Development of a predictive model algorithm to identify patients with hypophosphatasia, using Optimum Patient Care Record Database in United Kingdom

First published: 10/11/2021

Last updated: 21/02/2024

Study Finalised

Administrative details

EU PAS number

EUPAS44118

Study ID

47750

DARWIN EU® study

No

Study countries

United Kingdom

Study description

The study is a retrospective observational case-control study accessing de-identified primary healthcare records from patients enrolled in the Optimum Patient Care Record Database in the UK. The study observation period will start at 1st January 2000 and end on 31st March 2021. HPP cases will be identified based on Read or Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) codes, with index date defined as date of first HPP diagnosis during the study period. Controls will be a random selection of non-HPP patients matched by year of birth/age, gender, date of earliest record of index case and being alive at index date. Controls will be collected with a target ratio of 1 case to 20,000 controls. The pooled cases and controls will be randomly allocated to a (i) training (75%), or (ii) validating dataset (25%). For patients' electronic health records respectively in the training and validating datasets, predictor variables will include all available data items as Read or SNOMED-CT codes recorded any time prior to index date. A machine learning prediction model (the "scoring algorithm") will be developed in the training dataset and will then be tested in the validating set using statistical measures of accuracy, discrimination, and calibration. The validation step will involve estimating the predicted probability of HPP diagnosis for each control patient and rank-ordering of patients according to their predicted probabilities ("score"). As a next step, at least two clinical experts in HPP will perform a chart review of the top 10% ranked patients and score patients as 'highly likely HPP', 'likely HPP', 'unlikely HPP', 'highly unlikely HPP', 'not HPP', or 'unable to assess'. Based on this clinical assessment, a threshold for possible or likely HPP will be defined, and the scoring algorithm will be determined.

Study status

Finalised

Research institutions and networks

Institutions

OPEN Health

First published: 01/02/2024

Last updated: 01/02/2024

Institution

Contact details

Study institution contact

Heaton Dave Dave.Heaton@harveywalsh.co.uk

Study contact

Dave.Heaton@harveywalsh.co.uk

Primary lead investigator

Price David

Primary lead investigator

Study timelines

Date when funding contract was signed

Actual: 13/02/2021

Study start date

Planned: 17/12/2021

Actual: 21/12/2021

Data analysis start date
Planned: 17/12/2021

Actual: 04/01/2022

Date of interim report, if expected
Planned: 21/12/2021

Actual: 27/01/2022

Date of final study report
Planned: 30/12/2022

Actual: 22/07/2022

Sources of funding

- Pharmaceutical company and other private sector

More details on funding

Alexion Pharmaceuticals

Study protocol

Protocol_HPP Patient Identification_FINAL V2.0_09NOV21.pdf (1004.84 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:

Disease /health condition

Study type:

Non-interventional study

Scope of the study:

Disease epidemiology

Data collection methods:

Secondary use of data

Main study objective:

To develop and validate a scoring algorithm to aid in the diagnosis of HPP by primary care physicians, incorporating symptoms and risk factors recorded in patients’ primary care electronic health records prior to HPP diagnosis.

Study Design

Non-interventional study design

Other

Non-interventional study design, other

Retrospective observational case-control study

Study drug and medical condition

Medical condition to be studied

Hypophosphatasia

Population studied

Short description of the study population

Newly diagnosed hypophosphatasia (HPP) patients identified from the Optimum Patient Care Research Database (OPCRD) for the study period of 1st January 2000 to 31st March 2021 in UK.
Inclusion criteria:
1. All eligible HPP patients newly identified during the study observation period between 1st January 2000 and 31st March 2021 will be included in the study.

Exclusion criteria:
1. Patient aged less <2 years old at first HPP diagnosis will be excluded.

Age groups

- Adolescents (12 to < 18 years)
- Children (2 to < 12 years)
- Adults (18 to < 46 years)

- Adults (46 to < 65 years)
- Adults (65 to < 75 years)
- Adults (75 to < 85 years)
- Adults (85 years and over)

Special population of interest

Other

Special population of interest, other

Patients with hypophosphatasia

Estimated number of subjects

4600

Study design details

Data analysis plan

The data analysis will be performed using Python version 3.10.0 and will include the following stages: - Machine learning - Re-sampling - Machine learning algorithms - Model validation - Model Application - Descriptive analyses - Sensitivity analyses -

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 source(s), other

Optimum Patient Care Research Database (OPCRD)

Data sources (types)

Electronic healthcare records (EHR)

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