Breathlessness Rapid Evaluation and THErapy (BREATHE)

First published: 27/02/2024

Last updated: 27/02/2024

Study Ongoing

Administrative details

EU PAS number

EUPAS1000000018

Study ID

1000000018

DARWIN EU® study

No

Study countries

- ☐ Australia
- ☐ United Kingdom

Study description

To develop an algorithm to predict the probability of different conditions relating to breathlessness based on demographic, presenting symptoms,

observations, diagnostic test results and treatments. The aim of the algorithm is to optimise assessment pathway based on evidence and reduce the time taken for a patient to receive an accurate diagnosis of conditions causing their presenting breathlessness.

Study status

Ongoing

Research institutions and networks

Institutions

The George Institute for Global Health, UNSW Sydney

Networks

Optimum Patient Care (OPC) Network

United Kingdom (Northern Ireland)

First published: 26/09/2015

Last updated: 16/06/2025

Network **ENCePP partner**

Contact details

Study institution contact

Anthony Sunjaya asunjaya@georgeinstitute.org.au

Study contact

asunjaya@georgeinstitute.org.au

Primary lead investigator

Christine Jenkins 0000-0003-2717-5647

Primary lead investigator

ORCID number:

0000-0003-2717-5647

Study timelines

Date when funding contract was signed

Planned: 01/04/2023

Actual: 01/04/2023

Study start date

Planned: 01/12/2023

Actual: 26/01/2024

Date of final study report

Planned: 30/09/2024

Sources of funding

- Other public funding (e.g. hospital or university)

More details on funding

Australian Government Medical Research Future Fund (MRFF)

Regulatory

Was the study required by a regulatory body?

No

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

Non-EU RMP only

Methodological aspects

Study type

Study type list

Study topic:

Disease /health condition

Study type:

Non-interventional study

Scope of the study:

Assessment of risk minimisation measure implementation or effectiveness

Disease epidemiology

Validation of study variables (exposure outcome covariate)

Data collection methods:

Secondary use of data

Study design:

Predictive modelling study.

Main study objective:

To develop an algorithm to predict the probability of different conditions relating to breathlessness based on demographic, presenting symptoms, observations, diagnostic test results and treatments.

Study Design

Non-interventional study design

Cohort

Study drug and medical condition

Medical condition to be studied

Dyspnoea

Chronic obstructive pulmonary disease

Asthma

Interstitial lung disease

Ejection fraction decreased

Anxiety disorder

Additional medical condition(s)

Heart Failure; Breathing Pattern Disorder; Dysfunctional Breathing; Obesity

Population studied

Short description of the study population

People aged 18 years and over, presenting at General Practices with either breathlessness or a diagnosis of a condition that may present with breathlessness in the UK and Australia.

Age groups

- 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)

Estimated number of subjects

7000000

Study design details

Outcomes

Upon reception of the data, a refined list of outcomes will be constructed based on sample size and prevalence within the data population. Causes of breathlessness to be modelled include:

1. Asthma
2. Chronic obstructive pulmonary disease
3. Lung cancer
4. Heart attack
5. Heart failure
6. Pneumonia
7. Pulmonary thrombo-embolism

8. COVID-19
9. Dysfunctional breathing
10. Deconditioning
11. Anxiety

Data analysis plan

A Bayesian Network (BN) predictive model will be developed. The model development has three steps.

1. Structure learning: Structured learning will take place for the total data set in an expert informed iterative manner.

(i) a directed acyclic graph (DAG) model will be developed based on expert knowledge

(ii) a data driven DAG will be developed using different available BN algorithms

Iterations between the expert driven and data driven DAGs will continue until a simple but biologically plausible form is developed.

Multiple BN algorithms will be explored including:

• constraint-based
• score-based and
• hybrid learning

2. Parameter learning: A conditional probability table will be calculated for each symptom in the model. Conditional probabilities of the presence or absence of each symptom will be calculated, based on the presence or absence of all other variables that the symptom is directly or indirectly connected within the DAG.

3. Validation and assessment of the algorithm: The algorithm will be validated:

(i) Internally by using 10-fold cross-validation

(ii) Externally by training the algorithm with data from one country and validate it on the data from another country. For instance, train the algorithm using Australian data (from NPS Medicine Insight) and validate it using Vietnamese data (VCAPS-1).

4. Then, the predictive performance of the algorithm will be assessed through discrimination using Area Under the ROC Curve

Summary results

The result would be a Bayesian Network predictive model to incorporate into a breathlessness electronic decision support system for primary care.

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)

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