S sickling disorder with crisis
	4217370	Aase syndrome
	4218100	Hemolytic anemia due to drugs
	4218974	Hypoplastic anemia
	4219253	Anemia due to arsenic hydride
	4219359	G-6-PD class IV variant anemia
	4219853	Warm autoimmune hemolytic anemia
	4220697	Acute megaloblastic anemia
	4221567	Megaloblastic anemia due to disease of small intestine
	4223031	Anemia associated with AIDS
	4223896	Mycoplasmal anemia
	4225810	Aplastic anemia associated with AIDS
	4228194	Congenital hypoplastic anemia
	4228444	Acquired hemolytic anemia associated with AIDS
	4231887	Secondary acquired sideroblastic anemia
	4234973	Chronic acquired pure red cell aplasia
	4235788	Familial megaloblastic anemia
	4238904	Autoimmune hemolytic anemia due to IgA
	4241982	Congenital dyserythropoietic anemia, type I
	4242111	HNSHA due to phosphoglycerate kinase deficiency

Condition	Concept ID	Concept Name
	4243831	Anemia of pituitary deficiency
	4243950	Megaloblastic anemia due to blind loop syndrome
	4244129	Anemia due to decreased red cell production
	4246105	Hemolytic anemia with emphysema AND cutis laxa
	4247416	Megaloblastic anemia due to congenital deficiency of intrinsic factor
	4250028	Acute megaloblastic anemia due to nitrous oxide
	4254249	HNSHA due to pyruvate kinase deficiency
	4254380	Coombs positive hemolytic anemia
	4258685	HNSHA due to triosephosphate isomerase deficiency
	4260689	Anemia due to multiple mechanisms
	4261354	Megaloblastic anemia due to decreased intake of vitamin B-12
	4262948	Microcytic hypochromic anemia
	4263315	Normocytic normochromic anemia
	4264046	Sickle cell-hemoglobin E disease
	4265915	HNSHA due to diphosphoglycerate mutase deficiency
	4268894	Acute megaloblastic anemia due to severe illness
	4269764	Glucose-6-phosphate dehydrogenase deficiency anemia
	4269919	Autoimmune hemolytic anemia due to IgG plus complement
	4271197	Idiopathic paroxysmal cold hemoglobinuria
	4278920	Anemia due to lead
	4280070	Antibody-mediated anemia
	4280354	Nutritional anemia
	4282785	Megaloblastic anemia due to nontropical sprue
	4284415	Megaloblastic anemia due to increased requirements
	4286660	Congenital dyserythropoietic anemia, type II
	4287402	Anemia of parathyroid dysfunction
	4287574	Megaloblastic anemia due to error of cobalamin metabolism
	4291002	Megaloblastic anemia due to drugs
	4297024	Hemolytic anemia due to Bartonella
	4297537	Hemolytic anemia due to infection
	4298690	Immunologic aplastic anemia

Condition	Concept ID	Concept Name
	4298975	Hemolytic anemia due to malaria
	4300295	Drug-induced sideroblastic anemia
	4303199	Anemia due to pantothenic deficiency
	4306199	Perinatal anemia
	4307469	Sports anemia
	4307799	Anemia due to diabetes mellitus
	4308062	Diaphyseal dysplasia with anemia
	4308125	Macrocytic anemia
	4311676	Anemia due to vitamin A deficiency
	4312008	Anemia due to substance
	4312853	Anemia due to vitamin B-6 deficiency
	4313413	Anemia due to chlorate
	4313581	Hapten type high affinity hemolytic anemia
	4314111	Non megaloblastic anemia due to alcoholism
	4318674	Chronic idiopathic autoimmune hemolytic anemia
	4319914	Anemia due to radiation
	4323223	Anemia due to medication
	4329173	Anemia of gonadal dysfunction
	4330322	Anemia due to disturbance of proliferation AND/OR differentiation of hematopoietic stem cells
	4336555	G-6-PD variant enzyme deficiency anemia
	4338370	Megaloblastic anemia due to alcoholism
	4338976	Megaloblastic anemia due to tropical sprue
	35624317	Hemolytic anemia due to adenylate kinase deficiency
	35624756	Anemia due to and following chemotherapy
	36680584	Autosomal dominant aplasia and myelodysplasia
	36713571	Megaloblastic anemia due to vitamin B12 deficiency secondary to intestinal disease
	36713572	Vitamin B12 deficiency anemia caused by drug
	36713573	Acquired iron deficiency anemia due to increased iron requirement
	36713763	Autoimmune hemolytic anemia mixed type
	36715009	Adult-onset autosomal recessive sideroblastic anemia
	36715492	Megaloblastic anemia due to folate deficiency due to increased requirement
	36715580	Acquired thiamine deficiency anemia
	36715584	Refractory anemia with ringed sideroblasts associated with marked thrombocytosis

Condition	Concept ID	Concept Name
	36716029	Hyperuricemia, anemia, renal failure syndrome
	36716126	Iron-refractory iron deficiency anemia
	36716259	Pancreatic insufficiency, dyserythropoietic anemia, calvarial hyperostosis syndrome
	36716460	X-linked congenital dyserythropoietic anemia with thrombocytopenia
	37016121	Anemia following acute postoperative blood loss
	37016151	Aplastic anemia caused by antineoplastic agent
	37017132	Anemia co-occurrent with human immunodeficiency virus infection
	37017165	GATA binding protein 1 related thrombocytopenia with dyserythropoiesis
	37017285	Acquired hemolytic anemia co-occurrent with human immunodeficiency virus infection
	37018722	Anemia caused by zidovudine
	37019055	Aplastic anemia co-occurrent with human immunodeficiency virus infection
	37019193	Anemia co-occurrent and due to chronic kidney disease stage 3
	37110070	Mitochondrial myopathy with sideroblastic anemia syndrome
	37110336	Acquired iron deficiency anemia due to decreased absorption
	37110727	Nonspherocytic hemolytic anemia due to deficiency of adenosinetriphosphatase
	37110923	Severe congenital hypochromic anemia with ringed sideroblasts
	37111627	Central nervous system calcification, deafness, tubular acidosis, anemia syndrome
	37116297	Secondary autoimmune hemolytic anemia co-occurrent and due to chronic inflammatory disease
	37116298	Secondary autoimmune hemolytic anemia co-occurrent and due to lymphoproliferative disorder
	37116300	Secondary autoimmune hemolytic anemia co-occurrent and due to rheumatic disorder
	37116301	Secondary autoimmune hemolytic anemia co-occurrent and due to ulcerative colitis
	37117740	Secondary autoimmune hemolytic anemia co-occurrent and due to systemic lupus erythematosus

Condition	Concept ID	Concept Name
	37119138	Iron deficiency anemia due to blood loss
	37204236	X-linked dyserythropoietic anemia with abnormal platelets and neutropenia
	37204287	Hemoglobinopathy Toms River
	37204551	Hereditary isolated aplastic anemia
	37312032	Anemia due to chronic infectious disease
	37395652	Anemia in chronic kidney disease stage 5
	37397036	Autosomal recessive sideroblastic anemia
	37398911	Anemia in chronic kidney disease stage 4
	40478891	Erythropoietin resistance in anemia of chronic kidney disease
	40599994	X chromosome-linked sideroblastic anemia
	42536530	Hereditary vitamin B12 deficiency anemia
	42536531	Hereditary folate deficiency anemia
	42537687	Anemia due to metabolic disorder
	42872405	Anemia, pre-end stage renal disease on erythropoietin protocol
	44783626	Pulmonary arterial hypertension associated with chronic hemolytic anemia
	44806268	Refractory anaemia with multilineage dysplasia
	44810002	Recurrent anaemia
	45768812	Anemia in chronic kidney disease
	45768813	Anemia in end stage renal disease
	45768941	Chronic hemolytic anemia
	45773534	Anemia in malignant neoplastic disease
	46272744	Hypochromic microcytic anemia with iron overload
Hypertension	312648	Benign essential hypertension
	4215640	Benign essential hypertension complicating AND/OR reason for care during childbirth
	4034031	Benign essential hypertension complicating AND/OR reason for care during pregnancy
	4148205	Benign essential hypertension complicating AND/OR reason for care during puerperium
	4269358	Benign essential hypertension in obstetric context
	320128	Essential hypertension
	4083723	Essential hypertension complicating AND/OR reason for care during childbirth

Condition	Concept ID	Concept Name
	4302591	Essential hypertension complicating AND/OR reason for care during pregnancy
	4321603	Essential hypertension complicating AND/OR reason for care during puerperium
	4217486	Essential hypertension in obstetric context
	4058987	High-renin essential hypertension
	4159755	Labile essential hypertension
	4263067	Low-renin essential hypertension
	317898	Malignant essential hypertension
	45757787	Postpartum pre-existing essential hypertension
	4180283	Systolic essential hypertension
Diabetes	201826	Type 2 diabetes mellitus
	201254	Type 1 diabetes mellitus
	443731	Renal disorder due to type 2 diabetes mellitus
	200687	Renal disorder due to type 1 diabetes mellitus
	443729	Peripheral circulatory disorder due to type 2 diabetes mellitus
	318712	Peripheral circulatory disorder due to type 1 diabetes mellitus
	376065	Disorder of nervous system due to type 2 diabetes mellitus
	377821	Disorder of nervous system due to type 1 diabetes mellitus
	443733	Disorder of eye due to type 2 diabetes mellitus
	42538169	Disorder of eye due to type 1 diabetes mellitus
	443732	Disorder due to type 2 diabetes mellitus
	435216	Disorder due to type 1 diabetes mellitus
Osteoporosis	80502	Osteoporosis
	37204244	X-linked osteoporosis with fractures
	4109181	Osteoporosis with pseudoglioma
	44783850	Osteoporosis circumscripta
	36716194	Osteoporosis and oculocutaneous hypopigmentation syndrome
Chronic kidney disease	46271022	Chronic kidney disease
	75865	Disorder of the urinary system
	197331	Disorder of urinary tract
Skin Ulcer	4262920	Skin ulcer

Condition	Concept ID	Concept Name
	46269755	Chronic non-pressure ulcer of calf extending to fat level
	46269752	Chronic non-pressure ulcer of ankle extending to fat level
Ischeamic heart disease	4185932	Ischeamic heart disease
Heart Failure	316139	Heart failure
Cerebrovascular disease	381591	Cerebrovascular disease
Peripheral vascular disease	321052	Peripheral vascular disease
Atrial fibrilation	313217	Atrial fibrilation
Heart valve disease	4281749	Heart valve disorder
Hypotension/syncope	317002, 40350983	Low blood pressure
	40316030	Hypotension
	319041	Orthostatic hypotension
	135360, 40498271	Syncope
Foot problem	4101512	Foot problem
	4268887	Chiropody follow-up
	4053100	Domiciliary chiropody
	4136647	Seen by community-based podiatrist
	4138349	Seen by community-based podiatry service
	4140790	Seen by hospital-based podiatrist
	4140924	Seen by hospital-based podiatry service
	42539590	Seen by podiatric surgeon
	4083436	Seen by podiatrist
	4139895	Seen by podiatry service
	4085778	Seen in chiropody clinic
	4139217	Under care of community-based podiatrist
	4139218	Under care of hospital-based podiatrist
	42539494	Under care of podiatric surgeon
	4139705	Under care of podiatrist
Arthritis	4291025	Arthritis
	40555828	Arthritis
Chronic Respiratory disease	4063381	Chronic disease of the respiratory system
	261325	Pulmonary emphysema
	255573	Chronic obstructive lung disease
	317009	Asthma
Peptic ulcer	4027663	Peptic ulcer

Condition	Concept ID	Concept Name
	4057060	Acute Peptic ulcer
	4134146	Chronic Petic ulcer
Thyroid disease	141253	Disorder of thyroid gland
	4194160	Thyroid function tests abnormal
Fragility fracture	44791986	Fragility fracture
	40480160	Osteoporotic fracture
	4001142	Osteopathies, chondropathies and acquired musculoskeletal deformities
	44791986	Fragility fracture
	4174520	Fracture of vertebral column
	4050747	Fracture of upper limb
	4053828	Fracture of thoracic spine
	4302740	Fracture of sternum
	4142905	Fracture of rib
	4013613	Fracture of lumbar spine and/or pelvis
	4278672	Fracture of forearm
	442560	Fracture of femur
	4129393	Fracture of cervical spine
	4015350	Fracture at wrist and/or hand level
	4001458	Fatigue fracture of vertebra
	4344386	Disorder of continuity of bone
	4222001	Collapse of vertebra
Parkinsonism and tremor	381270	Parkinson's disease
	36716783	Atypical Parkinsonism
	37110549	Functional parkinsonism
	4140090	Parkinsonism
	372604	Movement disorder
	443782	Tremor

Concept IDs include descendants unless highlighted as being excluded. By OMOP standards descendants automatically include the ancestor.
Before finalising the concept sets, CohortDiagnostics will run on cohorts created using the initial concept sets to check code counts and patient characteristics which might give indications to adjust the concept sets.

Appendix II: ENCePP checklist for study protocols

ENCePP Checklist for Study Protocols (Revision 4)

Study title: DARWIN EU® – Frailty and polypharmacy among adults aged 18 and above with cancer at the time of diagnosis

EU PAS Register® number: N/A
Study reference number (if applicable): N/A

EU PAS Register® number:
Study reference number (if applicable):

Section 1: Milestones	**Yes**	**No**	**N/A**	**Section Number**
1. Does the protocol specify timelines for 1.1.1 Start of data collection 1.1.2 End of data collection 1.1.3 Progress report(s) 1.1.4 Interim report(s) 1.1.5 Registration in the EU PAS Register® 1.1.6 Final report of study results.	X			5. Milestones, 8.2 Data Sources

Comments:

Section 2: Research question	**Yes**	**No**	**N/A**	**Section Number**
2.1 Does the formulation of the research question and objectives clearly explain: 2.1.1 Why the study is conducted? (e.g. to address an important public health concern, a risk identified in the risk management plan, an emerging safety issue) 2.1.2 The objective(s) of the study? 2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised) 2.1.4 Which hypothesis(-es) is (are) to be tested? 2.1.5 If applicable, that there is no *a priori* hypothesis?	X			7. Research question and objectives 8. Research methods

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Section Number**
3.1 Is the study design described? (e.g. cohort, case-control, cross-sectional, other design)	X			8.1 Study type and Study Design
3.2 Does the protocol specify whether the study is based on primary, secondary or combined data collection?	X			8.2 Study Setting and Data Sources
3.3 Does the protocol specify measures of occurrence? (e.g., rate, risk, prevalence)	X			8.8 Analysis

3.4	Does the protocol specify measure(s) of association? (e.g. risk, odds ratio, excess risk, rate ratio, hazard ratio, risk/rate difference, number needed to harm (NNH))	X			8.8 Analysis
3.5	Does the protocol describe the approach for the collection and reporting of adverse events/adverse reactions? (e.g. adverse events that will not be collected in case of primary data collection)			X	

Comments:

Section 4: Source and study populations		**Yes**	**No**	**N/A**	**Section Number**
4.1	Is the source population described?	X			8.5 Study Population
4.2	Is the planned study population defined in terms of: 4.2.1 Study time period 4.2.2 Age and sex 4.2.3 Country of origin 4.2.4 Disease/indication 4.2.5 Duration of follow-up	X			8.3 Study Period 8.6.3. Other covariates 8.2 Study Setting and Data Sources 8.6.1. Exposures 8.4 Follow-up
4.3	Does the protocol define how the study population will be sampled from the source population? (e.g. event or inclusion/exclusion criteria)	X			8.5 Study Population with inclusion and exclusion criteria

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
5.1 Does the protocol describe how the study exposure is defined and measured? (e.g. operational details for defining and categorising exposure, measurement of dose and duration of drug exposure)	X			8.6.1. Exposures
5.2 Does the protocol address the validity of the exposure measurement? (e.g. precision, accuracy, use of validation sub-study)			X	
5.3 Is exposure categorised according to time windows?	X			8.6.1. Exposures
5.4 Is intensity of exposure addressed? (e.g. dose, duration)	X			8.6.1. Exposures
5.5 Is exposure categorised based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?			X	
5.6 Is (are) (an) appropriate comparator(s) identified?	X			8.8 Analysis

Comments:

Section 6: Outcome definition and measurement	**Yes**	**No**	**N/A**	**Section Number**
6.1 Does the protocol specify the primary and secondary (if applicable) outcome(s) to be investigated?			X	
6.2 Does the protocol describe how the outcomes are defined and measured?			X	

6.3 Does the protocol address the validity of outcome measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, use of validation sub-study)			X	
6.4 Does the protocol describe specific outcomes relevant for Health Technology Assessment? (e.g. HRQoL, QALYs, DALYS, health care services utilisation, burden of disease or treatment, compliance, disease management)			X	

Comments:

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

Comments:

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

Comments:

Section 9: Data sources	**Yes**	**No**	**N/A**	**Section Number**
9.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
9.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview)	X			8.6.1. Exposures
9.1.2 Outcomes? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics)			X	
9.1.3 Covariates and other characteristics?	X			8.6.3. Other covariates
9.2 Does the protocol describe the information available from the data source(s) on:				
9.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	X			8.2 Study Setting and Data Sources
9.2.2 Outcomes? (e.g. date of occurrence, multiple event, severity measures related to event)			X	
9.2.3 Covariates and other characteristics? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, lifestyle)	X			8.2 Study Setting and Data Sources
9.3 Is a coding system described for:				
9.3.1 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC) Classification System)	X			8.6.1. Exposures
9.3.2 Outcomes? (e.g. International Classification of Diseases (ICD), Medical Dictionary for Regulatory Activities (MedDRA))			X	

9.3.3 Covariates and other characteristics?	X			8.6.3. Other covariates
9.4 Is a linkage method between data sources described? (e.g. based on a unique identifier or other)			X	

Comments:

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Section Number**
10.1 Are the statistical methods and the reason for their choice described?	X			8.8 Analysis
10.2 Is study size and/or statistical precision estimated?			X	
10.3 Are descriptive analyses included?	X			8.8.2 Descriptive statistics
10.4 Are stratified analyses included?	X			8.8 Analysis
10.5 Does the plan describe methods for analytic control of confounding?			X	
10.6 Does the plan describe methods for analytic control of outcome misclassification?			X	
10.7 Does the plan describe methods for handling missing data?			X	
10.8 Are relevant sensitivity analyses described?			X	

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Section Number**
11.1 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	X			9. Data management
11.2 Are methods of quality assurance described?	X			10. Quality Control
11.3 Is there a system in place for independent review of study results?			X	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Section Number**
12.1 Does the protocol discuss the impact on the study results of: 12.1.1 Selection bias? 12.1.2 Information bias? 12.1.3 Residual/unmeasured confounding? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods).	X			11. Limitations of the research methods
12.2 Does the protocol discuss study feasibility? (e.g. study size, anticipated exposure uptake, duration of follow-up in a cohort study, patient recruitment, precision of the estimates)	X			Table 8.2. Description of the selected Data Sources.

Comments:

Section 13: Ethical/data protection issues	Yes	No	N/A	Section Number
13.1 Have requirements of Ethics Committee/ Institutional Review Board been described?	X			13. Governance board aspects
13.2 Has any outcome of an ethical review procedure been addressed?			X	
13.3 Have data protection requirements been described?	X			9.2 Data storage and protection

Comments:

Section 14: Amendments and deviations	Yes	No	N/A	Section Number
14.1 Does the protocol include a section to document amendments and deviations?	X			4. Amendments and updates

Comments:

Section 15: Plans for communication of study results	Yes	No	N/A	Section Number
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	X			14. Plans for disseminating and communicating study results
15.2 Are plans described for disseminating study results externally, including publication?	X			14. Plans for disseminating and communicating study results

Comments:

Name of the main author of the protocol: Talita Duarte-Salles

Date: 26/01/2024

Signature: _____