

UCL Cancer Domain
AI Resources and Methodologies

Use of Large Scale Electronic Health Records in Cancer Research

Dr Arturo González-Izquierdo
UCL Institute of Health Informatics

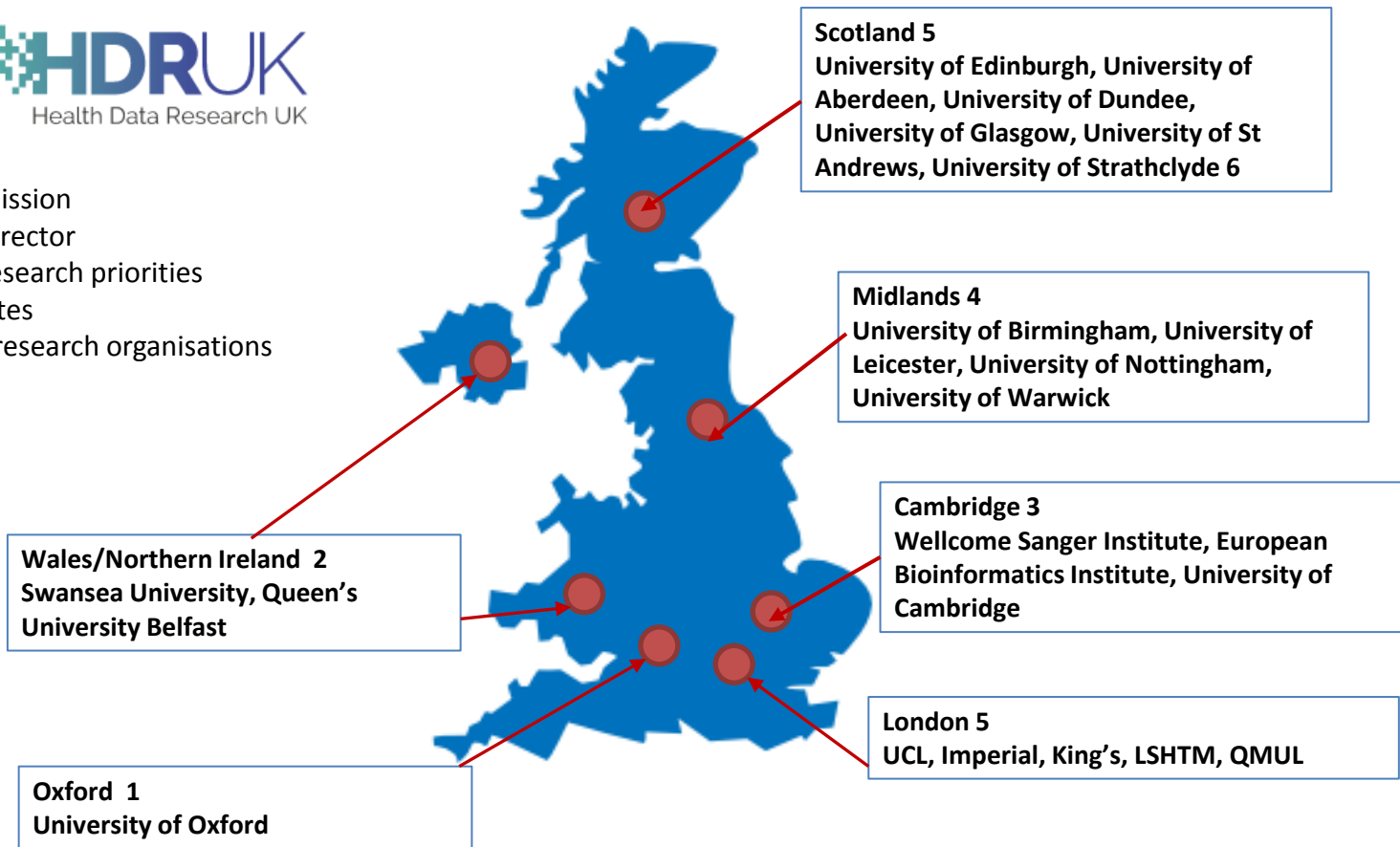
The CALIBER research platform as a resource for cancer research with potential for machine learning applications

UCL Institute of Health Informatics

“We aim to conduct high quality research that leverages large scale databases and health informatics approaches to improve health at local, national and international levels”



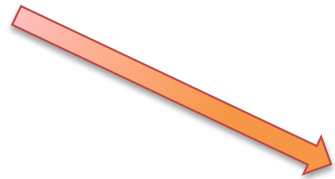
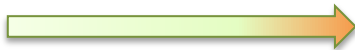
- 1 mission
- 1 director
- 4 research priorities
- 6 sites
- 20 research organisations



Electronic Health Records

Clinical Practice Research Datalink

Hospital Episode Statistics

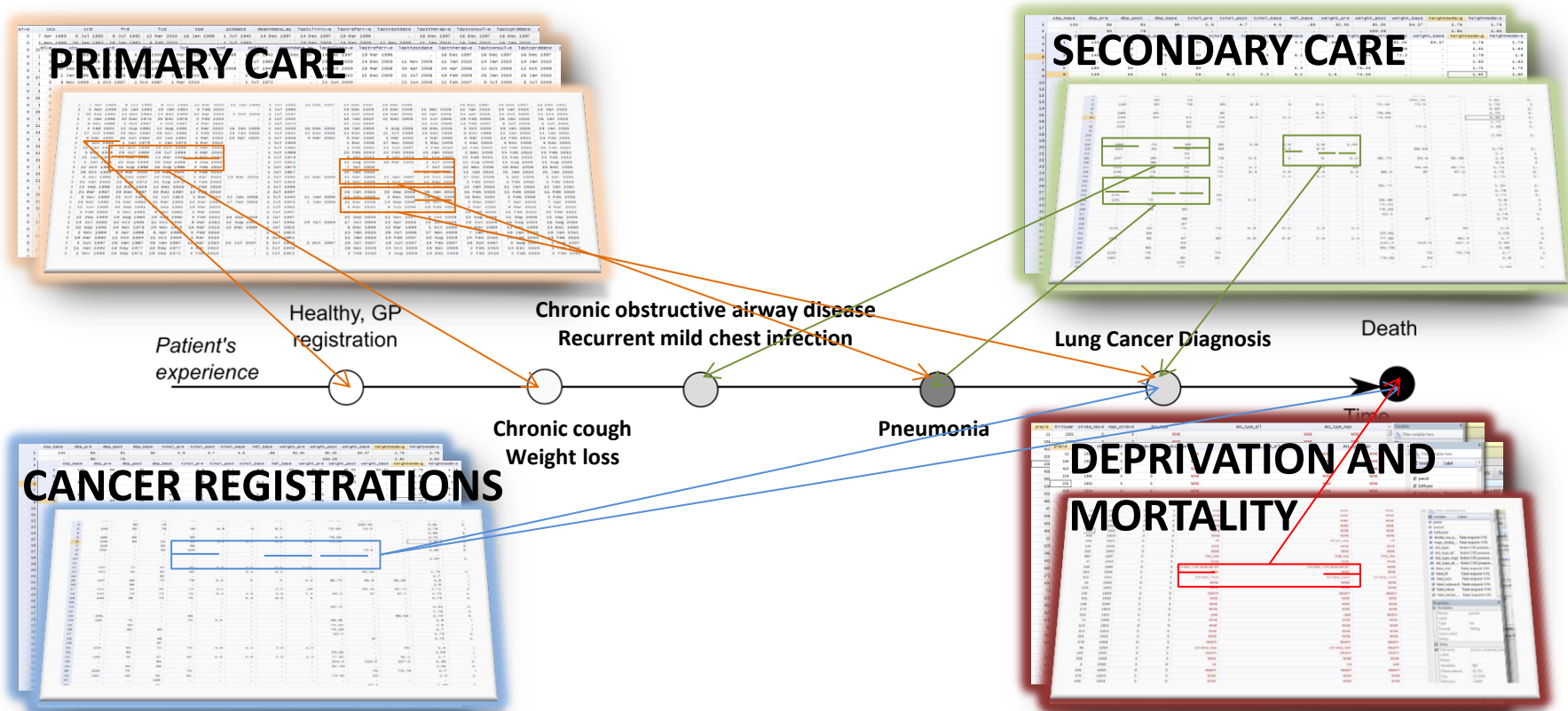


National Cancer Registration and Analysis Service

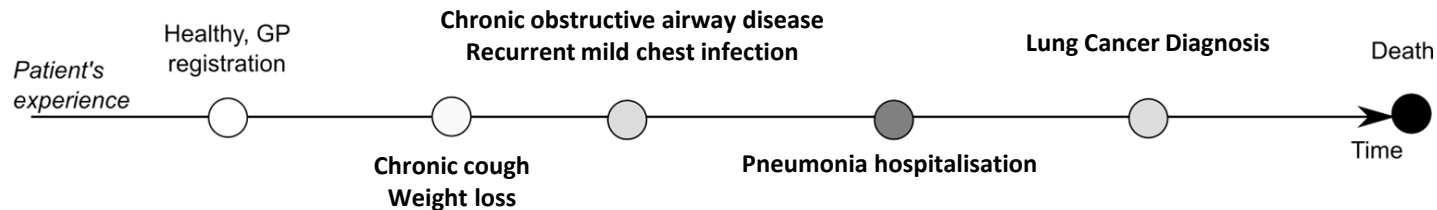
Office for National Statistics



Linked Electronic Health Records

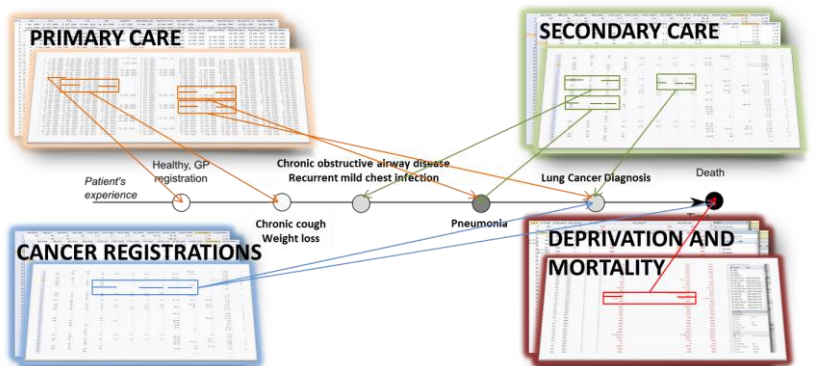


Linked Electronic Health Records



Registration date	Consultation date	Consultation date	Admission date	Admission date	Date of death
Date of birth	Blood tests (routine)	Sputum tests	Diagnosis	Diagnosis (cancer type, stage, metastases)	Underlying cause
Blood pressure	Spirometry	Diagnosis	Additional diagnosis (deep vein thrombosis pneumonia)	Procedures (surgery)	Subsidiary causes
Weight	Chest x-rays	Procedures	Procedure for biopsy of lesion (bronchoscopy, Chest drain)	Treatment (chemotherapy, radiotherapy)	
Height	CT chest		Chest imaging (x-ray, CT, PET CT, chest drain insertion)	Discharge date	
Physical activity	Treatment (antibiotic, Inhalers)		Discharge date		
Health history (heart, diabetes, stroke)					
Smoking					
Alcohol					
Contraception					
Immunisations					

The CALIBER Research Platform



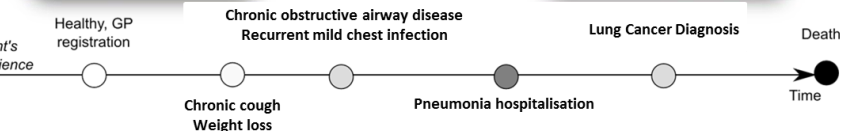
Cohort identification methods

Deep phenotyping algorithms



Longitudinal clinical trajectories

Precise temporal allocation of Exposures and outcomes



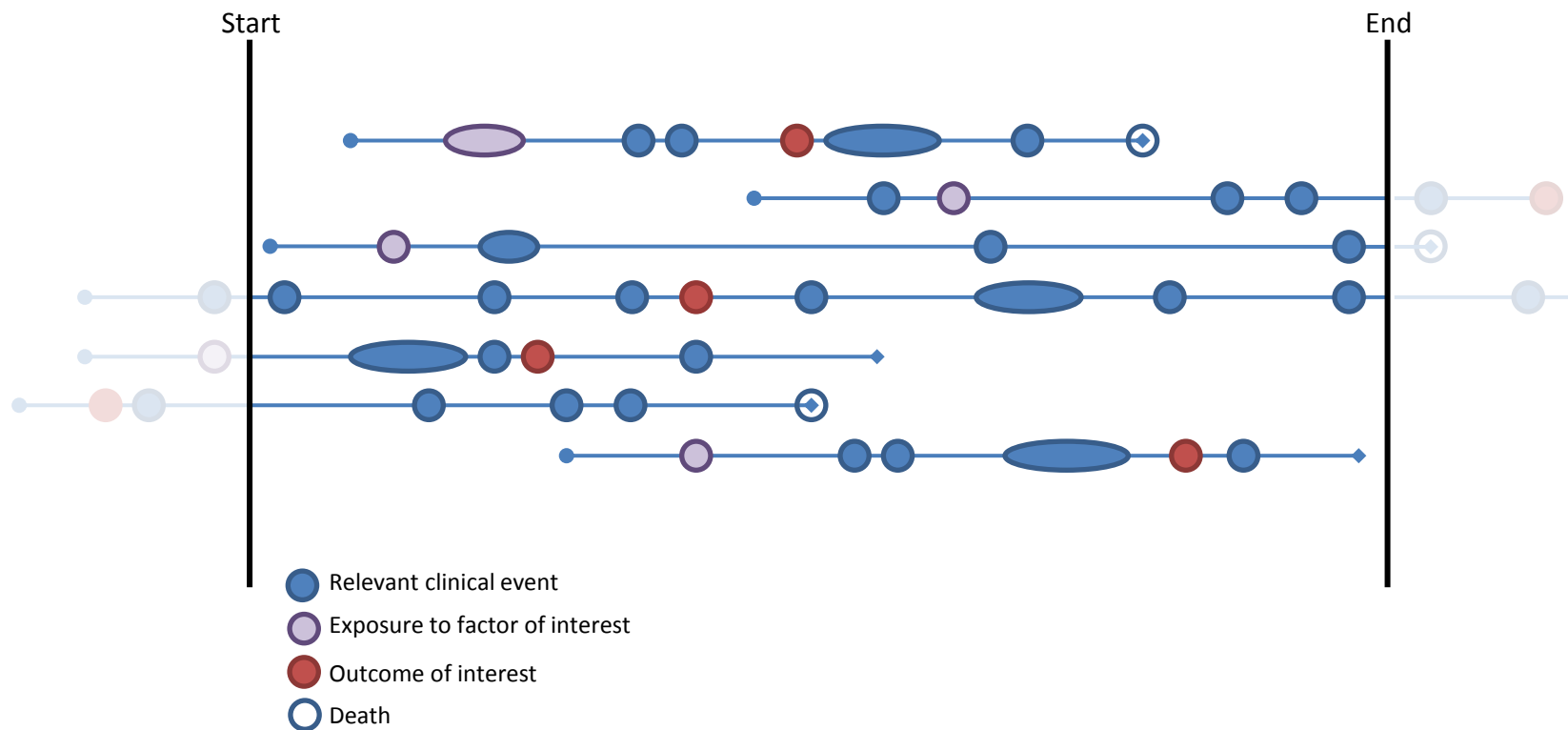
Registration date	Consultation date	Consultation date	Admission date	Admission date	Date of death
Date of birth	Blood tests (routine)	Sputum tests	Diagnosis	Diagnosis (cancer type, stage, metastases)	Underlying cause
Blood pressure	Spirometry	Diagnosis	Additional diagnosis (deep vein thrombosis)	Procedures (surgery)	Subsidiary causes
Weight	Chest x-rays	Procedures	Pneumonia)	Treatment (chemotherapy, radiotherapy)	
Height	CT chest		Procedure for biopsy of lesion (bronchoscopy, Chest drain)	Discharge date	
Physical activity	Treatment (antibiotic, Inhalers)		Chest imaging (x-ray, CT, PET CT, chest drain insertion)		
Health history (heart, diabetes, stroke)			Discharge date		
Smoking					
Alcohol					
Contraception					
Immunisations					

New cancer related research in CALIBER

- The CALIBER Research Platform
 - 5 Newly approved (or awaiting approval) cancer specific projects
 - Disease epidemiology
 - Characterisation and causal inference
 - Incidence and progression of disease
 - Diagnosis and risk of cancer-specific outcomes
 - Healthcare utilisation
 - Drug repurposing

Challenges

EHR's observation window



Opportunities

- Recent willingness by data custodians to research health data using machine learning based methodologies
- Wide range of exploratory or hypothesis generation/test studies
 - Patient classification (Machine Learning sub-phenotyping)
 - Detailed healthcare utilisation patterns (multi-state trajectory flows)
 - Sophisticated epidemiological/statistical methods computationally feasible for causal inference
 - EHR based decision/early detection tools (automation)

Academic members



Spiros Denaxas

Senior Lecturer in Biomedical Informatics



Arturo Gonzalez-Izquierdo

Research Associate in Electronic Health Records



Kenan Direk

Research Associate in Electronic Health Records



Michalis Katsoulis

BHF Research Fellow



Marina Daskalopoulou

Research Associate



Ghazaleh Fatemifar

AHA Research Fellow



Natalie Fitzpatrick

Research Data Coordinator



Maria Pikoula

Clinical Data Scientist



Vaclav Papez

Clinical Data Scientist



Marcos Barreto

Royal Society Newton Fellow



Colin Josephson

Assistant Professor of Neurology and Community Health Sciences (Univ. of Calgary)

Marie Erwood

IMAGINE Data Manager

The Data Lab

Natalie Fitzpatrick
Data Science Facilitator
n.fitzpatrick@ucl.ac.uk

CALIBER portal
<https://www.caliberresearch.org/portal>

Denaxas Lab
<http://denaxaslab.org/>