Lung Cancer Focus DATASET

First published: 01/02/2024

Last updated: 17/10/2024

Data source   Human   Cancer registry   Other

Administrative details

Administrative details

PURI

https://redirect.ema.europa.eu/resource/1111152

Data source ID

1111152

Data source acronym

LUCAS

Data holder

OAKS Consulting

Data source type

Cancer registry

Other

Data source type, other

Electronic health records

Main financial support

Funding by own institution

Funding from industry or contract research

Care setting

Hospital inpatient care

Hospital outpatient care

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.

No

Data source website

https://www.lucascz.cz

Contact details

Jaroslav Duba

(**Main**)

jaroslav.duba@oaks.cz

Data source regions and languages

Data source countries

Czechia

Data source languages

Czech

English

Data source establishment

Data source established

01/06/2018

Data source time span

First collection: 01/06/2018

The date when data started to be collected or extracted.

Publications

Data source publications

Lung Cancer in Non-smokers in Czech Republic: Data from LUCAS Lung Cancer Clinical Registry

Laboratory Parameters Associated With Inflammation Do Not Affect PD-L1 Expression in Non-small Cell Lung Cancer

Worse Prognosis in the Symptomatic Patients With Lung Cancer - Czech Multicentric Study

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.

Yes

Disease details (other)

Lung cancer

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)?

No

Hospital admission and/or discharge

No

ICU admission

Is information on intensive care unit admission available?

No

Cause of death

Captured

Cause of death vocabulary

ICD-10

SNOMED CT

Prescriptions of medicines

Captured

Prescriptions vocabulary

ATC

RxNorm

Dispensing of medicines

Captured

Dispensing vocabulary

ATC

RxNorm

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)].

No

Contraception

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

No

Indication for use

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

Captured

Indication vocabulary

ICD-10

SNOMED CT

Medical devices

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

No

Administration of vaccines

No

Procedures

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

Captured

Procedures vocabulary

Other

SNOMED CT

Procedures vocabulary, other

CZ codes

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

HGNC

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.

No

Unique identifier for persons

Are patients uniquely identified in the data source?

Yes

Diagnostic codes

Captured

Diagnosis / medical event vocabulary

ICD-10

SNOMED CT

Medicinal product information

Captured

Medicinal product information collected

Brand name

Dosage regime

Formulation

Package size

Strength

Medicinal product vocabulary

ATC

SNOMED

Quality of life measurements

Not Captured

Lifestyle factors

Captured

Lifestyle factors

Alcohol use

Tobacco use

Sociodemographic information

Captured

Sociodemographic information collected

Age

Gender

Quantitative descriptors

Population Qualitative Data

Population age groups

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

60–80 % population covered. 11/14 of the CZ pneumooncology centers are included in the the LUCAS datasource.

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)

LUCAS is the prospective follow-up study of patients with Malignant neoplasm of bronchus and lung (ICD-10 code C34) from diagnosis in a lifelong horizon.

Data on the population treated without health insurance or treated privately are not collected, but represent up-to 5% of an overall population.

Population

Population size

6245

Active population size

2835

Population by age group

Age group	Population size	Active population size
Adults (18 to < 46 years)	94	56
Adults (46 to < 65 years)	1288	589
Elderly (≥ 65 years)		1110
Adults (65 to < 75 years)	2782	1210
Adults (75 to < 85 years)	1903	911
Adults (85 years and over)	178	69

Median observation time

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

0.90

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

1.01

Data flows and management

Access and validation

Biospecimen access

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

Yes

Biospecimen access conditions

Some centers store the material in biobanks and if necessary it can be reused on request.

Access to subject details

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

Yes

Description of data collection

Data holders: Czech Pneumological and Phthiseological Society, Czech Medical Association of J. E. Purkyně.

Most data is extracted and processed from Electronic Health Record Systems

(EHRS) using the technology developed by OAKS.

The CLADE-IS web application https://lucas.data-warehouse.zone is used for manual data entry.

Data from EHRS and CLADE-IS are then compiled into the source file to the AWS S3 bucket, which is then imported into the OAKS database (PostgreSQL).

Event triggering registration

Event triggering registration of a person in the data source
Disease diagnosis

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

Event triggering de-registration of a person in the data source, other
The subject is discharged from an specialist outpatient clinic or reject consent.

Event triggering creation of a record in the data source
Confirmed diagnosis of ICD-10 C34

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

Linked data sources

Pre linked

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

Yes

Data source, other

other

Linkage strategy

Deterministic

Linkage variable

LUCAS-ID unique identifier for linking data sources.

Linkage completeness

Our process ensure, that the combined data provides a complete and consistent representation of the information. Different data sources have been accurately connected, resulting in a reliable and consistent integrated dataset.

Data management specifications that apply for the data source

Data source refresh

Monthly

Informed consent for use of data for research

Required for general use

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

14/04/2023

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

1,00 month

Data source ETL status

Completed