Hospital District of Helsinki and Uusimaa patient cohort (FinOMOP)

First published: 01/02/2024

Last updated: 12/11/2024

Data source | Human | Other

Administrative details

Administrative details

Data source ID

1111134

Data source acronym

FinOMOP_Helsinki

Data holder

Hospital District of Helsinki and Uusimaa (HUS)

Data source type

Other

Data source type, other

Electronic health records

Main financial support

Funding by own institution

Funding from public-private partnership

Care setting

Hospital inpatient care

Hospital outpatient care

Secondary care – specialist level (ambulatory)

Data source qualification

If the data source has successfully undergone a formal qualification process (e.g., from the EMA, ISO or other certifications), this should be described.

Yes

Description of the qualification

ISO 13485:2016, ISO 9001:2015, EHDEN Data Partner

Data source website

https://aineistokatalogi.fi/catalog/studygroups/7c79b717-ad67-4688-98c9-806d03faf407

Contact details

Eric Fey eric.fey@hus.fi

Main

eric.fey@hus.fi

Data source regions and languages

Data source countries

Finland

Data source languages

Finnish

Data source regions

Uusimaa

Data source establishment

Data source established

15/06/2018

Data source time span

First collection: 15/06/2004

The date when data started to be collected or extracted.

Publications

Data source publications

POSC128 Healthcare Resource Utilization in Adults Diagnosed with Acute Myeloid Leukemia (AML) in Helsinki and Uusimaa Hospital District

Implementing a Functional Precision Medicine Tumor Board for Acute Myeloid Leukemia

Machine Learning of Bone Marrow Histopathology Identifies Genetic and Clinical Determinants in Patients with MDS

Studies

List of studies that have been conducted using the data source

DARWIN EU® - Chondrosarcoma: patient demographics, treatments, and survival in the period 2010-2023

DARWIN EU® - Prescription trends of ketamine and esketamine

DARWIN EU® – Incidence rates of venous thromboembolic events in patients with selected cancers

DARWIN EU® - Monitoring prescription of essential medicines administered in ICU

DARWIN EU® - Eye disorders in women with breast cancer treated with anastrozole, letrozole or tamoxifen

DARWIN EU® - Characterisation of acute renal outcomes and diabetic complications among patients with concomitant use of metformin and iodinated contrast agents

DARWIN EU® - Drug utilisation study on antibiotics in the ‘Access’ category of the WHO AWaRe classification of antibiotics for evaluation and monitoring of use

DARWIN EU® - Drug Utilisation Study on Antibiotics in the ‘Watch’ category of the WHO AWaRe classification of antibiotics for evaluation and monitoring of use

DARWIN EU® - Coverage of meningococcal vaccines in the target population in Europe

DARWIN EU® - RR Childhood hypertension and sartans prescribing in children

DARWIN EU® - Incidence, period prevalence, and characterisation of individuals with paediatric pulmonary arterial hypertension

DARWIN EU® - Multiple myeloma: patient characterisation, treatments, and survival in the period 2012–2024

DARWIN EU® - Acute myeloid leukaemia: incidence, patient characteristics, treatments, and survival in the period 2015–2024

DARWIN EU® – Drug Utilisation Study of terbinafine-containing products

DARWIN EU® – Proof-of-concept: Preparedness for annual seasonal influenza vaccine effectiveness studies - Vaccine coverage and incidence of influenza-related outcomes

DARWIN EU® - Time to onset of thromboembolic events in adults with selected types of cancer

Data elements collected

The data source contains the following information

Disease information

Does the data source collect information with a focus on a specific disease? This might be a patient registry or other similar initiatives.

No

Rare diseases

Are rare diseases captured? In the European Union a rare disease is one that affects no more than 5 people in 10,000.

Yes

Pregnancy and/or neonates

Does the data source collect information on pregnant women and/or neonatal subpopulation (under 28 days of age)?

Yes

Hospital admission and/or discharge

Yes

ICU admission

Is information on intensive care unit admission available?

Yes

Cause of death

Captured

Prescriptions of medicines

Captured

Prescriptions vocabulary

ATC

Dispensing of medicines

Captured

Advanced therapy medicinal products (ATMP)

Is information on advanced therapy medicinal products included? A medicinal product for human use that is either a gene therapy medicinal product, a somatic cell therapy product or a tissue engineered products as defined in Regulation (EC) No 1394/2007 [Reg (EC) No 1394/2007 Art 1(1)].

Yes

Contraception

Is information on the use of any type of contraception (oral, injectable, devices etc.) available?

Yes

Indication for use

Does the data source capture information on the therapeutic indication for the use of medicinal products?

Captured

Medical devices

Is information on medicinal devices (e.g., pens, syringes, inhalers) available?

Yes

Administration of vaccines

Yes

Procedures

Does the data source capture information on procedures (e.g., diagnostic tests, therapeutic, surgical interventions)?

Captured

Procedures vocabulary

Other

Procedures vocabulary, other

Nomesco NCSP

Healthcare provider

Is information on the person providing healthcare (e.g., physician, pharmacist, specialist) available? The healthcare provider refers to individual health professionals or a health facility organisation licensed to provide health care diagnosis and treatment services including medication, surgery and medical devices.

Yes

Clinical measurements

Is information on clinical measurements (e.g., BMI, blood pressure, height) available?

Yes

Genetic data

Are data related to genotyping, genome sequencing available?

Captured

Genetic data vocabulary

Other

Genetic data vocabulary, other

OMOP Genomics

Biomarker data

Does the data source capture biomarker information? The term “biomarker” refers to a broad subcategory of medical signs (objective indications of medical state observed from outside the patient), which can be measured accurately and reproducibly. For example, haematological assays, infectious disease markers or metabolomic biomarkers.

Captured

Patient-reported outcomes

Is information on patient-reported outcomes (e.g., quality of life) available?

Yes

Patient-generated data

Is patient-generated information (e.g., from wearable devices) available?

Yes

Units of healthcare utilisation

Are units of healthcare utilisation (e.g., number of visits to GP per year, number of hospital days) available or can they be derived? Units of healthcare utilisation refer to the quantification of the use

of services for the purpose of preventing or curing health problems.

Yes

Unique identifier for persons

Are patients uniquely identified in the data source?

Yes

Diagnostic codes

Captured

Diagnosis / medical event vocabulary

ICD-10

Other

Diagnosis / medical event vocabulary, other

ICD-10-FI

Medicinal product information

Captured

Medicinal product information collected

Active ingredient(s)

Dose

Route of administration

Medicinal product vocabulary

ATC

Quality of life measurements

Captured

Quality of life measurements vocabulary

15D

Lifestyle factors

Captured

Lifestyle factors

Tobacco use

Sociodemographic information

Captured

Sociodemographic information collected

Age

Gender

Quantitative descriptors

Population Qualitative Data

Population age groups

Paediatric Population (< 18 years)

Preterm newborn infants (0 – 27 days)

Term newborn infants (0 – 27 days)

Infants and toddlers (28 days – 23 months)

Children (2 to < 12 years)

Adolescents (12 to < 18 years)

Adults (18 to < 46 years)

Adults (46 to < 65 years)

Elderly (≥ 65 years)

Adults (65 to < 75 years)

Adults (75 to < 85 years)

Adults (85 years and over)

Estimated percentage of the population covered by the data source in the catchment area

100%

Description of the population covered by the data source in the catchment area whose data are not collected (e.g., people who are registered only for private care)

All inhabitants of the region are entitled to public healthcare, specialized and emergency health care. Very few people use only private servives.

Family linkage

Family linkage available in the data source permanently or can be created on an ad hoc basis

Ad hoc

Population

Population size

2606149

Active population size

2185750

Population by age group

Age group	Population size	Active population size
Term newborn infants (0 – 27 days)	1597	1128
Infants and toddlers (28 days – 23 months)	38052	37491
Children (2 to < 12 years)	216646	215472
Adolescents (12 to < 18 years)	147016	145878
Adults (18 to < 46 years)	881099	854305
Adults (46 to < 65 years)	595414	514639
Adults (65 to < 75 years)	308049	223830
Adults (75 to < 85 years)	260161	145616
Adults (85 years and over)	158115	47391

Median observation time

Median time (years) between first and last available records for unique individuals captured in the data source

7.00

Median time (years) between first and last available records for unique active individuals (alive and currently registered) capt

8.00

Data flows and management

Access and validation

Governance details

Documents or webpages that describe the overall governance of the data source and processes and procedures for data capture and management, data quality check and validation results (governing data access or utilisation for research purposes).

FinOMOP_data_governance

English (696.7 KB - PPTX) View document

https://www.hus.fi/tutkimus-ja-opetus/tutkijan-palvelut/sahkoiset-palvelut-tutkijalle/hus-acamedic-tietoturvallinen#ohjelmavalikoima-hus-acamedici

Biospecimen access

Are biospecimens available in the data source (e.g., tissue samples)?

Yes

Biospecimen access conditions

Hospital biobank can collect all kinds of biospecimen in different phases of the treatment based on patients' concent. E.g. DNA, serum, and plasma samples, fresh frozen tumor samples, SCF-samples and FFPE samples.
https://helsinginbiopankki.fi

Access to subject details

Can individual patients/practitioners/practices included in the data source be contacted?

Yes

Description of data collection

All patients who visit the hospital are recorded in our IT system. All visits, all procedures and given treatments have been recorded systematically in the electronic format. We use more than hundred different operational IT systems. For secondary use, the individual level data is pooled into a data lake. Data relevant for research is then collected through an ETL-process and provided to dedicated secure analytics environments for a specific research purpose (project). Furtehr, we have the OMOP mapping and ETL processes built in collaboration with the other Finnish University Hospitals and the Institute of Health and Welbeing (FinOMOP Consortium). Patient events are mainly of the following types: outpatient visit (with a specific start and end time), inpatient episodes (with a specific start and end time) or additional clinical events (lab, radiology, procedure) with specific timestamps. The observation period for a patient starts from the earliest event and end at the latest event.

Event triggering registration

Event triggering registration of a person in the data source

Birth

Other

Event triggering registration of a person in the data source, other

Contact with the HUS, visit

Event triggering de-registration of a person in the data source

Death

Event triggering creation of a record in the data source

Any contact with HUS

Data source linkage

Linkage

Is the data source described created by the linkage of other data sources (prelinked data source) and/or can the data source be linked to other data source on an ad-hoc basis?

Yes

Linkage description, pre-linked

Information about deaths.

Linkage description, possible linkage

Defined patient group in the hospital data base

Linked data sources

Pre linked

Is the data source described created by the linkage of other data sources?

No

Data source, other

Data from many nation-wide health registries such as drug purchase, visual impairment, cancer, retirement due to a disease registries and many others, can be combined to the Tampere University Hospital patient registry. The

combination needs a specific research plan and data permit.

Linkage strategy

Combination

Linkage variable

social security number

Linkage completeness

High completeness

Pre linked

Is the data source described created by the linkage of other data sources?

Yes

Data source, other

Digital and Population Data Services Agency of Finland

Linkage variable

ssn

Linkage completeness

Almost complete (emigration).

Data management specifications that apply for the data source

Data source refresh

Monthly

Informed consent for use of data for research

Not Required

Possibility of data validation

Can validity of the data in the data source be verified (e.g., access to original medical charts)?

Yes

Data source preservation

Are records preserved in the data source indefinitely?

Yes

Approval for publication

Is an approval needed for publishing the results of a study using the data source?

Yes

Data source last refresh

29/04/2023

Common Data Model (CDM) mapping

CDM mapping

Has the data source been converted (ETL-ed) to a common data model?

Yes

CDM Mappings

CDM name

OMOP

CDM website

https://www.ohdsi.org/Data-standardization/

Data source ETL CDM version

5.4

Data source ETL frequency

0,25 months

Data source ETL status

Completed