Validity of using prescription records to identify patient comorbidities

First published: 07/11/2024

Last updated: 07/11/2024

Study Planned

Administrative details

EU PAS number

EUPAS1000000359

Study ID

1000000359

DARWIN EU® study

No

Study countries

United Kingdom

Study description

The aim of this study is to assess the validity of using prescription records for identifying several patient comorbidities in routinely collected electronic health records and to identify predictors for misclassification. The study design is a

retrospective data linkage study, following the design of an evaluation of medical tests for classification and prediction.

Study status

Planned

Research institutions and networks

Institutions

Queen's University Belfast

First published: 01/02/2024

Last updated: 01/02/2024

Institution | Educational Institution

Contact details

Study institution contact

Christian Schnier c.schnier@qub.ac.uk

Study contact

c.schnier@qub.ac.uk

Primary lead investigator

Liam Heaney 0000-0002-9176-5564

Primary lead investigator

ORCID number:

0000-0002-9176-5564

Study timelines

Date when funding contract was signed
Actual: 23/05/2023

Study start date
Planned: 01/11/2024

Data analysis start date
Planned: 01/11/2024

Date of final study report
Planned: 01/05/2025

Sources of funding

- Non-EU institutional research programme

More details on funding

Health Data Research UK (HDR-UK, https://www.hdruk.ac.uk/)

Study protocol

schnier et al Validation.pdf (167.79 KB)

Regulatory

Was the study required by a regulatory body?

No

Is the study required by a Risk Management Plan (RMP)?

Not applicable

Methodological aspects

Study type

Study type list

Study topic:

Other

Study topic, other:

Validation study in electronic health data research

Study type:

Non-interventional study

Scope of the study:

Validation of study variables (exposure outcome covariate)

Data collection methods:

Secondary use of data

Study design:

The study design is a retrospective data linkage study, following the design of an evaluation of medical tests for classification and prediction.

Main study objective:
The primary aim of this study is to assess the validity of using prescription records for identifying patient comorbidities. We additionally identify predictors for misclassification.

Study Design

Non-interventional study design
Cross-sectional

Study drug and medical condition

Medicinal product name, other
Several combinations of drugs indicative of a comorbidity

Additional medical condition(s)
Several comorbidities that are common, non-communicable, are pharmacological treated and where medications are reasonably specific to treating the condition.

Population studied

Short description of the study population
Adult patients of UK GP practices who, between 2004 and 2024, have been registered with a GP practice that contributes to Optimum Patient Care

Research Database (https://opcrd.optimumpatientcare.org/).

Age groups

- **Adult and elderly population (≥18 years)**
  - Adults (18 to < 65 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)

Study design details

Setting

A random subset of the study population is drawn from the OPCRD database. For each patient in the subset, a random index date is created between 2004 and 2024 that allows a 3-year follow-up (follow-up period). For the follow-up period, all medical records are interrogated for a diagnosis indicative of a specific disease (gold standard). For the same period, all prescription records are interrogated for a prescription record indicative of the same disease (index test).

Outcomes

Sensitivity, Specificity, Predictive Values and Likelihood Ratios will be calculated for each disease. To identify predictors for misclassification, test accuracy is calculated for different strata of the study population.

Data analysis plan

The probability of misclassification (positive predictive value and negative predictive value) conditional on potential predictors (covariates) is calculated using multilevel logistic regression models. Observations from patients of the same GP practice are most likely correlated, therefore the practice identifier will be included in any model as random effect.
In sensitivity analysis, we will potentially include (a) a shorter study period (1 year); (b) compare alternative code lists (e.g., only include codes that are indicative of a severe case); (c) analyze time-to-prescription in a survival analysis; (d) modify the definition of a test positive record (e.g., 2 or more prescriptions).
Data management and analysis will be conducted using Microsoft SQL Server and Stata (V18), respectively.

Data management

ENCePP Seal

The use of the ENCePP Seal has been discontinued since February 2025. The ENCePP Seal fields are retained in the display mode for transparency but are no longer maintained.

Data sources

Data source(s)

Optimum Patient Care Research Database

Data sources (types)

Administrative healthcare records (e.g., claims)

Drug dispensing/prescription data

Use of a Common Data Model (CDM)

CDM mapping

No

Data quality specifications

Check conformance

Unknown

Check completeness

Unknown

Check stability

Unknown

Check logical consistency

Unknown

Data characterisation

Data characterisation conducted

No