

# Cohort Discovery and the OMOP Common Data Model in the United Kingdom

Bradley Kirby, Charlotte Weaver, Tom Giles, Emily Jefferson

Health Data Research (HDR) UK



# Health Data Research (HDR) UK

- **Mission:**
  - To unite the UK's health data to enable discoveries that improve people's lives
- **20yr Vision:**
  - For large-scale data and advanced analytics to benefit every patient interaction, clinical trial and biomedical discovery, and to enhance public health
- **Core Objectives:**
  - Accelerate trustworthy use of data
  - Empower researchers
  - Promote partnerships

## The Gateway...

The screenshot shows the Health Data Research Gateway website. The main heading is "Cohort Discovery Tool". Below it, a description reads: "Compare the numbers of relevant records across different datasets: query federated datasets to find out how many records match a specified criteria e.g. the number of females with asthma." A large graphic on the right says "Cohort Discovery". Below this is a navigation menu with options: "Datasets & BioSamples", "Data Uses / Research Projects", "Cohort Discovery" (highlighted), and "Analysis Scripts & Software". Other options include "Publications", "Data Custodians", "Data Custodian Networks", "Collections", "Phenotypes", and "Courses". At the bottom, there are sections for "Support Centre" and "Community".

Metadata descriptions for +850 health datasets

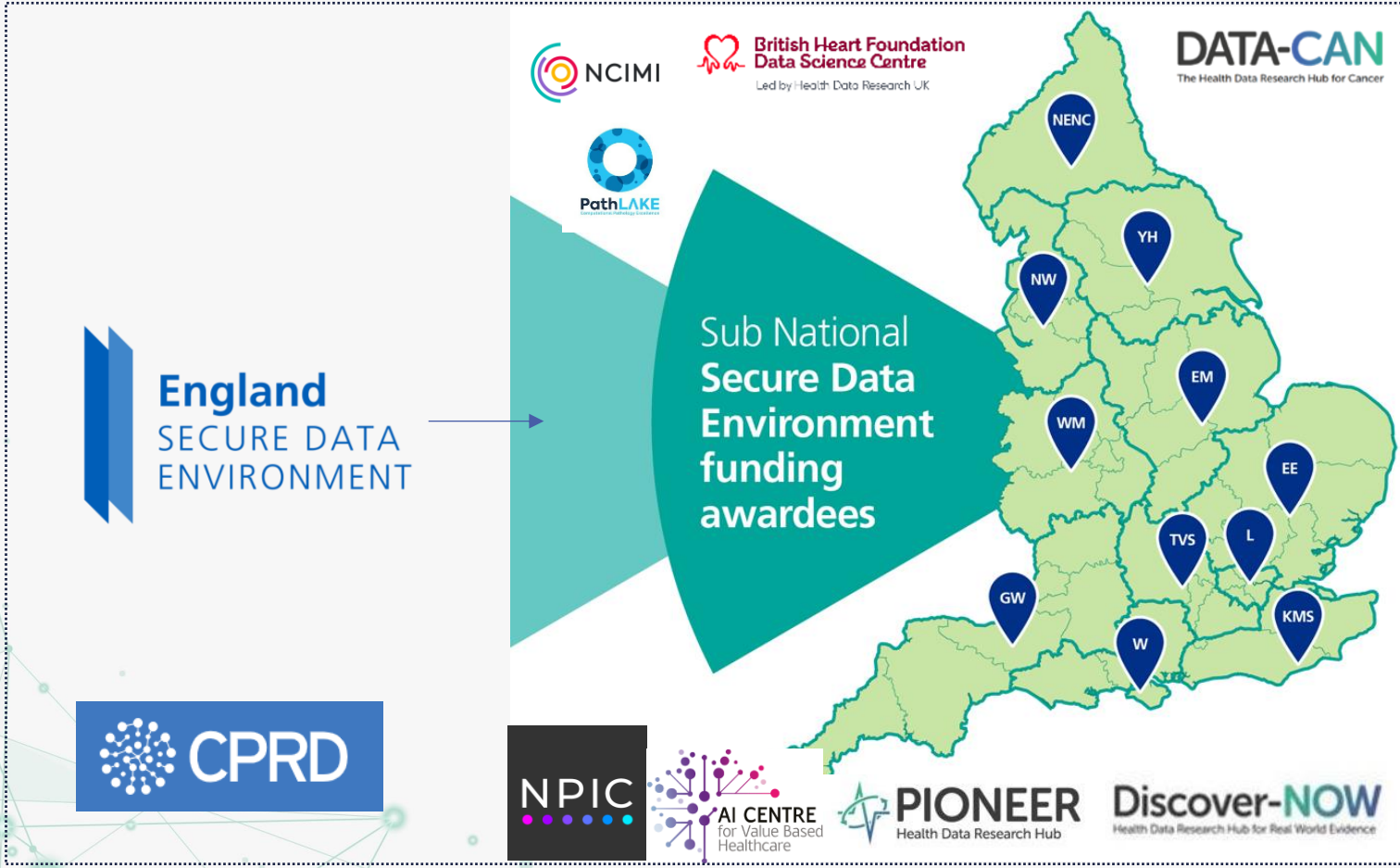
Links to +3,000 health data resources

+3,000 users

# Working in partnership



## England



## Scotland



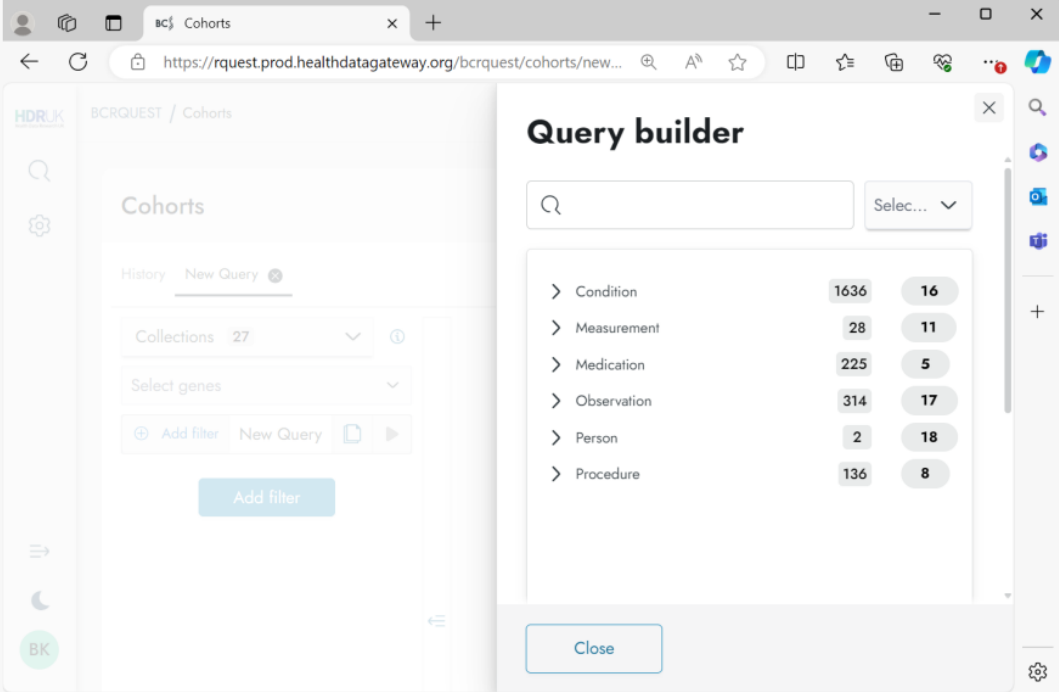
# Standardised OMOP attributes for Cohort Discovery

OMOP Table	OMOP Field Names
Person	person_id, <b>gender_concept_id</b> , year_of_birth, <b>race_concept_id</b>
Condition_Occurrence	condition_occurrence_id, person_id, <b>condition_concept_id</b> , <b>condition_start_date</b>
Procedure_Occurrence	procedure_occurrence_id, person_id, <b>procedure_concept_id</b> , <b>procedure_date</b>
Drug_Exposure	drug_exposure_id, person_id, <b>drug_concept_id</b> , <b>drug_exposure_start_date</b>
Observation	observation_id, person_id, <b>observation_concept_id</b> , <b>observation_date</b> , value_as_number(*), value_as_concept_id(*)
Measurement	measurement_id, person_id, <b>measurement_concept_id</b> , <b>measurement_date</b> , <b>value_as_number</b> , <b>value_as_concept_id</b>
Death	person_id, death_date, <b>death_type_concept_id</b> , <b>cause_concept_id</b>
Visit_Occurrence	visit_occurrence_id(*), person_id(*), visit_concept_id(*), visit_start_date(*)

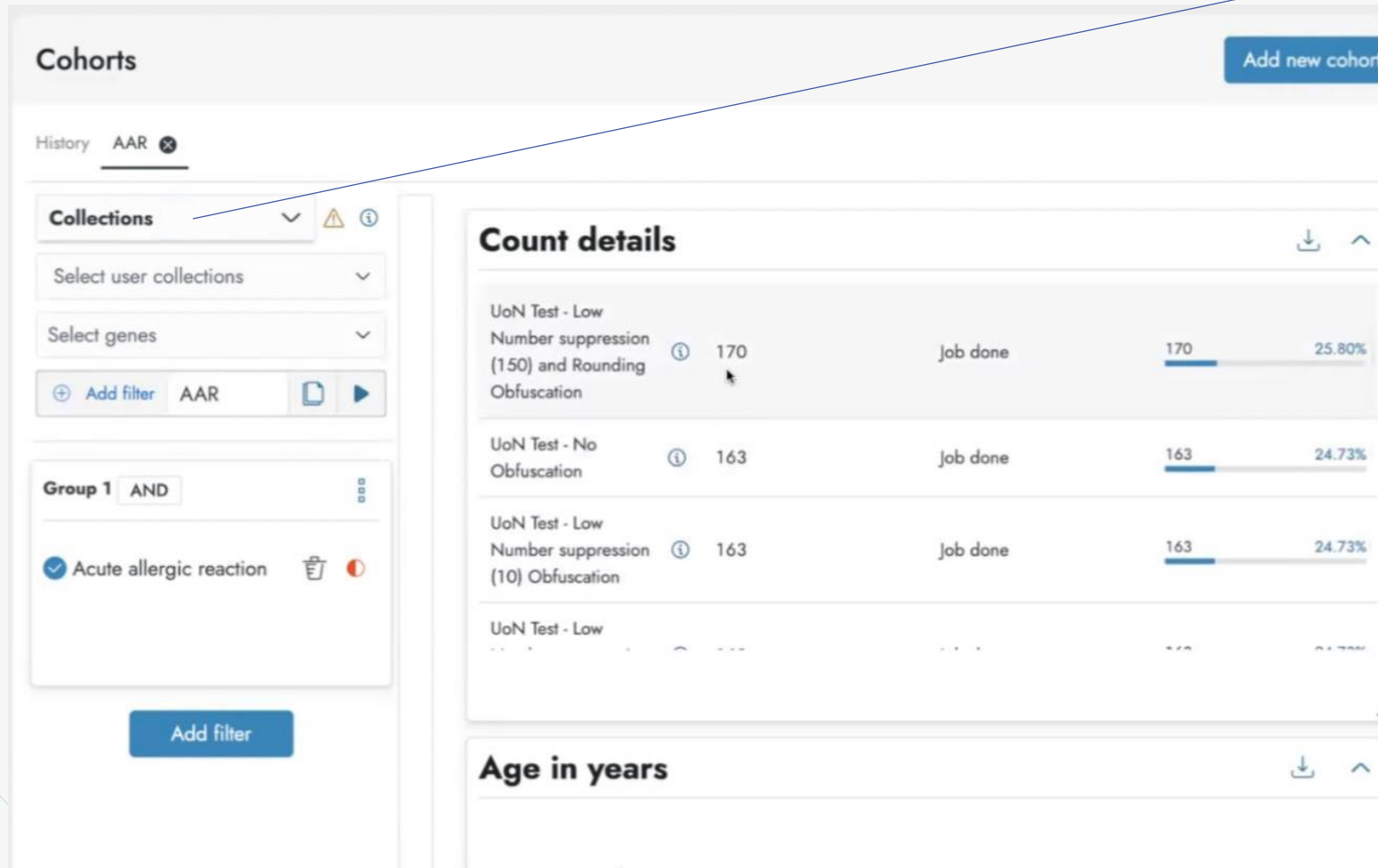
Fields in bold are the core content of interest for analysis. The remaining non-bold fields are required for linking tables together.

\* Optional, but useful

## Cohort Discovery Query Builder



# Cohort Discovery Query Results



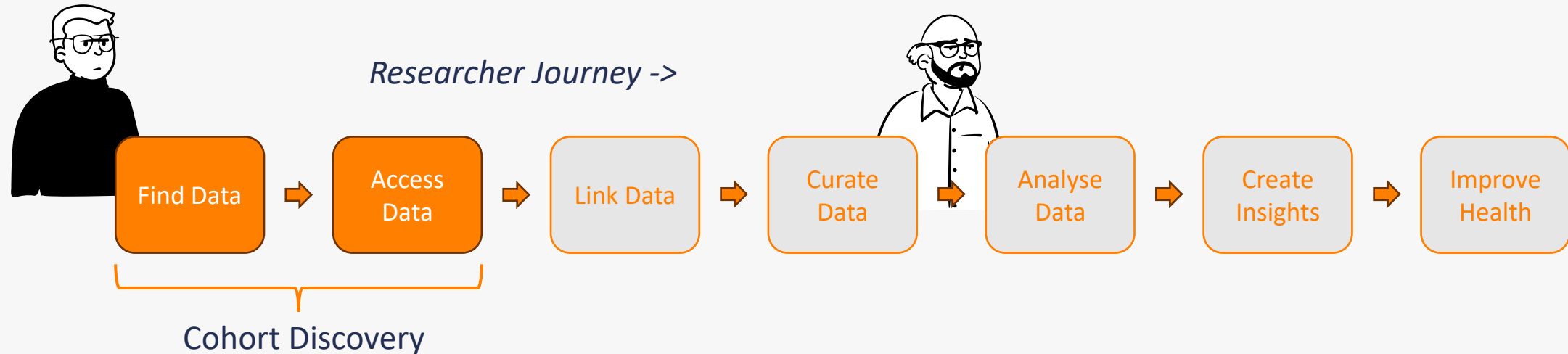
## Data Collections

### Active (n=22):

Kings College London  
NHS Great Ormond Street  
Hospital (GOSH)  
NIHR BioResource  
Public Health Scotland  
UCL Barts  
University of Bristol  
University of Cambridge  
University of Dundee  
University of Edinburgh (EPCC)  
University of Nottingham  
University of Oxford

### In progress (n=many!):

Clinical Practice Research  
Datalink (CPRD)  
NHS Secure Data Environment  
(SDE) Network



- Using the OMOP CDM, Cohort Discovery **streamlines** the discovery phases of research
- This **saves time, effort and costs** for both researcher and data custodian
- A key **challenge** is encouraging data custodians to transform their data to OMOP due to the **complexities of transformation** and **slow returns on investment**.