AIM RSF Event

AI-based modelling of EHRs of patients with learning disabilities and multiple long-term health conditions

Professor
Georgina Cosma

Dr Rania
Kousovista

Dr Emeka
Abakasanga

Tuesday, 9 April 2024

13:30 - 14:30

ZOOM

www.www.turing.ac.uk

AI for Multiple Long-term Conditions Research Support Facility

SUPPORTED BY NIHR | National Institute for Health and Care Research

Learning disability in the UK

- **1.1 million (2.16%)** adults with a learning disability **in the UK**
- **956,000** adults with a learning disability **in England**
- **54,000** adults with a learning disability **in Wales**
- **31,000** adults with a learning disability **in Northern Ireland**
- **Median age of death is 63 years old**
- Susceptible to developing multiple long term health conditions (MLTC)
- Common associated health conditions include mental health problems, epilepsy, and being underweight or overweight (MENCAP)

Avoidable deaths

- Mencap defines "avoidable deaths" as deaths preventable by good quality healthcare

- In 2013, 38% of learning disabled people died from avoidable causes vs 9% in the general population (Heslop et al. 2013, p. 92)

- In 2023, LeDeR found 42% of learning disabled deaths were avoidable
- Avoidable death odds: 39% mild, 32% moderate, 26% severe, 3% profound learning disabilities

- Barriers to good healthcare: inaccessible transport, unidentified disabilities, staff misunderstanding, misdiagnosis, patient anxiety, poor collaboration and aftercare

DECODE

- Data-driven machinE-learning aided stratification and management of multiple long-term COnditions in adults with intellectual disabilitiEs (DECODE)
- Funded by the NIHR AI for Multiple Long-term Conditions (AIM) Programme.
- This work uses data provided by patients and collected by the NHS as part of their care and support.

Structure of the talk

- SAIL data preparation and description
- Analysing Hospitalisation data and predicting length of hospital stay
- Identifying clusters of LTCs in patients with LD
- Identifying common temporal patterns and LTC trajectories in patients with LD
- Conclusions

SAIL dataset

- What is the SAIL dataset?
- What defines a long-term condition?
- How was it prepared?
- Patient characteristics of the SAIL dataset

SAIL database

- SAIL Database:
  - **Anonymized individual-level data** from hospitals and GPs in Wales
  - **Over 3 million** patients
- Study Population:
  - **14 323 LD** patients aged 18+
  - Welsh residents with GP records and LD diagnosis at study start
- LTC Identification:
  - Identified LTCs in study population
  - **13 069 LD** patients with at least 1 LTC

Individuals 18 or older and alive at study start date
(3,246,171)

↓

Individuals 18 or older, alive, Welsh resident, with GP records and record of **LD** at study start date
(14,323)

↓

Individuals 18 or older, alive, Welsh resident, with GP records, record of **LD** and **at least one LTC** at study start date
(13,069)

EHR dataset preparation

1. Identifying patients with LD

- Individuals have **LD code**
- Individuals **≥ 18 years** and **alive** at study start date
- **Resident** in the respective nations/TREs
- **Registered at a GP** providing data for research between the start date and the end date

2. Long-term conditions list

- **A professional advisory panel** reviewed and formed a consensus on **relevant LTCs for people with LD**
- Some **conditions were combined** following several clinical discussions
- **40 LTCs** were chosen from an initial list of 110 conditions

3. Phenotypes and outcomes

ICD10
READ
codes

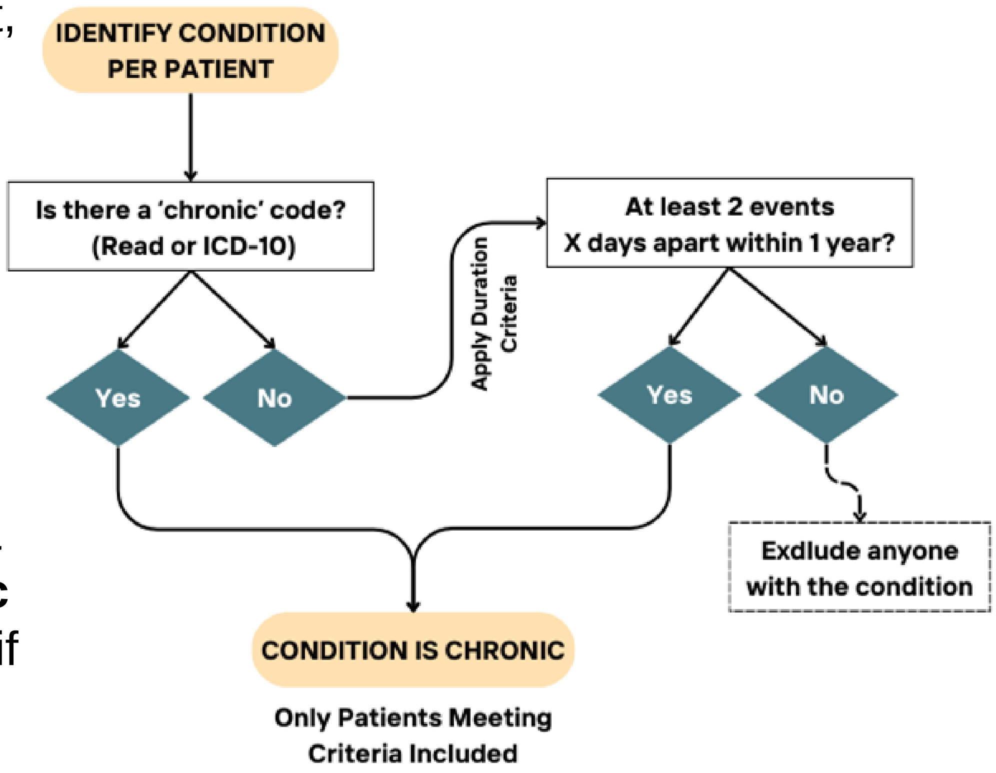
- **Primary care (READ codes):** visits, diagnoses, prescribed medications
- **Secondary care (ICD-10):** hospital admissions, outpatient attendances, diagnoses
- **Social aspects:** age, gender, ethnicity, socioeconomic status, etc.
- **Mortality:** death registrations

4. Data cleaning/harmonisation

- **READ** and **ICD-10** codes reviewed and updated
- **Merged** primary (GP) and secondary (hospital) care tables
- A specific **health condition variable** was created that the code refers to
- Applied several **duration criteria** to define chronic conditions

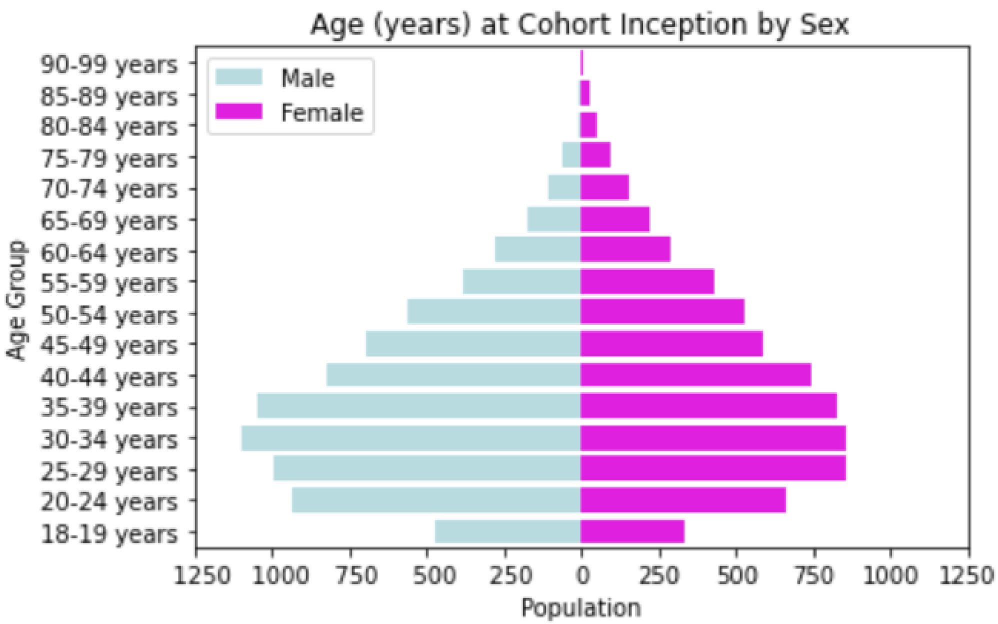
What defines a long-term condition?

- **Standard definition:** a long-term condition is a condition that cannot, at present, be cured but is controlled by medication and/or other treatment/therapies.
- **Examples of LTCs** include Diabetes, Hypertension and others.
- For conditions like **constipation, diarrhoea, back pain**, and **pneumonia** that **do not always fall** into the **chronic (or long-term)** category, we determine if a condition is considered long-term or chronic assessing **its duration**.



Patient characteristics

SAIL database (Wales)



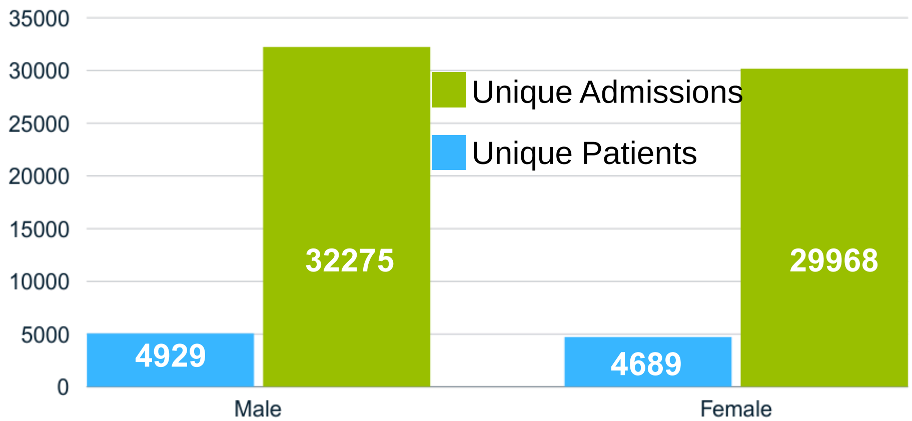
Characteristics	N (%)
Patients	13,069 (100)
Gender	
Female	6,239 (47.7)
Male	6,830 (52.2)
Age	
<20	684 (5.2)
20-29	2221 (16.9)
30-39	3156 (24.4)
40-49	3096 (23.6)
50-59	2157 (16.5)
60-69	1116 (8.5)
70-79	502 (3.8)
80+	137 (1.0)
Ethnicity	
White	9161 (70.0)
Asian	179 (1.3)
Black	43 (0.3)
Mixed/Other	48 (0.3)
Unknown	3638 (27.8)
Welsh Index of Multiple Deprivation	
1 (most deprived)	3195 (28.6)
2	2582 (23.1)
3	2087 (18.7)
4	1949 (17.4)
5 (least deprived)	1346 (12.0)

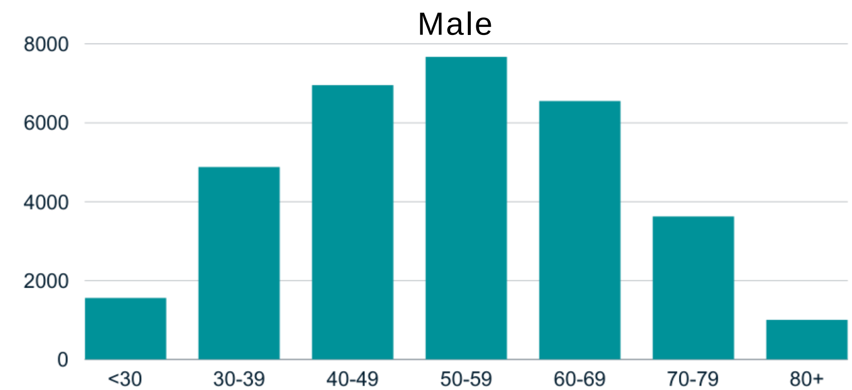
Hospitalisation

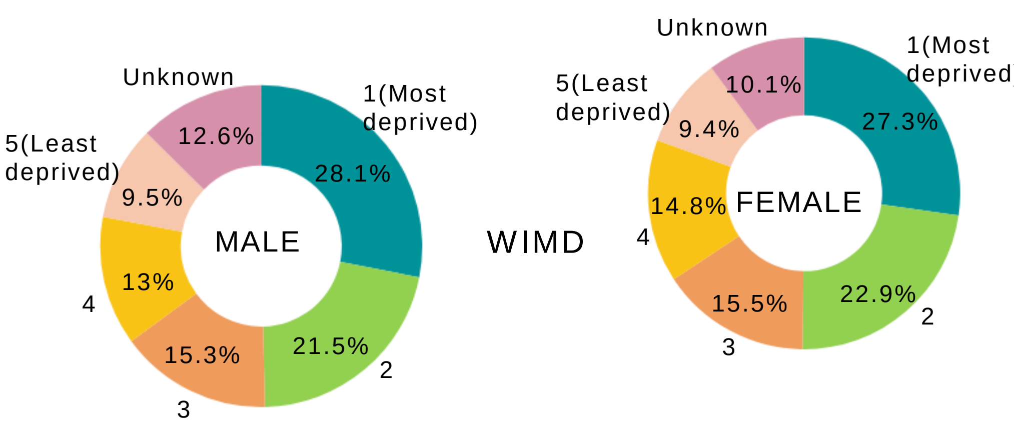
- Demographic analysis of hospitalised Welsh patient with LD from SAIL database.
- Identification of common and top primary conditions treated during hospitalisation.
- Identification of prevalent LTCs linked to prolonged hospital stays.
- Development and evaluation of machine learning models for binary prediction of length of hospital stays, using patient data available up to the first 24 hours of admission.
- Assessment of model performance differences across sensitive groups. Application and comparison of bias mitigation algorithms for equitable prediction

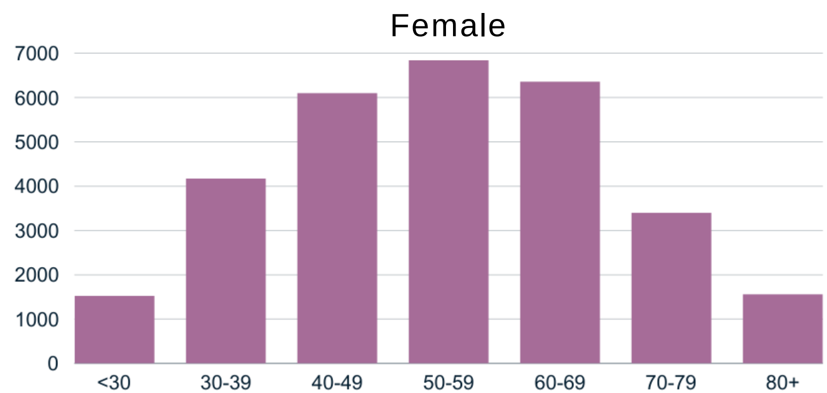
Hospitalisation

Between 2000 and 2021



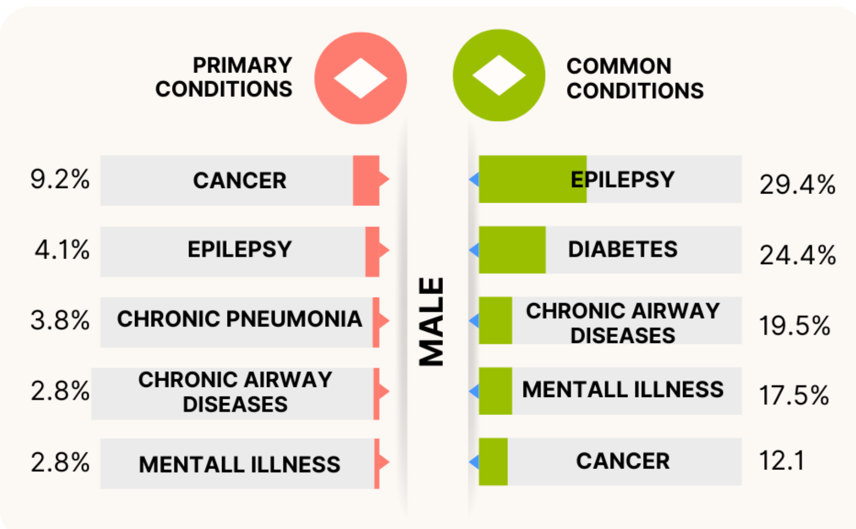


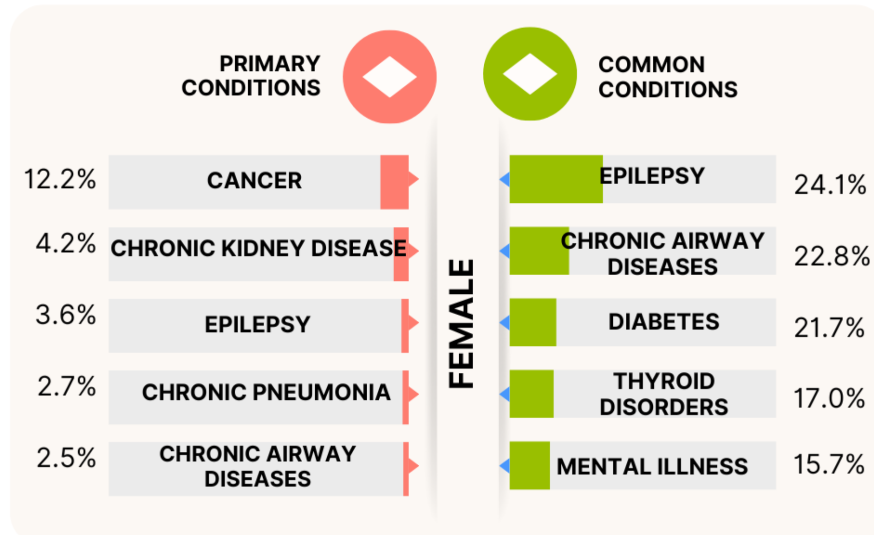




Hospitalisation

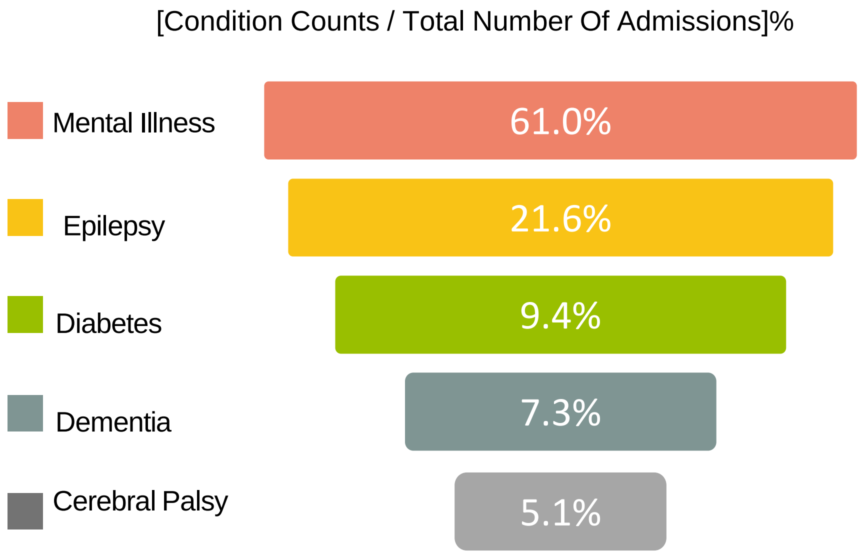
Common and primary LTCs treated between 2011-2021

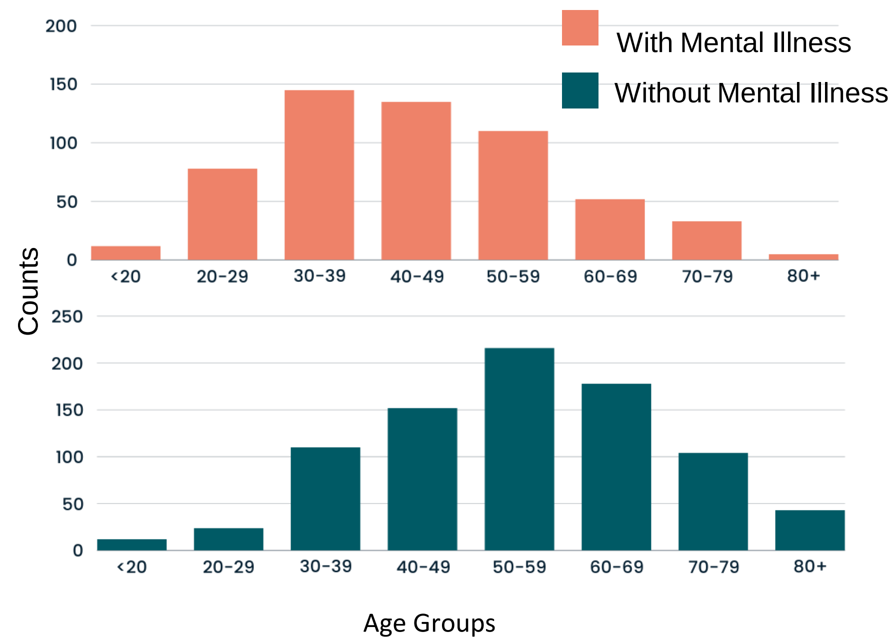




Hospitalisation

Long Stays >= 129 days





This analysis was done on all hospitalisations from birth for patients with LD within the extracted cohort

Hospitalisation

- Premature discharge
- Patients stuck in hospital
- Need to manage patients' hospitalisations

UNIVERSITY OF BIRMINGHAM

Home › News

Search

Report calls for more ordinary lives for people with learning disabilities stuck in hospital

Report into the challenges faced by people with learning disabilities and/or autistic people trying to leave 'long-stay' hospital published in full today.

Published 6 February 2024 · 4 min read

Share

Support the Guardian

Fund independent journalism with £10 per month

Support us →

The Guardian

Newspaper of the year

News | Opinion | Sport | Culture | Lifestyle | More

UK | World | Climate crisis | Ukraine | Football | Newsletters | Business | Environment | UK politics | Education | Society | Science | Tech | Global development | Obituaries

Mental health

This article is more than 2 months old

Thousands of mental health patients readmitted within a month in England

Experts say being discharged prematurely can be 'disastrous', setting back chances of full recovery

Denis Campbell Health policy editor

NHS England

About us | Our w… | …volved | Coronavirus

Search

Urgent and emergency care

Reducing length of stay

Reducing long stays: Empowering patients' campaign

Reducing long stays: Where best next campaign

Urgent treatment centres

About urgent and emergency care

Home › Urgent and emergency care › Reducing length of stay

Reducing length of stay

The Reducing Length of Stay (RLoS) programme aims to provide patients with a better care experience by ensuring they are discharged from hospital without unnecessary delay.

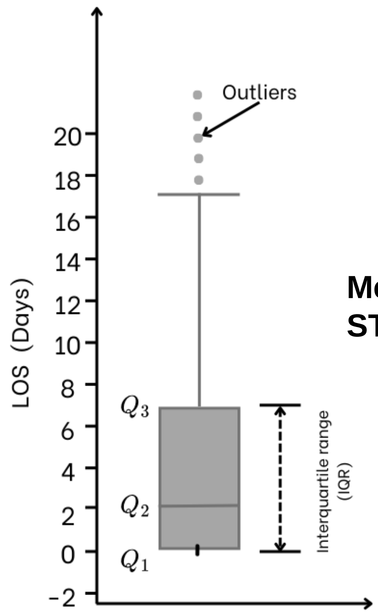
Prolonged stays in hospital are bad for patients, especially for those who are frail or elderly. Spending a long time in hospital can lead to an increased risk of falling, sleep deprivation, catching infections and sometimes mental and physical deconditioning. Despite this, nearly 350,000 patients spend more than three weeks in acute hospitals each year.

Hospitalisation

- In 2023, almost **5,000** people (children and adults) were readmitted to a mental health facility within a month of leaving. [NHS]
- ~**2 000** patients with LD and/or autism in long-stay hospitals have been hospitalised, with over half having spent **> 2 years** in hospital care [University of Birmingham].
  - Includes **350** LD patients admitted for more than a decade
- Highlights the need to proactively manage patient discharges as early as possible during their hospitalisation
- Explore the use of Machine learning models to predict length of hospital stay (LOS), to optimise healthcare resource allocation

Hospitalisation

Predicting Length of Hospital Stay (LOS)



Mean=3.015 days,
STD=$\pm$4.064

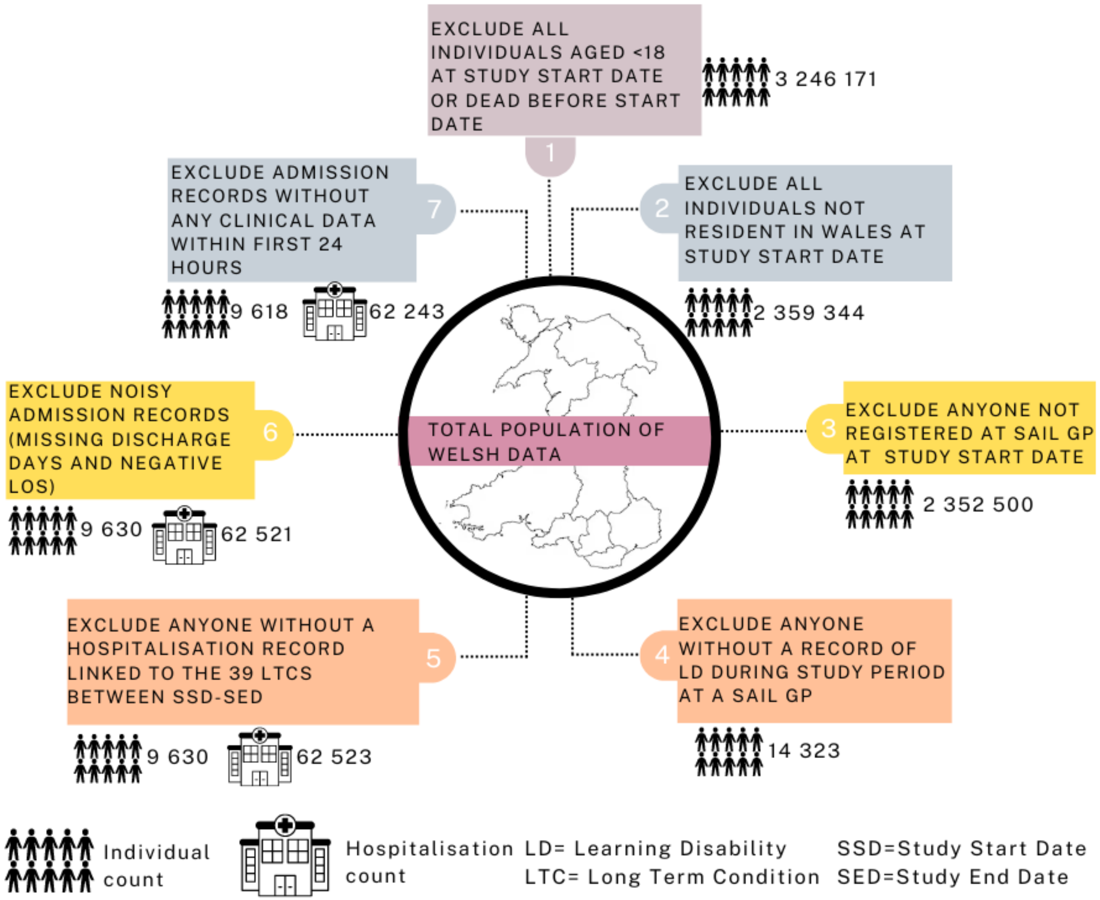
The days range from 0 to >5000 days
0 days indicate outpatient visits

Binary prediction of LOS using ceil(Mean) as threshold

0: LOS < 4days; 1: LOS>= 4days

Hospitalisation

Predicting Length of Hospital Stay (LOS)

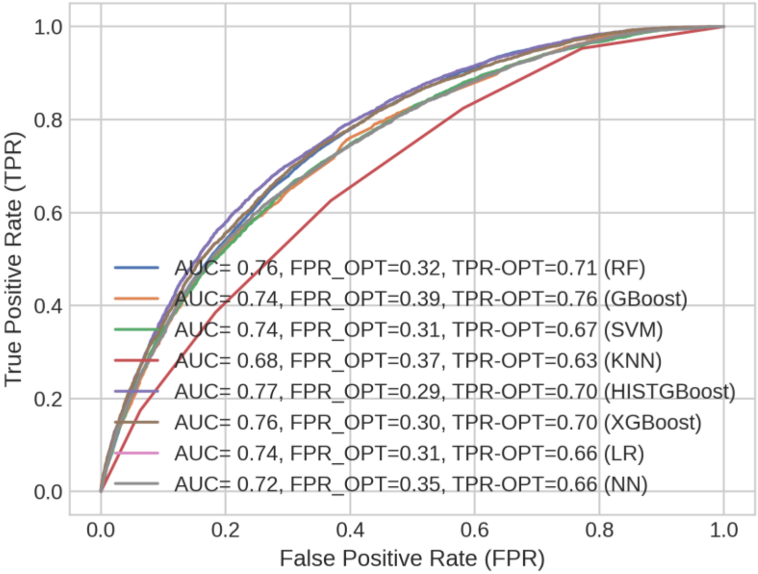


Inclusion/extraction criteria for predicting LOS

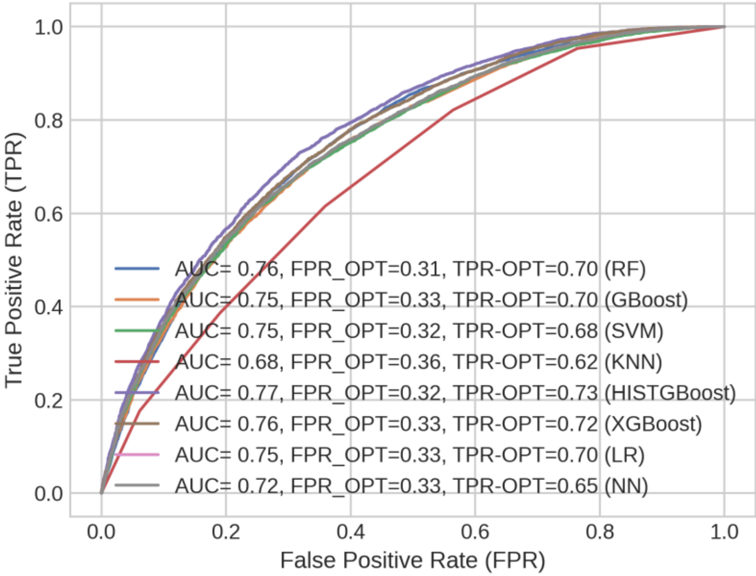
Hospitalisation

Predicting Length of Hospital Stay (LOS)

BMI
SMOKING HISTORY
ALCOHOL HISTORY
PHYSICAL ACTIVITY
AUTISM
NUM PRVADMISSION 1RY
NUM PRVEPISODES 1RY
NUM PRVCOMORBID 1RY
NUM PRVADMISSION 3RY
NUM PRVEPISODES 3RY
NUM PRVCOMORBID 3RY
NUM PRVHOSPITAL DAYS 1YR
NUM PRVHOSPITAL DAYS 3YR
MEDICATIONS
TOTAL COMORBIDITY
NUMEPISODES 24HRS
NUMCOMORBIDITIES 24HRS
CONDITIONS



Male



Female

Hospitalisation

Predicting Length of Hospital Stay (LOS)

Bias analysis and mitigation

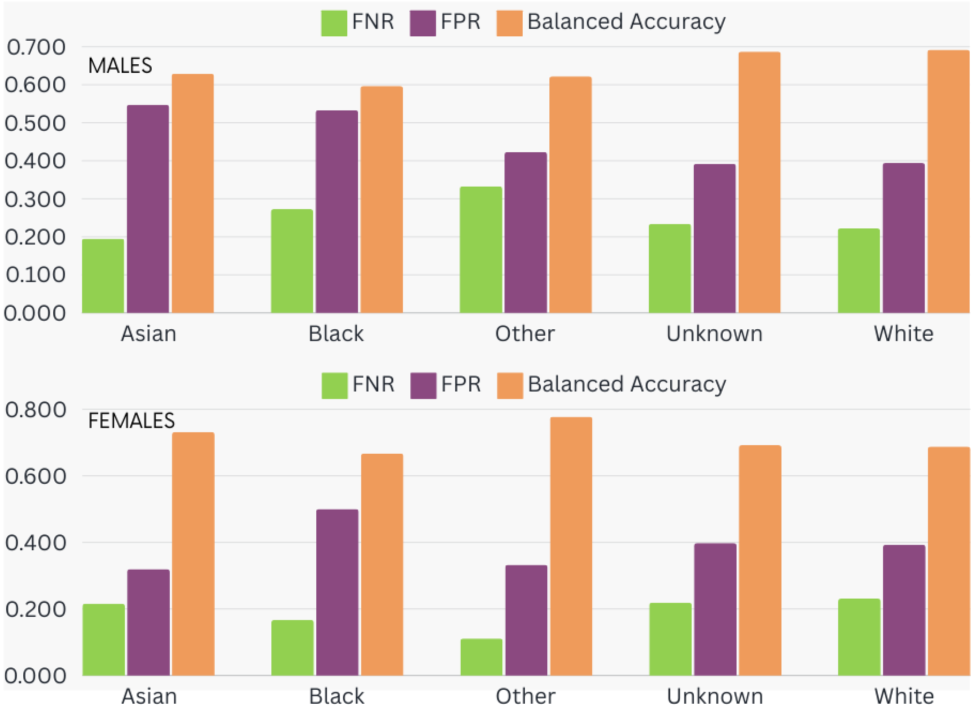
Unmitigated

Male

FNR	FPR	Balanced Accuracy
0.224	0.396	0.690

Female

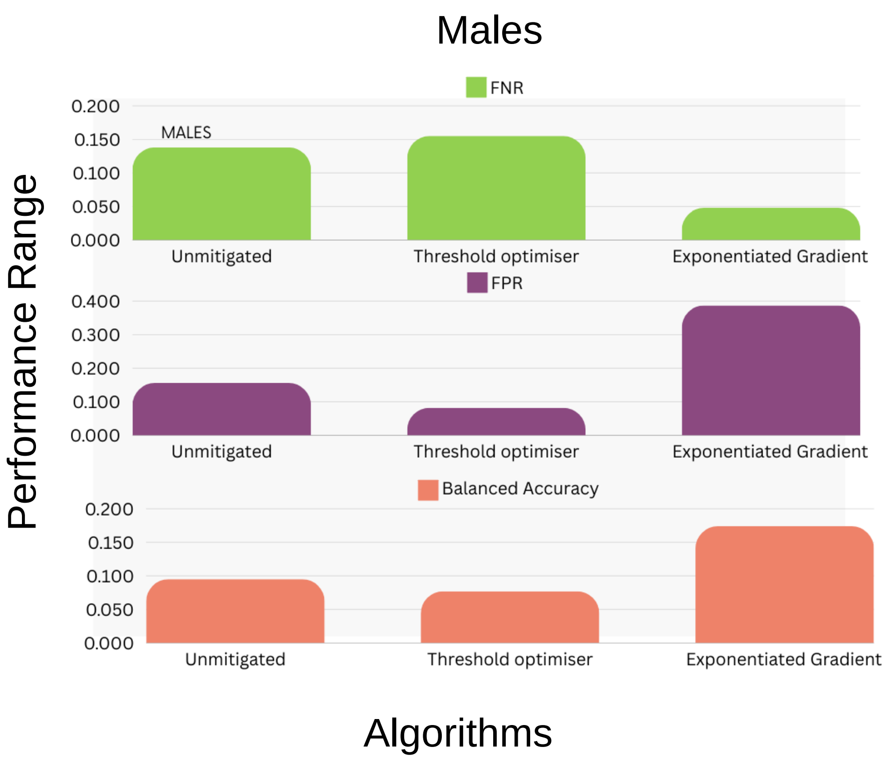
FNR	FPR	Balanced Accuracy
0.229	0.392	0.689

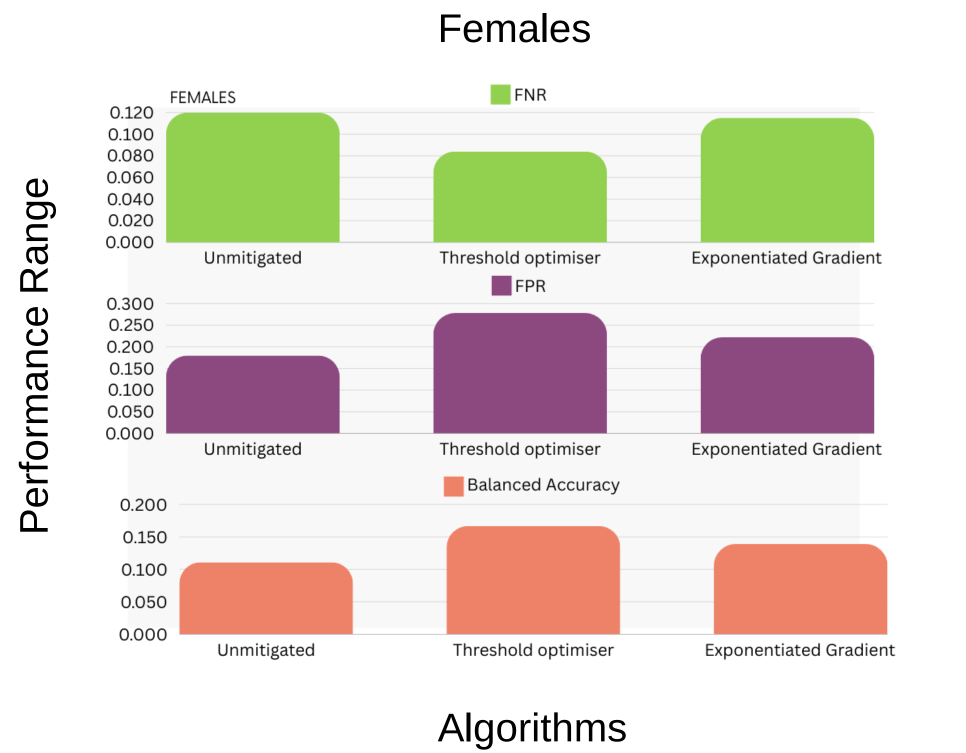


Apply bias mitigation algorithms: a) Threshold optimizer, and b) exponentiated gradient

Hospitalisation

Predicting Length of Hospital Stay (LOS)





Performance Range is the difference between the maximum and minimum values for each metric across ethnic groups

e.g., $Performance\ range_{\{FNR\}} = \max\{\mathrm{FNR}\} - \min\{\mathrm{FNR}\}$

Concluding remarks

- Analysed electronic health records of 9,618 patients with LD and MLTCs in Wales, examining 62,243 hospital admissions.
- Cancer was the top primary condition for hospital admissions in both males and females with LD. Epilepsy was the most commonly co-occurring condition across all admissions between 2011 and 2021.
- Long stays exceeding 129 days were commonly related to mental illness.
- A random forest ML model achieved optimal performance in predicting the LOS using data up to the first 24 hours of admission.
- Two bias mitigation approaches were tested, with the threshold optimizer outperforming the exponentiated gradient approach in minimising some performance discrepancies across groups.

Identifying clusters of LTCs

- Clustering in healthcare helps uncover patterns and structures embedded within vast and intricate healthcare datasets

- Provides insights into homogeneous groups characterised by shared clinical profiles, disease trajectories, or treatment responses.

- These patient clusters inform various facets of healthcare research
  - disease stratification, personalised medicine, predictive modelling, and healthcare resource allocation.

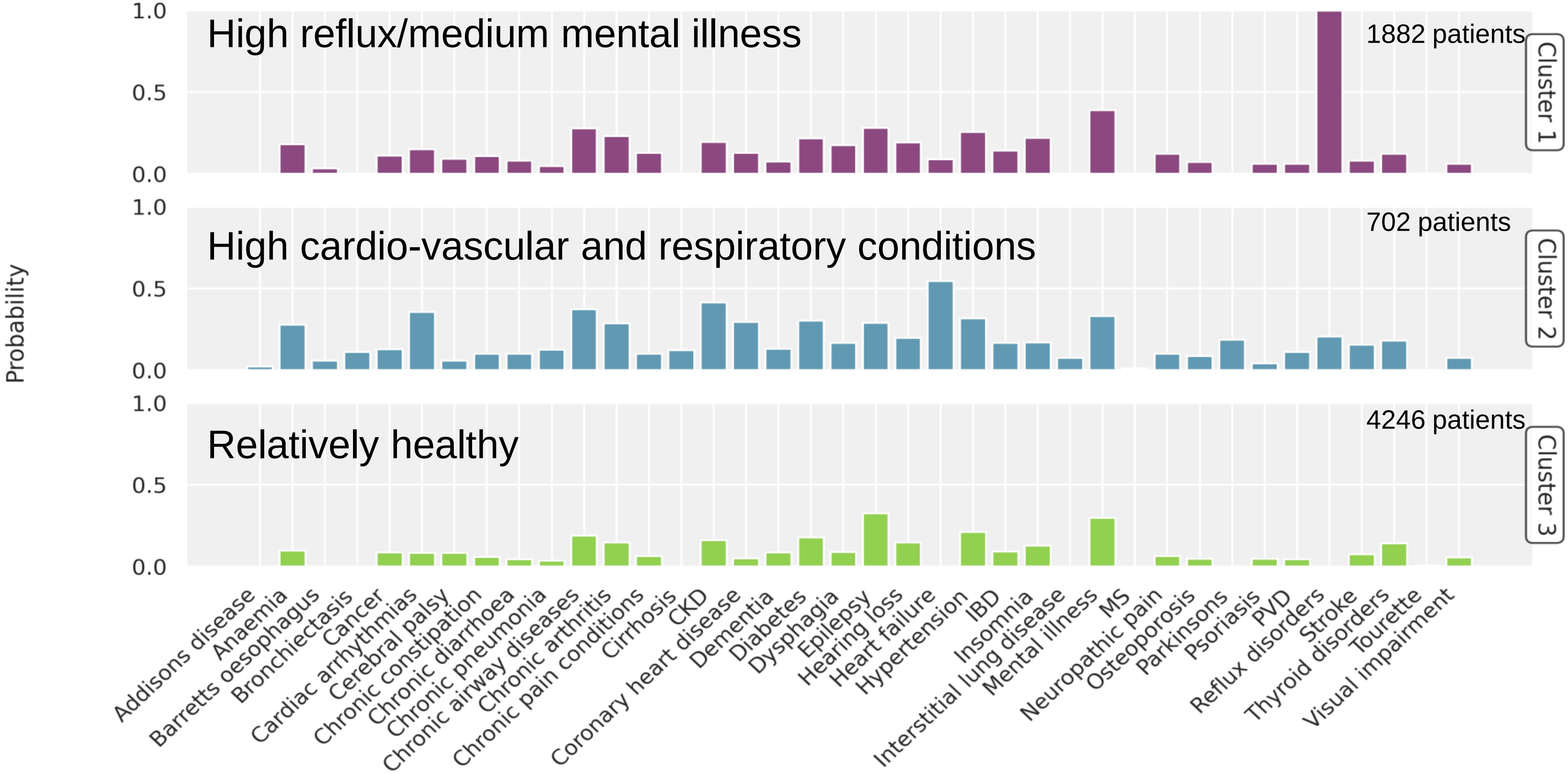
- Clustering has advanced diagnosis, symptom management, and prognosis in complex diseases
  - cancer, psoriasis, diabetes, cardiovascular diseases, and drug discovery and development to manage diseases.

- **There is limited clustering research to analyse patterns of co-occurrence of MLTCs among individuals with LD.**

Clusters of LTCs

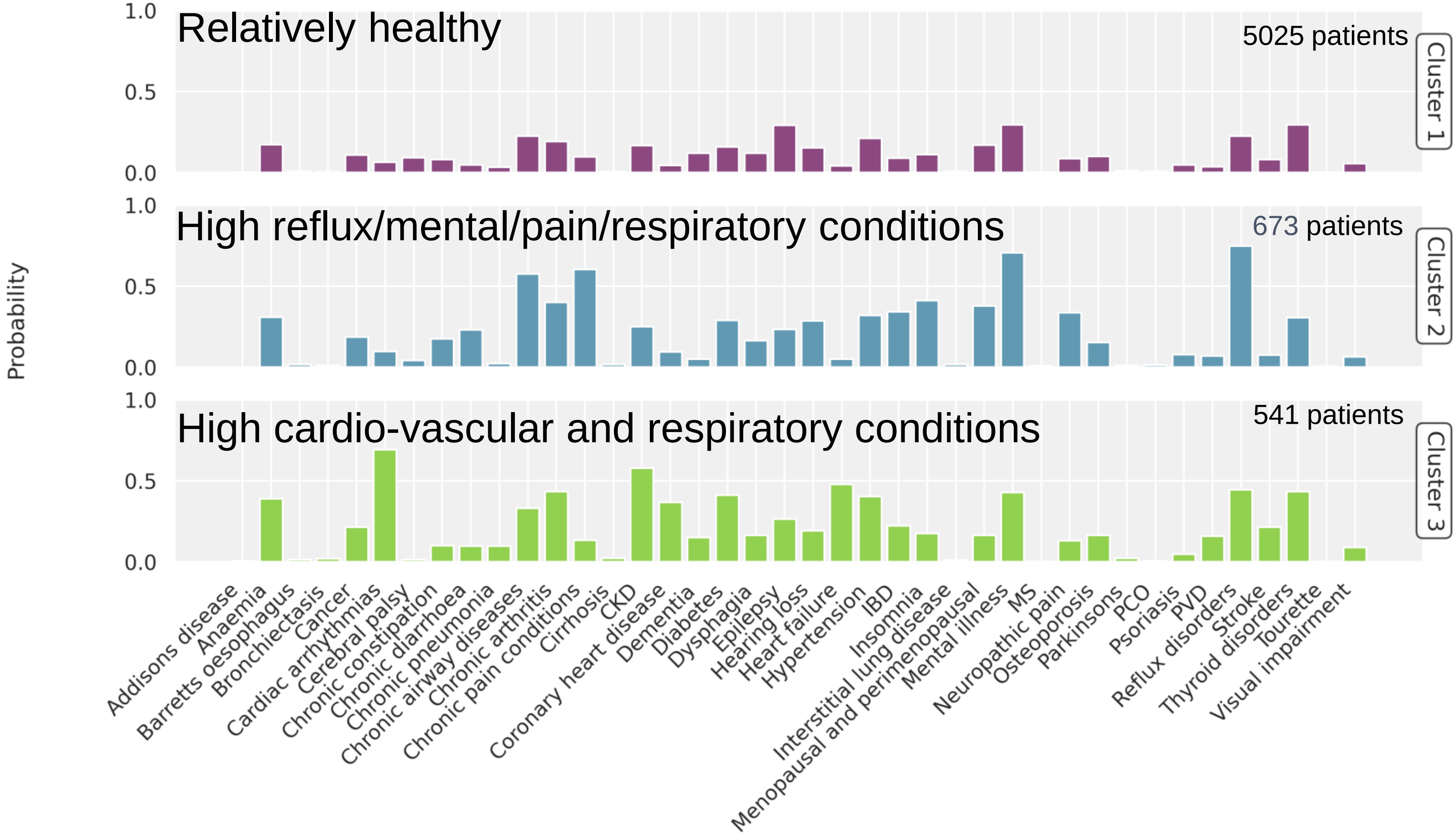
- SAIL dataset: 6,239 Females (47.7%) & 6,830 Males (52.2%)
- Evaluated categorical ML and statistical clustering algorithms:
  - Agglomerative, Birch, Kmeans, Kmodes, Latent Class Analysis (LCA), Gaussian Mixture Models (GMM).
- Obtain optimal number of clusters per algorithm:
  - Average silhouette width, Elbow heuristics on the SSE or BIC,
  - Optimal between 3-4 clusters
- Compare performance of algorithms with respect to Separability
  - Male: 3-cluster GMM
  - Female: 3-cluster Birch

Clusters of LTCs



Male: 3-cluster GMM

Clusters of LTCs



Female: 3-cluster Birch

Risk factors and trajectories

- For each **cluster** identified, we will evaluate **associations** with several **risk factors**, such as:
  - Medications
  - Physical activity
  - BMI and
  - other sociodemographic factors

- Identify significant LTC **pairs** and their **temporal directions**

- Development of **MLTC trajectories**

Why temporal trajectories analysis?

- LTCs frequently **co-occur**, and understanding their **progression** is crucial, especially for LD patients
- Traditional care management may **not capture** complex LTC temporal relationships
- Innovative methods are needed to:
  - Understand LTC **temporal patterns**
  - **Predict outcomes** based on trajectories
  - Tailor care strategies for **patient subgroups**

Aim: To develop LTC trajectories that take into account temporal directionality and uncover patterns among these trajectories using advanced statistical and machine learning techniques, ultimately informing personalised care strategies for patients with learning disabilities.

Identifying temporal patterns in LTCs trajectories

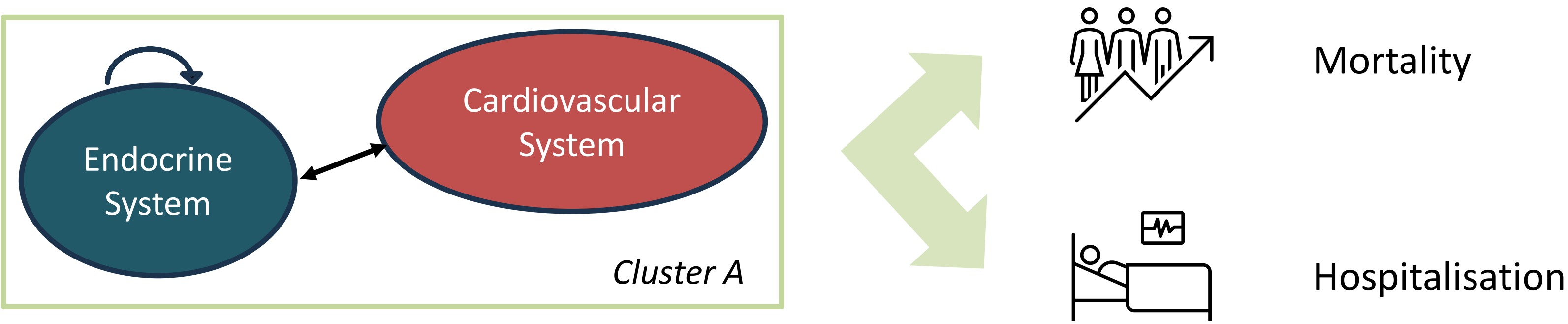
- Initially created trajectories using statistical methods.

Diabetes → Asthma → Hypertension → Arthritis → Retinopathy

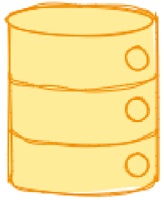
Diabetes → Hypertension → Bronchitis → Retinopathy

Share a pattern

AI-based models to group trajectories by similar diagnostic **patterns they share** over time

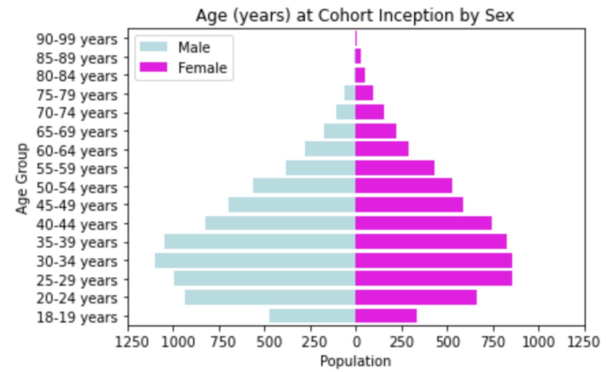


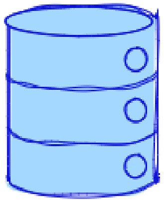
Next steps



SAIL database (Wales)

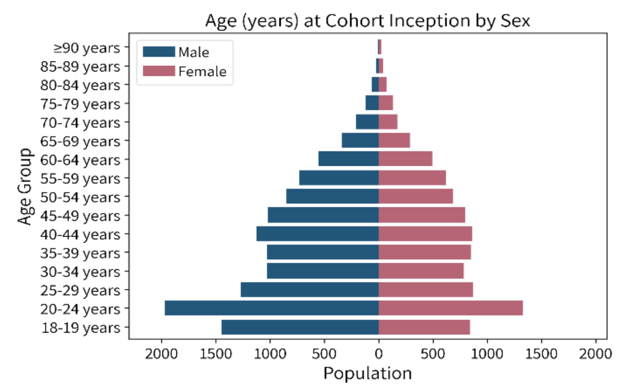
Characteristics	N (%)
Patients	13,069 (100)
Female	6,239 (47.7)
Male	6,830 (52.2)
Mean Age	30-34 years





CPRD dataset (England)

Characteristics	N (%)
Patients	20,646
Female	8,867 (43.0)
Male	11,779 (57.0)
Mean age	35-39 years



Conclusion

- Healthcare providers can understand **disease progression**
  - Given a patient's profile we can identify the trajectories closest to their profile over time.

- Address premature and avoidable **mortality**
  - Greater awareness of the nature of trajectories and their clusters
  - Better model of coordinated care improving the quality of life and longevity

- Reduce avoidable **hospitalisations** and excessive **lengths of stay**

Thank you for listening

Any questions?

Contact: Georgina Cosma
g.cosma@lboro.ac.uk