Auria Clinical Informatics (FinOMOP)

First published: 01/02/2024

Last updated: 18/12/2024

Data source | Human | Biobank | Hospital inpatient records | Other

Administrative details

Administrative details

PURI

https://redirect.ema.europa.eu/resource/1111111

Data source ID

1111111

Data source acronym

ACI

Data holder

Southwest Finland Wellbeing Services County (VarHa)

Data source type

Biobank

Hospital inpatient records

Other

Data source type, other

Electronic health records

Main financial support

National, regional, or municipal public funding

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

Part of the FinOMOP consortium (a DARWIN EU data partner).

Data source website

Auria Clinical Informatics at the Wellbeing services county of Southwest Finland (VarHa)

Contact details

Tutkimuksen tietopalvelut Auria Clinical Informatics

Main

tutkimuksentietopalvelut@varha.fi

pia tajanen-doumbouya

Alternate

pia.tajanen@varha.fi

Data source regions and languages

Data source countries

Finland

Data source languages

Finnish

Swedish

Data source regions

Varsinais-Suomi

Data source establishment

Data source established

01/01/2004

Data source time span

First collection: 01/01/2004

The date when data started to be collected or extracted.

Publications

Data source publications

How Reliable Are Trial-based Prognostic Models in Real-world Patients With Metastatic Castration-resistant Prostate Cancer? Seyednasrollah F, Mahmoudian M, Rautakorpi L, Hirvonen O, Laitinen T, Jyrkkiö S, Elo L. Eur Urol. 2017 May; 71(5):838-840. doi:10.1016/j.eururo.2017.01.043. Epub 2017 Feb 8.

Machine learning-based dynamic mortality prediction after traumatic brain injury. Raj R, Luostarinen T, Pursiainen E, Posti J, Takala R, Bendel S, Konttila T, Korja M. Scientific Reports. 2019 Nov 27; 9(1):17672. doi:10.1038/s41598-019-53889-6.

The burden of adult asthma in Finland: impact of disease severity and eosinophil count on health care resource utilization Viinanen A, Lassenius MI, Toppila I, et al. [published online ahead of print, 2019 Jul 3]. J Asthma. 2019;1-11. doi:10.1080/02770903.2019.1633664

CA 19-9 doubling time in pancreatic cancer as a predictor of venous thromboembolism: a hospital database study Peippo MH, Kurki S, Seppänen H, Lassila R, Carpén O. Acta Oncol. 2020 Feb;59(2):237-241. doi: 10.1080/0284186X.2019.1679881. Epub 2019 Oct 25

Cardiovascular event rates increase after each recurrence and associate with poor statin adherence. Lassenius MI, Toppila I, Bergius S, Perttilä J, Airaksinen KJ, Pietilä M. European Journal of Preventive Cardiology (2020 Feb) DOI:10.1177/2047487320904334

Studies

List of studies that have been conducted using the data source

Incidence and Prevalence of Interstitial Lung Disease and their progressive fibrosing phenotypes in 6 European Countries (PERSEIDS)

DARWIN EU® Drug utilisation of valproate-containing medicinal products in women of childbearing potential

DARWIN EU® Multiple myeloma: patient characterisation, treatments and survival in the period 2012-2022

DARWIN EU® Drug utilization study of prescription opioids

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

Cause of death vocabulary

ICD-10

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

Source vocabulary ID’s: MIKROBI_FIN_VSSHP (microbes list), NCSP_MODIFIER_FIN (procecure modifiers from the Nomesco NCSP vocabulary), ICD-10-FIN (Finnish modification of ICD-10), PATO_DGN_FIN (Snomed II –based pathology organ-diagnosis pairs), Hilmo_eala (Finnish national medical specialty classification), UNIT_FIN (measurement units list found in data), NCSP-FIN (Finnish modification of Nomesco NCSP procedure classification), LAB-FIN (Finnish national lab test classification), LAB-FIN-VSSHP (local additions to Finnish national lab test classification), ATC.
Available in prescribed medication data: Vnr (Nordic medicine product identifiers https://www.laaketietokeskus.fi/en/pharmaceutical-information/vnr-services)

Medicinal product information

Captured

Medicinal product information collected

Active ingredient(s)

Dose

Package size

Route of administration

Strength

Medicinal product vocabulary

ATC

Other

RxNorm

If 'other,' what vocabulary is used?

Vnr (Nordic medicine product identifiers

https://www.laaketietokeskus.fi/en/pharmaceutical-information/vnr-services)

Quality of life measurements

Captured

Quality of life measurements vocabulary

15D

EQ5D

Lifestyle factors

Captured

Lifestyle factors

Tobacco use

Sociodemographic information

Captured

Sociodemographic information collected

Age

Ethnicity

Gender

Other

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% of the population of the Wellbeing services county of Southwest Finland (Varha) needing specialized or emergency health care.

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)

Everyone in Finland is entitled to the public healthcare. Number of people/amount of data not included in our collections is next to none, so not scientifically/statistically significant.

Comment to question 9) ACI's Analytical Data warehouse from which the OMOP CDM has been extracted, was established in 2014. The actual patient data records are much older, some from 1986 onwards, most from 2004 onwards.

Comment to question 26 and 28) We can not make a difference between preterm and term newborn infants.

Comment to question 27) There are approximately 268,000 patients who are alive and have an event in the CDM data within a period of one year, counting backwards from today. This is our best estimate of the number of active patients. Our hospital treats approximately this number of patients yearly.

Comments to question 46) Data is available for medicines prescribed, coded in ATC and Vnr. Route of administration is also in structured format. Indication is not specifically recorded, but in previous studies we have attempted to deduce it based on relevant ICD10 diagnoses marked near the prescription date. Strength is available. For inpatient (=hospital-administered) drugs, dose is available but the format varies: If the dose is always the same, then it is available in structured format. If the dose varies, it is usually written in free-text format. If the dose is specifically calculated based on patient’s surface area (chemotherapy), it is available in structured format. For outpatients drugs, the dose is always in free-text format.

We are currently working on extracting structured information from free-text drug dosage data. These developments would allow us and our FinOMOP partners to compute the duration of treatment as well as the total daily dose. The duration of treatment is simpler to deduce for inpatient drugs. We know the drug ends at least when the patient leaves the hospital, even if the prescription is left unclosed (which it often is). For IV chemothera

Family linkage

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

Ad hoc

Population

Population size

835895

Active population size

701734

Population by age group

Age group	Population size	Active population size
Paediatric Population (< 18 years)	142464	141648
Preterm newborn infants (0 – 27 days)	19216	18934
Term newborn infants (0 – 27 days)	19216	18934
Infants and toddlers (28 days – 23 months)	16936	16828
Children (2 to < 12 years)	65186	64992
Adolescents (12 to < 18 years)	41126	40894
Adults (18 to < 46 years)	281639	276644
Adults (46 to < 65 years)	164398	144623

Age group	Population size	Active population size
Elderly (≥ 65 years)	231459	128230
Adults (65 to < 75 years)	94836	70625
Adults (75 to < 85 years)	81637	43678
Adults (85 years and over)	54986	13927

Median observation time

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

8.16

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

7.53

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

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.

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 (Oberon). 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 into a single analytical SQL data base. On top the research data base, 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

Disease diagnosis

Start of treatment

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

Other

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

In the event of relocation to another county person is registered at the destination county. Patient history still stays in the data source.

Event triggering creation of a record in the data source

Any contact with HDSF/VarHa, first event.

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 is loaded into ACI DW as batch transfer a few times a year

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 Turku University Hospital patient registry. The combination needs a specific research plan and data permit.

Linkage strategy

Combination

Linkage variable

SSN

Pre linked

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

Yes

Data source, other

Statistics Finland

Linkage strategy

Deterministic

Linkage variable

Social security number

Linkage completeness

High completeness, deaths abroad can be missed

Data management specifications that apply for the data source

Data source refresh

Every 6 months

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

11/04/2024

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

5.3.1

Data source ETL frequency

3,00 months

Data source ETL status

Completed