





Data Access and Discovery Webinar

Data journeys from two Health Data Research Hubs

Tues 6 Feb 2024, 14:00 - 15:00

Online





House Rules

Welcome!

- **Transparency**: Anyone is free to use information from our discussion, particular comments will not be attributed to individuals
- Optimism: Providing contributions that are positive and solutions-focused
- **Respect**: Promoting a culture of tolerance and inclusivity
- Courage: Be prepared to try new things and challenge
- **Humility**: Learning and respecting others and their diverse views

To read our meeting guidelines and etiquette in full visit our website:

www.hdruk.ac.uk/news-opinion-events/events/



Agenda

- 14:00 14:05: Hello and welcome by Alison Elderfield, Programme Manager, HDR UK
- 14:05 14:10: A brief history of the health data research hubs, **David Seymour**, Director of Infrastructure and Services, HDR UK
- 14:10 14:15: A patient perspective on health data research by Antony Chuter, PPIE Lead for Alleviate and Trustee of PainUK
- 14:15 14:30: Spotlight on Alleviate, **Gordon Milligan**, Deputy Director of Alleviate and **Chris Cole**, Director of Alleviate and Senior Lecturer, Health Informatics Centre at the University of Dundee
- 14:30 14:45: Spotlight on DATAMIND, **Ann John**, Co-Director of DATAMIND and Professor of Public Health and Psychiatry at Swansea University
- 14:45 15:00: Q&A with speaker panel, chaired by Alison Elderfield
- 15:00: Close



A Q&A with...

David Seymour

Director of Infrastructure and Services, HDR UK

Antony Chuter

PPIE Lead for Alleviate and Trustee of PainUK

Chris Cole

Director of Alleviate and Senior Lecturer, Health Informatics Centre, University of Dundee

Gordon Milligan

Deputy Director of Alleviate

Ann John

Co-Director of DATAMIND and Professor of Public Health and Psychiatry at Swansea University



Thanks for attending.
We hope to see you again soon!





An introduction and brief history of the health data research hubs

Data Access and Discovery Webinar

6th February 2024

David Seymour, Director of Infrastructure & Services, Health Data Research UK

Our mission & vision



HDR UK's mission is to unite the UK's health data to enable discoveries that improve people's lives

Our 20-year vision is for large scale data and advanced analytics to benefit every patient interaction, clinical trial, biomedical discovery and enhance public health.

Core funders 2023-28













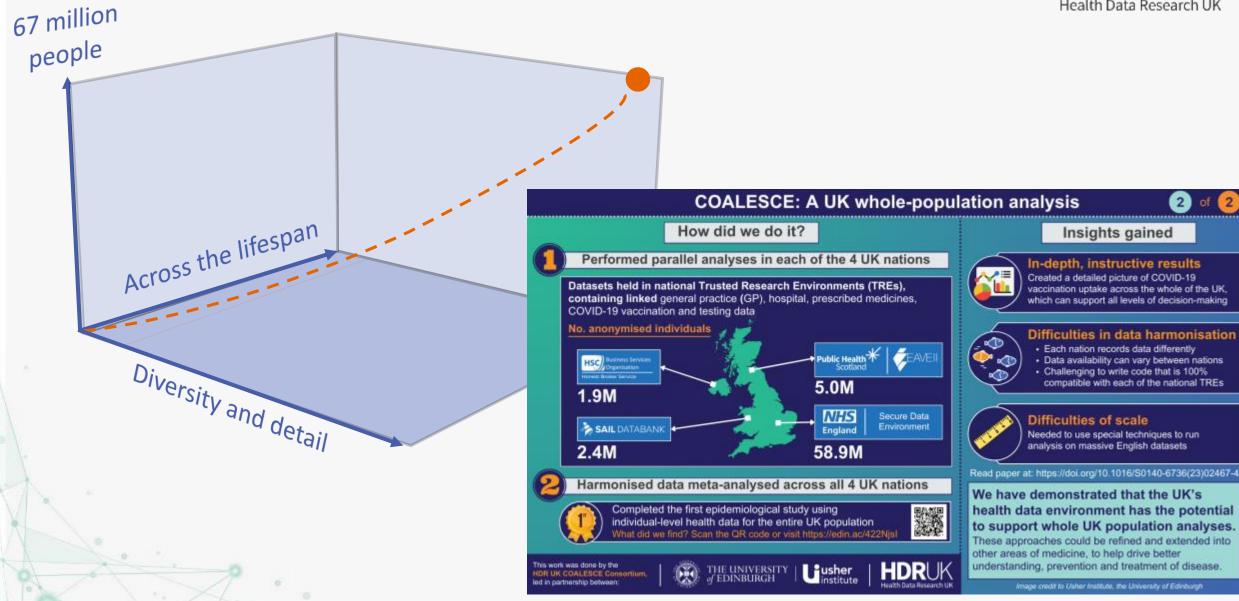






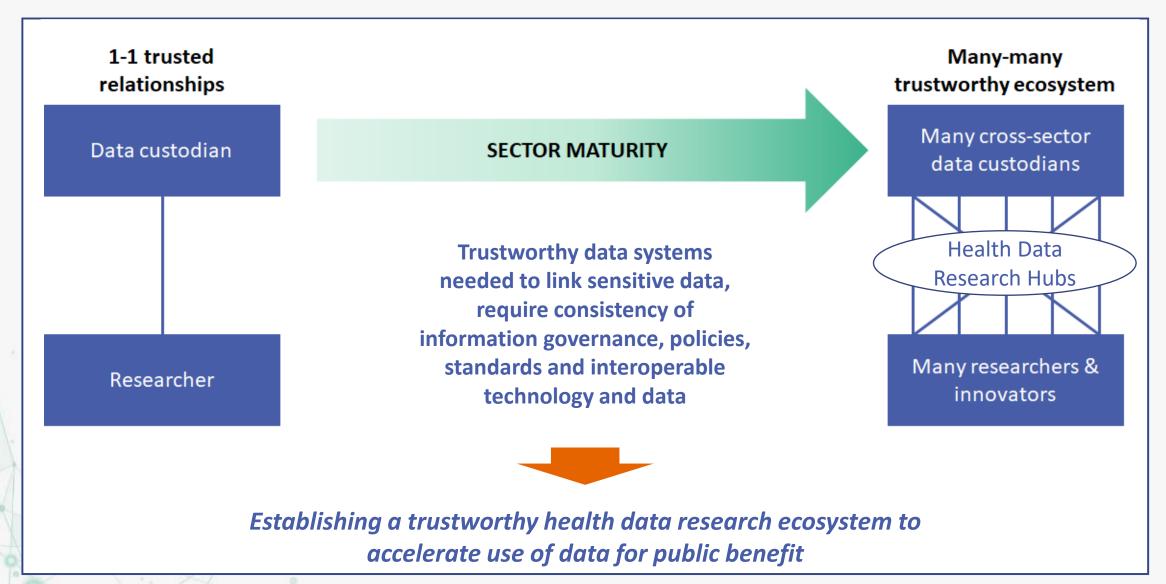
The scale of the ambition





We need a *trustworthy* health data research ecosystem to translate the potential of health data research at scale requires



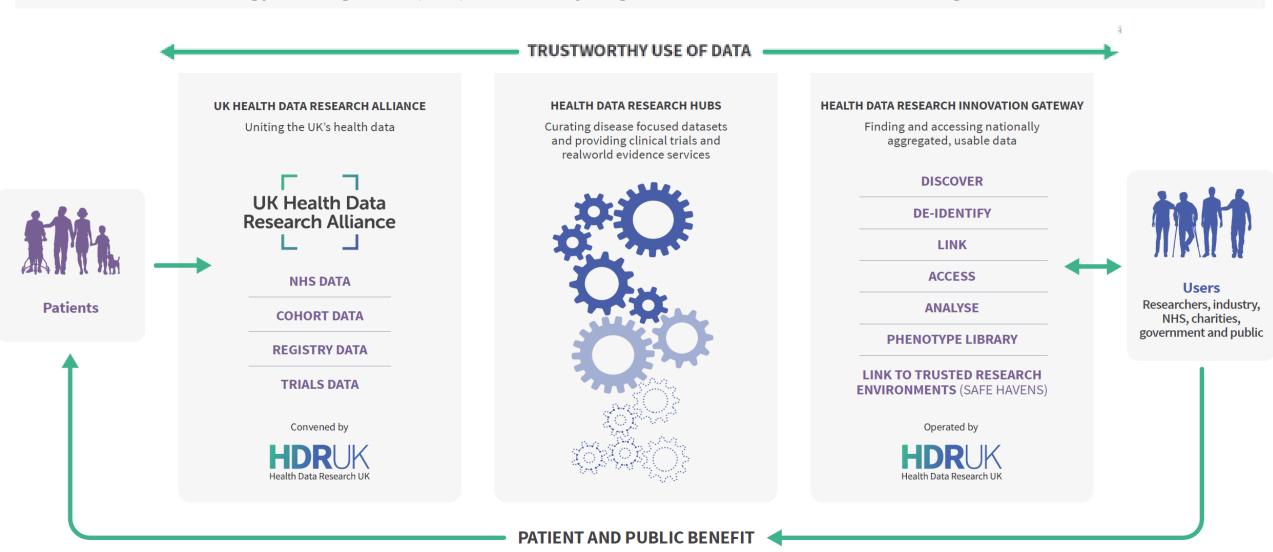


Where it all began...the Digital Innovation Hub Programme





Part of the Industrial Strategy Challenge Fund (ISCF) Data to Early Diagnosis and Precision Medicine challenge



Read more about the Digital Innovation Hub Programme and its evaluation, published August 2023







Visit the Gateway Visit the Alliance Visit HDR UK Futures

About Advancing Health Research Access Health Data Helping with Health Data Study and Train

Home | News

News

A pioneering programme to establish a national infrastructure for health data science

12 September 2023 | Author: Ruth Milne, Communications Manager, Infrastructure and Services

The Digital Innovation Hub Programme represents a defined period of funding but was delivered to support the advancement of a much wider agenda – to harness the power o UK health data for public benefit. The programme's instrumental effort has reshaped the health data research landscape and created lasting impact.



Visit the Gateway Visit the Alliance Visit HDR UK Futures

About Advancing Health Research Access Health Data Helping with Health Data Study and Train

Home | News

Opinion

Celebrating impact from our Health Data Research Hubs

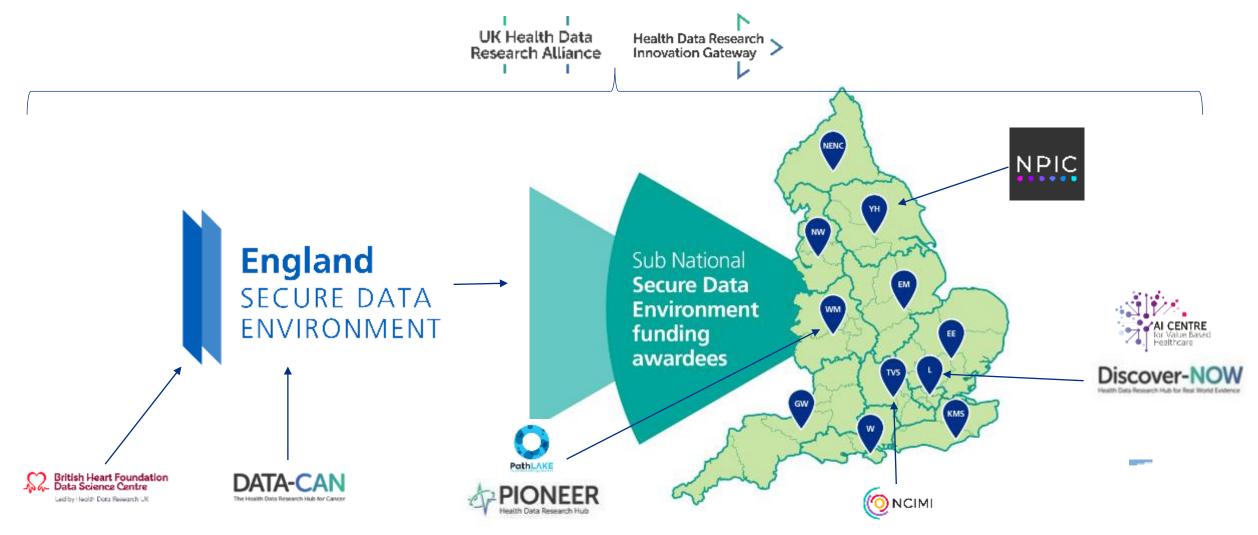
31 August 2023

Since the Hubs were set up in 2019 as part of the Digital Innovation Hubs Programme, they have developed a long way. Here, we look back to see how they've evolved and how they're enabling impactful research in a safe and secure way.





Helping to shape and deliver the NHS Data for R&D Programme - NHS Research SDE Network + NHS DigiTrials



Condition specific Hubs, Platforms and Data Science Centres





Centres of excellence with expertise and tools enhancing data foundations and providing FAIR data & services



Led by Health Data Research UK



Acute Care (Birmingham)



Inflammatory Bowel Disease (Cambridge)





Dementia (Oxford, SeRP/Swansea)



Respiratory Health (Edinburgh, SeRP)



Cancer (Leeds, National SDE)



Eye Health (Moorfields)



Pain (Dundee)



Mental Health (Swansea, Maudsley)

NB. <u>Discover-NOW</u> (Imperial), the health data research hub for real world evidence is condition agnostic and part of London Secure Data Environment | 8 development

Latest addition, focused on Traumatic Brain Injury





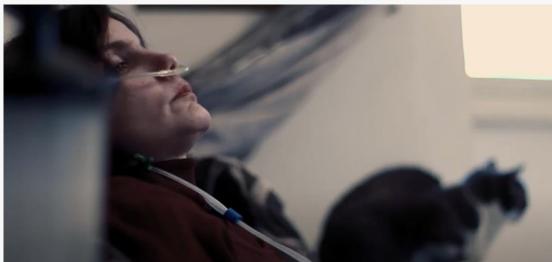
Jointly funded by the Medical Research Council (MRC), the National Institute for Health and Care Research, the Ministry of Defence, Alzheimer's Research UK

Why the Data to Early Diagnosis and Precision Medicine challenge still matters











"Patients want their data to be used to improve care. In fact, they're often surprised it's not used already."

Jacqui, cancer survivor and patient advocate



The Advanced Pain Discovery Platform (APDP) Data Hub

> Dr Christian Cole and













Background



- Chronic pain is defined as Pain lasting for three months or longer
- According to 2016 research by the British Pain Society around 28 million adults in the UK are living with Chronic pain
- Chronic pain is poorly captured in health data
- Chronic Pain first recognized as a condition under ICD-11 by WHO
 - Improved evaluations
 - Better treatment decisions

Feedback and opinions of people who live with pain is invaluable for pain related research



The Advanced Pain Discovery Platform (APDP)





- Five year programme initially
- £34m from MRC, ESRC, BBSRC, Versus Arthritis and Lilly









APDP Consortia





Neuropathic Pain

ADVANTAGE
Visceral Pain





Adverse Childhood Experience & Pain

Alleviate is the data hub of the APDP, and will collaborate with the consortia to streamline access to data and on-board data into the Alleviate federated platform where appropriate





FAIR Principles



Findable

Metadata and data should be findable for both humans and computers

Accessible

Once found, users need to know how data can be accessed

• Interoperable

Data needs to work with applications or workflows for analysis, storage and processing

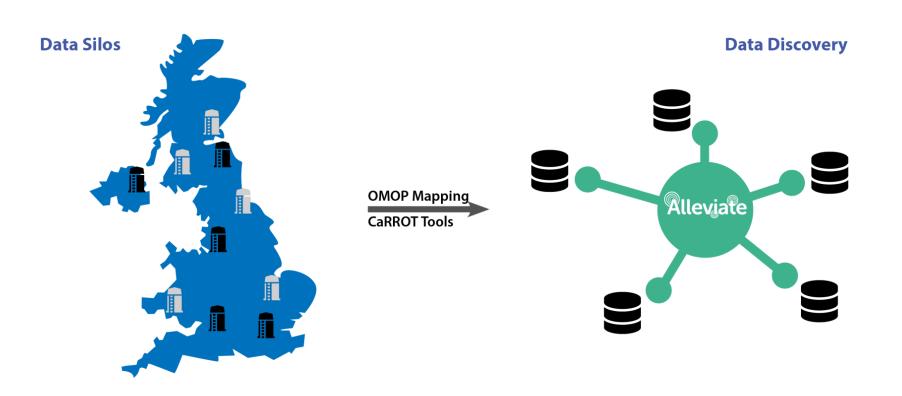
Reusable

The goal of FAIR is to optimise data re-use via well described metadata

Solution to a Common Problem



- Pain datasets are in silos
- Hard to find
- Even harder to query
- Datasets are in different formats



Alleviate.ac.uk

HDR UK Metadata Catalogue





HDR UK Gateway

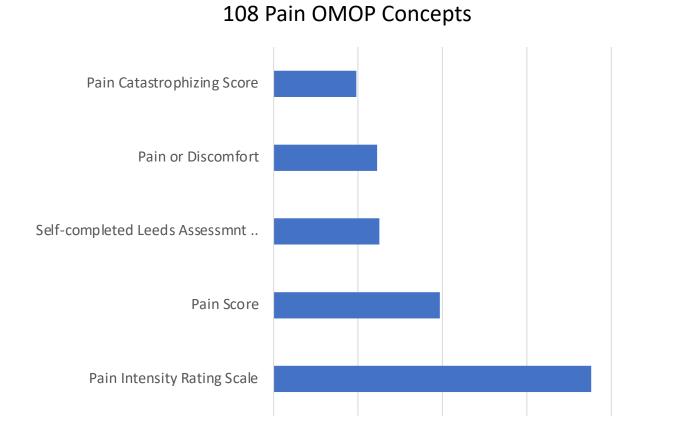


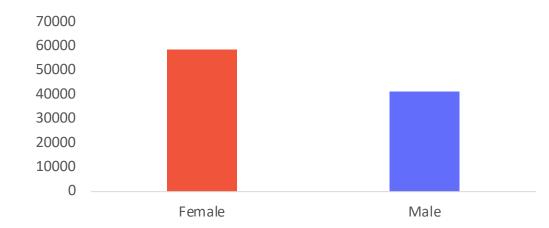
- Registered 20 Pain datasets with HDR UK's Metadata Catalogue
 - Available in the <u>Alleviate Collection</u>
- Currently added three Pain specific data sets to HDR UK Cohort Discovery Tool
 - Many further datasets in the pipeline for CDT

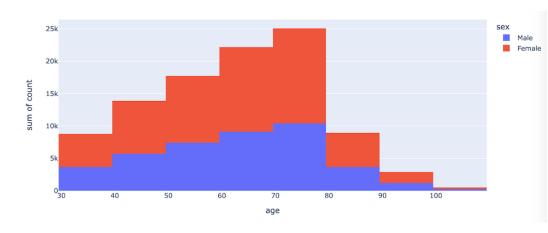


Generation Scotland



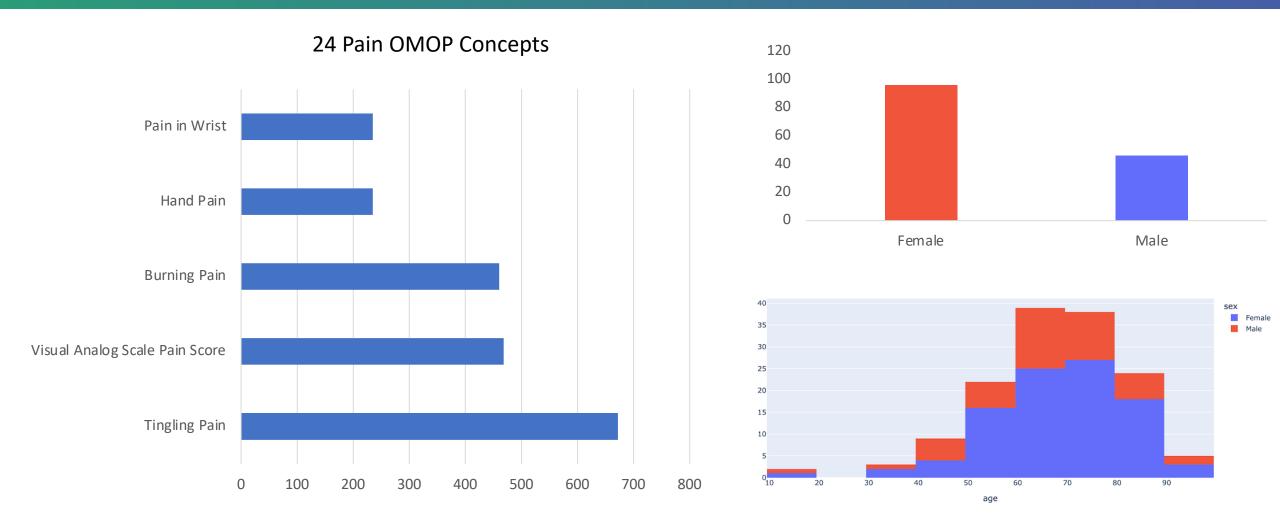






Oxford Carpal Tunnel Syndrome

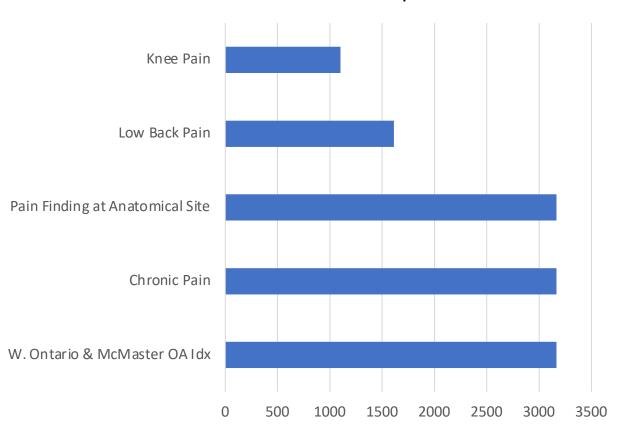


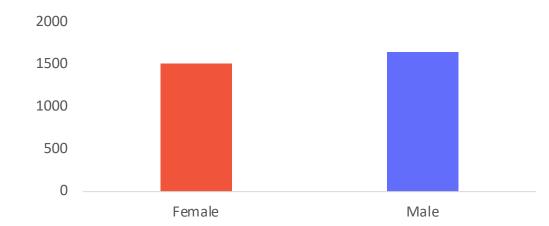


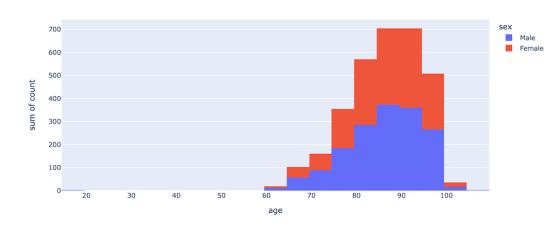
GOAL











Data Harmonisation & Standardisation



- Data harmonisation across data sources
 - Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)
- Using CaRROT Mapper and CDM Tools
 - Initially developed during CO-CONNECT
 - Continued development by Alleviate Team, Digital Research Services, University of Nottingham and HDR UK

Visit https://carrot4omop.ac.uk for more information on these tools

What Data is Being Mapped?





Alleviate Strengths

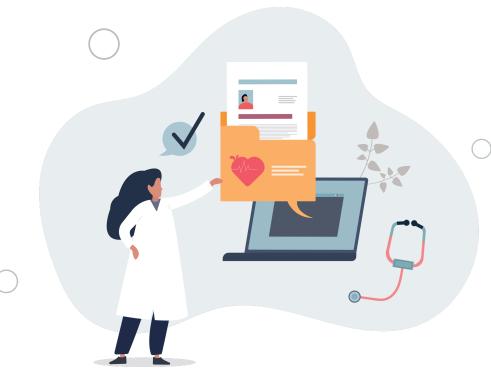


Data Services	Public and Patient Engagement	Data Sharing
 Advice on mapping to OMOP End to end service to map your data to OMOP Guidance on setting up studies for efficient data management OMOP Data Linkage Clinical Data Genomics Data Trusted Research Environment (TRE) 	 Access to an active patient group for input on e.g. academic proposal Time needs to be remunerated Awareness and visibility of Chronic Pain and its impacts Raising awareness of chronic pain through marketing campaigns Ran UK National Pain Survey 	 Data sets standardised to OMOP Data sets discoverable through HDR UK Gateway Data access requests through HDR UK Gateway Data extracts available via Alleviate TRE

Data Mapping Experience

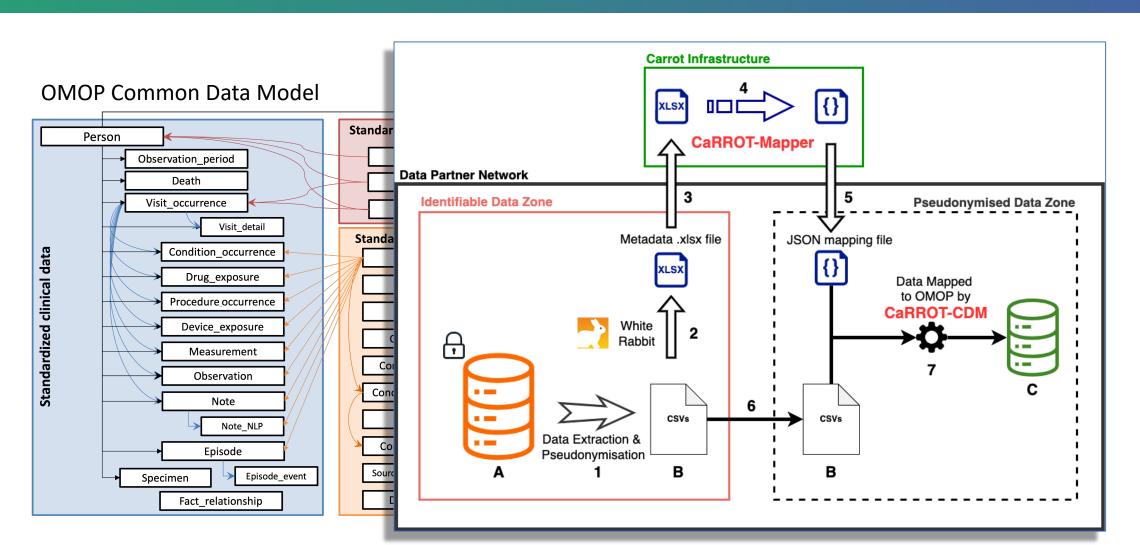


- Data Discovery and Re-use (FAIR)
- Data Team through Alleviate (+CO-CONNECT) have been involved in mapping >10 million health records:
 - Research cohorts
 - Population scale clinical datasets
- All 19 datasets currently in HDR UK Cohort Discovery Tool



Data Mapping Experience





Our Data Mapping Service



- Mapping health data to the OMOP CDM
- Review mappings to provide advice
- Provide guidance on your OMOP health data journey
- Provide training on best practices for OMOP use
- Review existing plans for implementing OMOP and providing feedback on improvements
- Including OMOP as part of a research collaboration



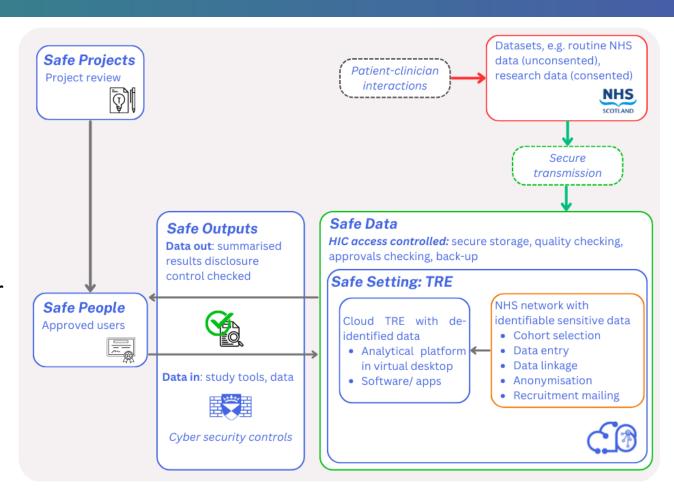
Trusted Research Environment



The Advanced Pain Discovery Platform (APDP) Data Hub

- Hosted by HIC, University of Dundee
 - >10 years experience
 - Flexible cloud compute
- "Five Safes" compliant
- We provide access to project data (following approvals), which is always pseudonymised and provisioned in a segregated workspace (green box for TRE). Diagram shows 5 Safes in blue and colour coded traffic light system for risk.
- Strict monitoring and auditing capabilities





Collaborative Partners





University of **Dundee**





















Questions?





@Alleviate_Data



@alleviate-hub



Alleviate.ac.uk



Ann John- Co-Director







Research and data and MH

Studies face problems:

- People most likely to experience poor MH may be least likely to take part
- Many measures are self-reported and may be subject to bias and/or their
- clinical meaning may be unclear
- People are lost to follow up
- These partly overcome through using linkage to routinely collected data
- Real world data
- Cohorts important but numbers
- It speaks to policy makers, the public and practitioners
- Issues re trust and innovation

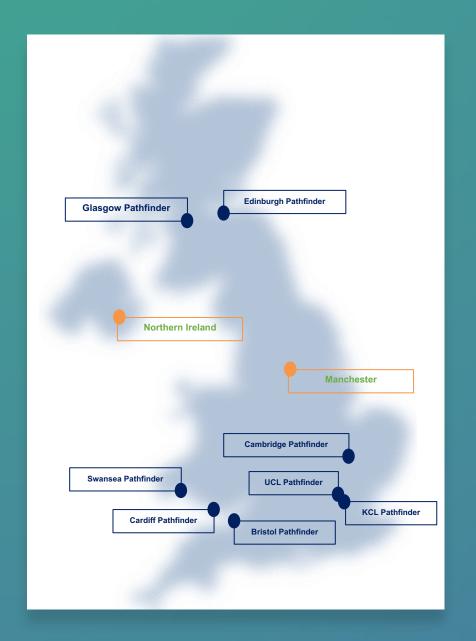








Overview



Workforce Public Business FAIR Curated Capacity, Training Participation, Development & Core Activities & Development Partnerships & Data Sustainability Governance Challenge Areas Children & Young People **Excluded and Underrepresented Group** Interfaces between Physical and Mental Health Severe Mental Illness (SMI) Milestone 1 Datasets available Discoverable excluded and Linking physical & mental Discoverable Schools Milestone 2 under-served populations health in SMI Milestone 3 Reusable free text data Digitally enhanced trials Drug discoverability



Meet the Team

Leadership Team



Ann John
Principal Investigator
and Co-director



Rob Stewart Co-director



Andrew McIntosh
Chief Scientific
Officer

Management Team



Matthew Broadbent
Chief Operating
Officer



Dara Almeida Medina Project Manager



Matthew Broadbent
Chief Operating
Officer



Dermot O'Reilly
Chief Data
Officer



Parisa Mansoori Core Activity Lead



Rudolf Cardinal Core Activity Lead



Louise Arsenea ult Core Activity Lead



Matthew Iveson
PM Business
Development &
Sustainability



Linda Jones PPI Lead



Elena Triantafillopoulou PM FAIR Curated MH Data



Hala Rachidi
PM Workforce
Capacity, Training &
Development



Claire McLoughlin Roadbuilders Innvations



Public Participation, Partnerships and Governance











SRAG
Shortlisted
for HDRUK
PPIE
Award
winner to be announced

Forum
Contribution
SRAG shape
guidelines on
data sharing with
industry

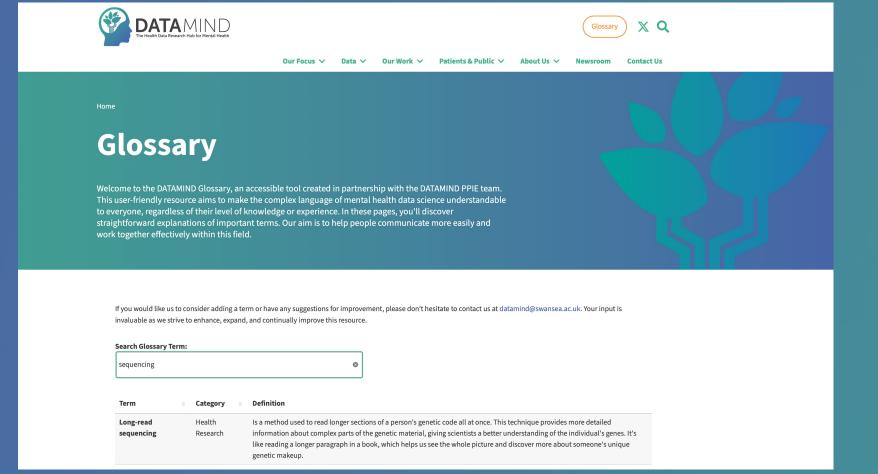
Opinion Article:

SRAG publish article in Frontiers in Psychiatry on Al and Mental Health

Engage with Other Org Collaboration with CRUK, HDRUK, HRA, MQ, McPin, and others

Empowering







Search Glossary Term: linkage							
Term \$	Category	‡	Definition				
Linkage of data (data linkage)	Processes		Joining (linking) data from more than one source. For example, to study the relationships between mental and physical health conditions, it might be necessary to link data from NHS mental health services to primary care (GPs) or acute hospital data. To study the relationships between health conditions and education, it might be necessary to link data from health services and a government education department.				
			Linkage may be legally complex because it involves data from more than one data controller. Linkage may be based on straightforward rules ("two records with the same NHS number are from the same person") or based on probability ("if two records share the same forename, surname, and date of birth, they are more likely to be from the same person"). Links may be made using identifiable data (e.g. NHS number) or de-identified data (e.g. a research pseudonym).				



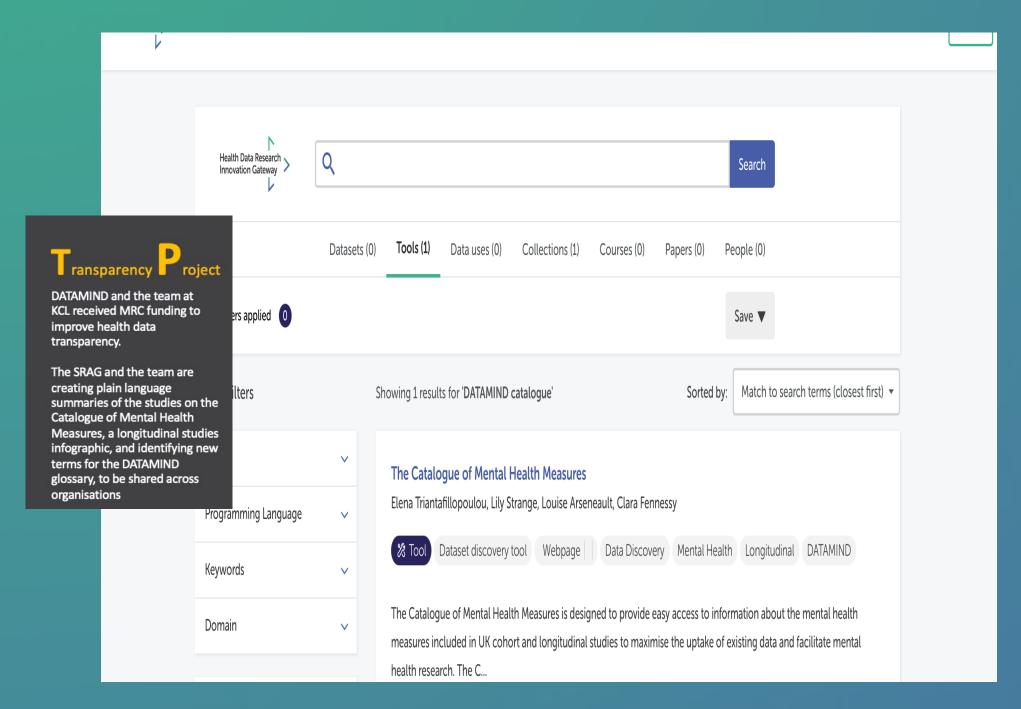
FAIR Curated DATA







SPECIALIST CATALOGUE AND ANNOTATION



Updates

Added new studies: British Autism Study of Infant Siblings (BASIS), Longitudinal Outcomes of Gender Identity in Children (LOGIC), Edinburgh Study of Youth Transitions and Crime (ESYTC)

Collaborated with other discoverability platforms: The Catalogue of Social Care Measures, HARMONY, Center for Pediatric Traumatic Stress

Developed and embedded an API

55 UK-based longitudinal studies reviewed

Over 3000 mental health measures from over 300 waves of data collection

More than 30 mental health and wellbeing topics





BCS70

1970 British Cohort Study





AESOP-10





































National Child

Development Study

OUTHAMPTON

























Airwave Health Monitoring Study

CFAS









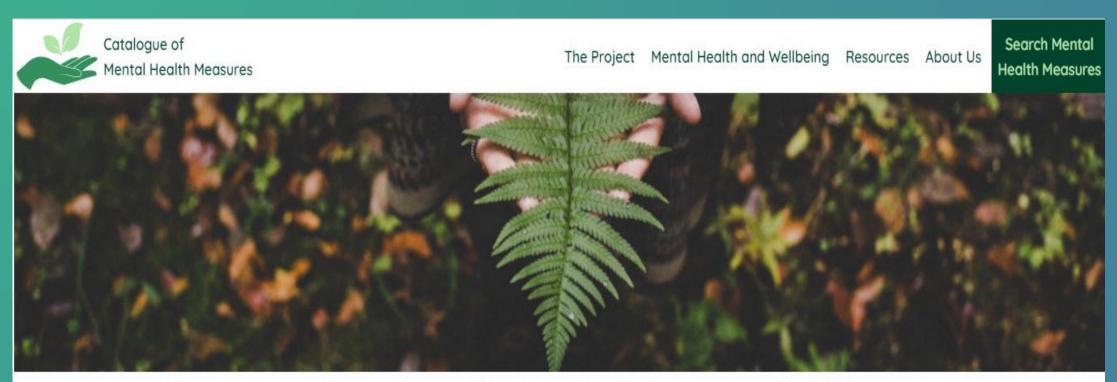








Harmonisation



An interactive catalogue of mental health and wellbeing measures in British cohort and longitudinal studies



The UK boasts rich and world-renowned cohort and longitudinal studies which offer unique opportunities to answer key questions about mental health and wellbeing.



The Catalogue provides information about thousands of standard and non-standard measures of mental health and wellbeing collected in UK longitudinal studies.



Measures detailed in the catalogue include indicators of mental health problems, treatment, service use, impairment and psychological wellbeing.

Search Mental **Health Measures**





Physical health data

MH across cohorts

Genetic summary tables

Data harmonization- HARMONY

Data from from third sector











Gateway to health data and tools for research

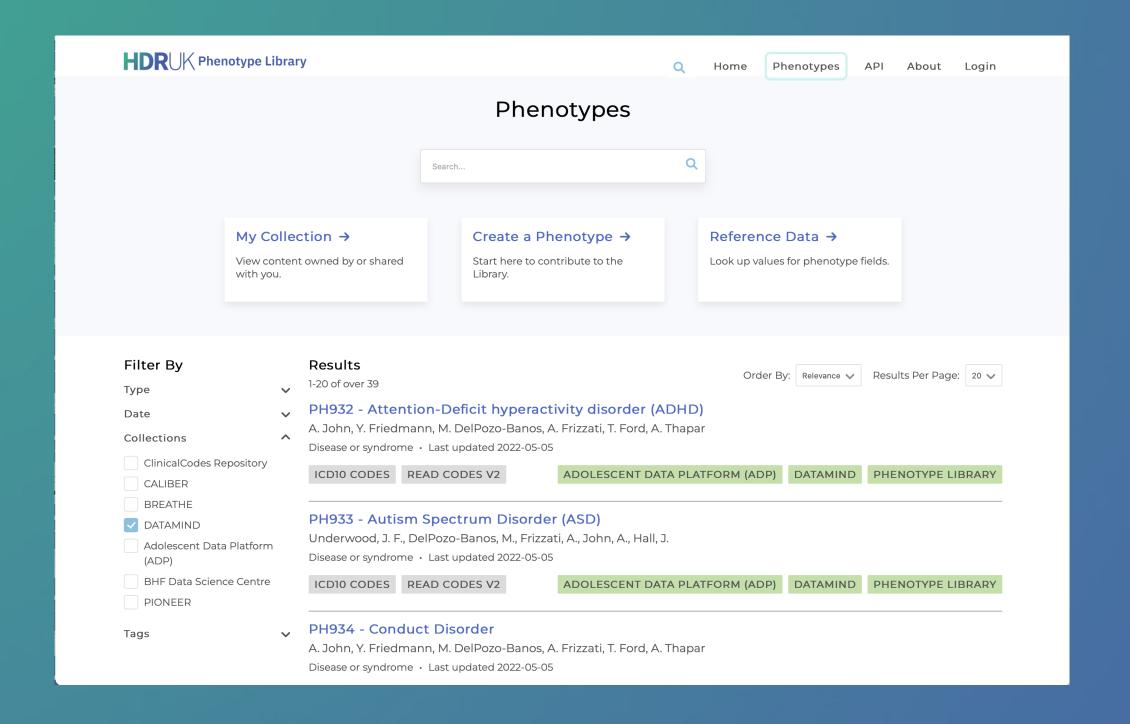
Q I'm looking for.













Children's Health & Wellbeing Dashboard

Local Authority	Gender	Value	2017	2019	2021
Monmouthshire	Persons	%	59.1	54.3	52.1
Newport	Persons	%	50.5	50.6	45.6

Show/hide technical info

Survey question: How often do you usually have breakfast (more than a glass of milk or fruit juice) on weekdays?

This indicator presents the percentage of children who eat breakfast every weekday (Monday-Friday). The data includes 11-16 year olds who answered that five days out of than a glass of milk or fruit juice for breakfast.

secondary and middle schools in Wales are SHRN members. For further details on response rates and survey methodology please see the SHRN well

No weights were applied to the data presented in this dashboard as the demography of the survey sample closely matched the equivalent population of Wales due to the high most variables, responses to the questionnaire item are presented and then a binary indicator is defined. For each indicator, results are presented by: gendler, age (where a fifture), local authority & local health board. Confidence intervals have been calculated using the profile likelihood method (Venative).

In the data table, values are highlighted in dark or light grey in the data table where the selected breakdowns have a significantly higher or lower result than Wales, respectivally is significantly different if its confidence interval range does not include the Wales value for the relevant year. For example, if the 2021 Wales value is lower than the value is significantly builder than the Wales value, the converse is also true for significantly lower than Wales value interval overlaps with the Wales value, then the difference is not significant and therefore not highlighted in the data table. The HBSC Flamily Affluence Scale (FAS) has been young people's socio-economic status, based on a set of questions which measure the material conditions of the household in which young people live. A binary affluence so between 0 and 7, with high scoring between 8 and 100 has been used in place of the usual high, mid and low standard scoring system, due to one of the six questions make no longer being an appropriate measure as a result of pandemic restrictions (Has your family traveled outside of Wales on heliday?). This method has been used across all to comparability in this deshboard and is therefore not comparable with results in the SHRN National Reports. Those who did not provide a response to all of the questions which were not given a high or low result; measure; have been excluded in 2017/28, 5,3 were excluded in 2019/20, and 5,732 (4.7%) responses were excluded in 2017/28, 5,3

as calculated using results provided for month and year of birth, which were then subtracted from month and year of survey being completed. If month of birth was no has been as proxy and age was assumed (only 0.57% total results were given a proxy age from just year data across all three years of surveys).

For presentational purposes, the scale of each figure's axis is variable and relative to the data being presented. Care should therefore be taken when reading the figures, to

Please note, data produced in this dashboard may not fully match the data found in the SHRN flational reports due to extra exclusion criteria set by Public Health Wales. Fur rounding in SHRN reports may lead to small discrepancies in reported percentages. Similarly, numbers within SHRN's Local Authority reports may not match numbers in this methods may differ.

If you have any comments or feedback, then please get in touch with us by emailing; zoe strawbridgeshwales nhs.uk













Linking physical & mental Discoverable excluded and Discoverable Schools Milestone 2 under-served populations health in SMI Milestone 3 Reusable free text data Digitally enhanced trials Drug discoverability

Select a chart view

- 2017 data, all local authorities
- 2019 data, all local authorities
- 2021 data, all local authorities
- Custom view

Select Gender

- Persons
- Female
- Male
- Male and female

Sort by value?

- On
- Off

Mean mental wellbeing scores out of a total of 35 using SWEMWBS, persons, aged 11-16, local authorities, Wales, 2021

Produced by Public Health Wales Observatory, using SHRN data

95% confidence interval Wales, 2021 Isle of Anglesey Gwynedo Conwy Denbighshire **Flintshire** Wrexham Powys Ceredigion Pembrokeshire Carmarthenshire Swansea

Neath Port Talhot





Training and Capacity





Core Activity 4: Workforce Capacity Training & Development





CONFERENCES & WORKSHOPS:

Build capacity on research practices, Data Science methods, PPIE, Leadership



VIDEO COURSES

Adapt courses & training material into accessible video content



MH DATASCIENCE COMMUNITY

Convene ECRs to share career development and collaboration opportunities and experiences





Workshop Series

PROGRAMMING

Programming principles
Practice session in R & Python

SEP 2022

DATA MANAGEMENT

Processing & linkage of routinely collected data Data catalogue & Phenotype library

APR 2023

04

DATA ANALYSIS

Epidemiology Advance statistical analysis

NOV 2023

MACHINE LEARNING 2

02

05 ANNs NLP

AUTUMN 2024

MACHINE LEARNING 1

Methodological principles
Distance based classifiers
Decision trees
SVMs

SPRING 2024





Workshops & Meetings



















Our Focus V

Data ∨ Our Work ∨

Patients & Public >

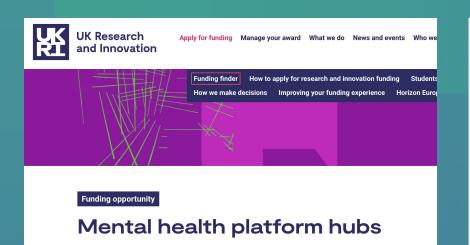
About Us V

Newsroom

Contact Us

Home > Data

Trusted Research Environments





Core Outcome Measures in Effectiveness Trials

"A core outcome set (COS) is an agreed standardised set of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or health care."



An individual's mental and physical health are closely linked. However, this fails to be reflected in starkly disconnected mental and physical healthcare provision, or in the public nealth and research sectors (DoH, 2017). Enhanced understanding of the links between physical and mental health is now a priority for patients, research councils and the NHS. The

are from studies concerned with physical ill health. A significant gap in these data is information about participants' mental health and wellbeing. However, public data collection initiatives require substantial PPIE, as well as multi-organisation agreement about the content and format of the datasets collected. Moreover, the technical infrastructure to capture mental health data routinely within physical health studies needs to be designed.

health research projects. This involves workshops with the public and other researchers to understand the concerns people may have about sharing mental health data when partaking in a physical health study – ensuring mental health outcomes collected in the Core Mental Health Data Set (CMHDS) reflect what the public and researchers feel are important elements of mental health to be captured in all studies. This is in line with COMET's goal to create outcomes that are relevant to patients and health professionals.

We will then test the feasibility of embedding the CMHDS in 3 physical health studies. The ultimate aim of the research is to have a core set of mental health questions asked acros all physical health studies, giving us consistent data which can be compared and analysed and providing a broader understanding of the mental health status of the population at large, as well as an insight into the links between mental and physical health.

Prof Kathryn Abel (University of Manchester)

Prof Ann John (Swansea University)

Prof Paul Dark (University of Manchester)

Dr Nawar Bakerly (Salford Royal NHS Foundation Trust)



THANK YOU!



- https://twitter.com/DatamindUK
- datamind@swansea.ac.uk
- https://popdatasci.swan.ac.uk/centres-of-excellence/datamind/