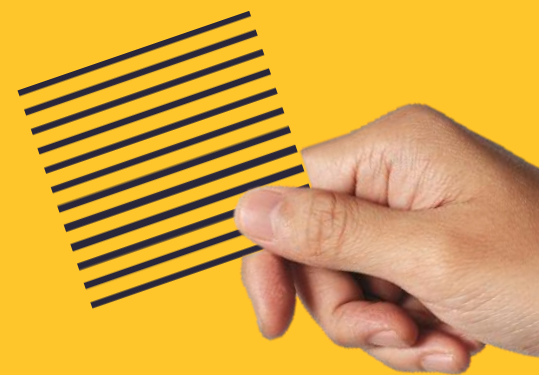




Our Future Health: A
new prospective
cohort study to
prevent, detect, and
treat diseases

Julia Lanoue & Ben Cairns

OMOP UK Symposium
27 Sept 2023





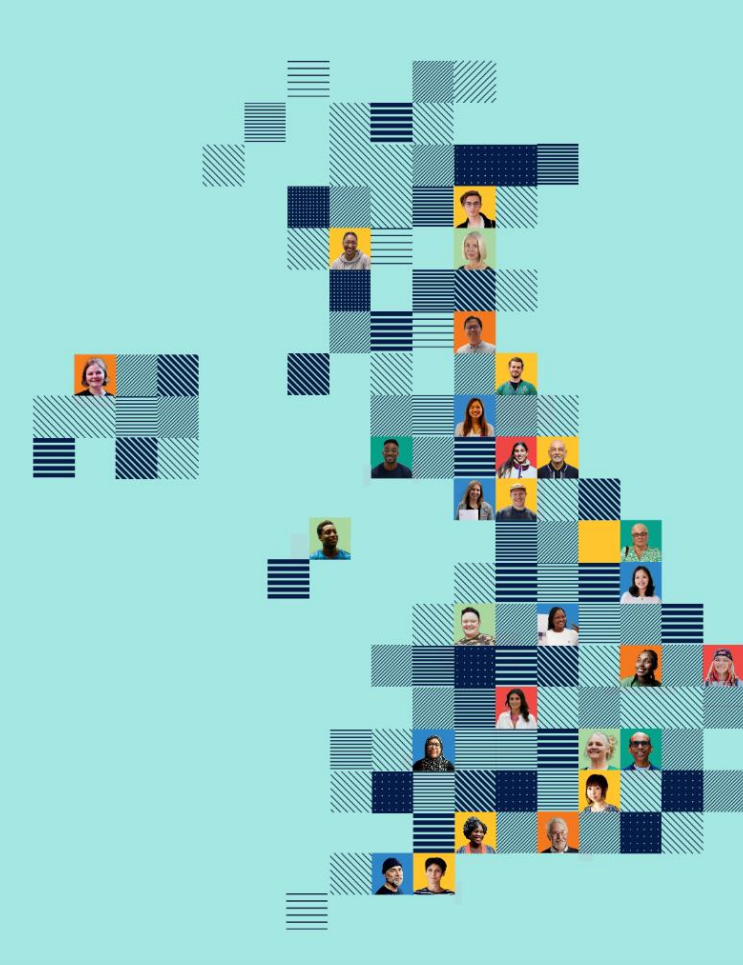
Our Future Health will be the UK's largest ever health research programme.

It is designed to help people live longer and healthier lives through the discovery and testing of more effective approaches to prevention, earlier detection and treatment of diseases.

The **5 million** volunteers sharing information will create an incredibly detailed picture that reflects the whole of the UK.

Aims of the programme

- 1. Build a programme** to generate and link multiple types of health and health-relevant information on five million people in the UK to facilitate **discovery** and **translational** research by academic, charity and commercial researchers
- 2. Estimate personal disease risk** information for participants, based on **genetic** and **non-genetic** information, and offer this to participants who wish to receive it
- 3. Re-contact sub-groups of participants** generally or based on personal characteristics or indicators of disease risk for additional biological samples, questionnaires, or invitations to take part in clinical trials and other studies by academic, charity and commercial researchers



Funders

Affiliate charities





Recruitment and
data collection



Taking part in Our Future Health

Recruitment started July 2022, following successful pilots in 2021.

Most participants have received an invitation via NHS Digital, but registration is open via our website.

Sign-up/invitation



Welcome! You've been invited to join Our Future Health.

Thank you for your interest in joining the programme. With your help, we can help researchers tackle the growing burden of disease in the UK so future generations

Let's prevent disease together

Our Future Health is the UK's largest ever health research programme.

At Our Future Health, our goal is to help future generations live in good health for longer.

We are aiming to collect information and samples from up to 5 million adults to build the most detailed picture ever of the UK's health.

Researchers will be able to apply to use this information to make new discoveries about health and disease.

Discoveries made through Our Future Health could lead to new ways to predict, prevent and detect diseases like dementia, cancer, diabetes, heart disease, and stroke.

We're inviting people aged over 18 from all backgrounds across the UK. Because to help all kinds of people, we need every kind of person.

[Start now](#)



Register

Enter your details to register

About you

First name
Joe

Last name
Bloggs

Date of birth
For example, 31 3 1980
Day Month Year



Information

You'll need to complete the following steps online:

1. Register by creating an account
2. Read about the programme (Participant Information)
3. Consent to take part
4. Complete a questionnaire about you and your health (5 sections)

After joining, you'll be sent a kit to provide a sample of your saliva, which you'll need to send back to us.



Consent

Consent Form

Chief Investigator: Dr Andrew Roddam of Our Future Health

If you wish to take part in the Our Future Health research programme, please complete the following form to show you understand and agree to what it involves. You need to tick all boxes from 1 to 11 to be eligible to take part.

1. I have read and understood the Our Future Health participant information sheet dated 5 March 2021 (version 1.1). I have had the opportunity to ask questions, and these have been answered fully.
2. I understand that taking part in Our Future Health is my choice.
3. I understand that Our Future Health will access, store and link to health-related records about me held by NHS Digital and other UK NHS bodies. I understand that Our Future Health can collect and use this information for research analysis at any point during my lifetime and beyond.
4. I understand that samples of blood and/or saliva that I give to Our Future Health will be stored and analysed in future health-related research. I understand that my de-personalised samples and data could be sent to approved processors outside the UK for analysis.
5. I understand that DNA will be extracted from my blood and/or saliva samples. I understand that my DNA will be analysed for health-related research using a technology called SNP array, and that other technologies might also be used, such as genome sequencing.
6. I understand that Our Future Health might contact me again in the future, including to:
 - send me news and updates about the research programme
 - ask me about my experiences of taking part
 - ask me to complete another questionnaire
 - ask my permission to collect health-relevant information about me from other sources
 - invite me to attend an appointment to give a further blood sample
 - invite me to attend an appointment for other assessments, such as imaging
 - ask if I would like to receive personal information arising from my samples or data
 - invite me to take part in other research studies.
7. I understand that researchers approved to access the information collected by Our Future Health could be from

What our participants donate to Our Future Health



Physical measurements and blood samples for genotyping and for long-term sample storage of plasma, buffy coat, and DNA



Responses to a baseline questionnaire with questions on demographics & household, work & education, lifestyle, family history, health history



Permission for data linkage to health-related datasets (e.g., primary care, secondary care, cancer, death)



Permission to re-contact for further surveys/biospecimens/assessments and to invite participants to re-contact studies.

Current and future development of data resources





Released data



Our Future Health Trusted Research Environment

Data management

Powerful, intuitive cohort browser

Combined clinical data and omics information

Latest data releases available quarterly

Analysis

Scalable analysis, with JupyterLab & Spark

Genomics Apps, with 'Swiss Army Knife' & PLINK

CLI for power users and multiple workflow languages

Collaboration

Share data, tools and analyses efficiently with collaborators

Refined access, security, and compliance made simple

Organisation tools simplify billing + £50 starting credits



Number of participants

In our Trusted Research Environment, we have:

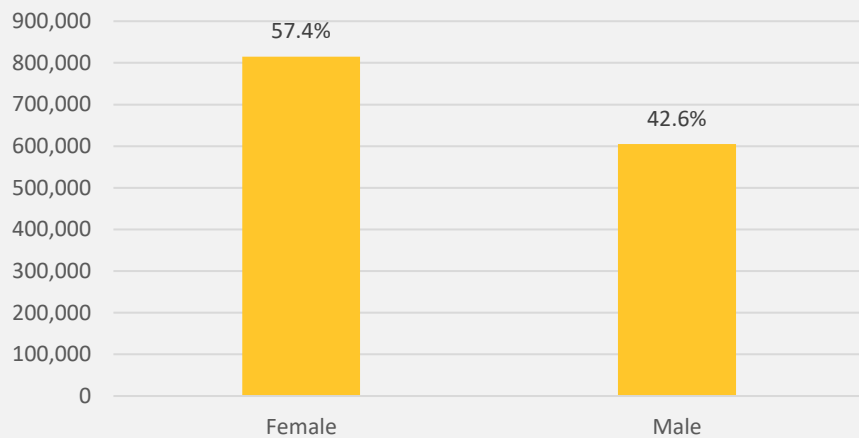
1,193,001 participants with completed
baseline health questionnaires

330,058 participants with genotype array data

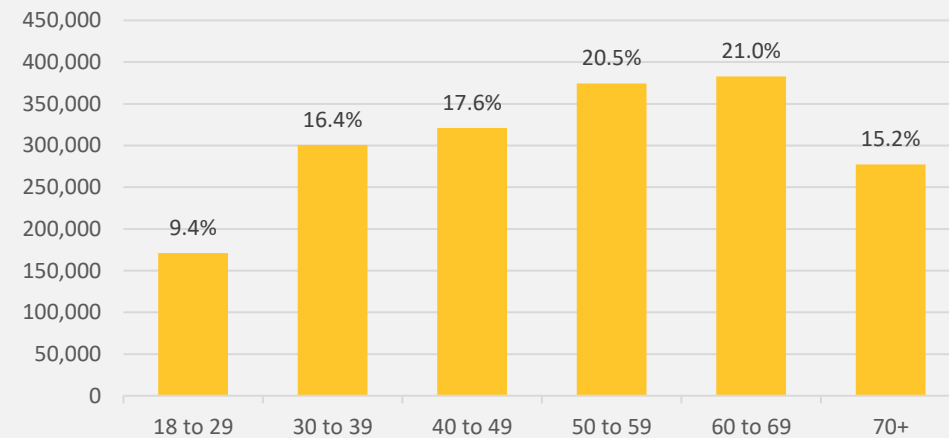
957,444 participants with NHS-E data
(HES, ONS deaths, Cancer)

Demographics – 1.8m consented participants

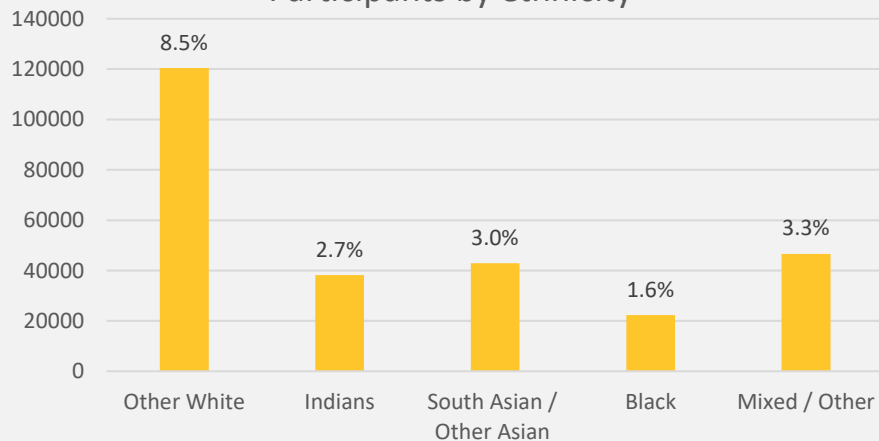
Number of male and female participants



Number of participants by age group

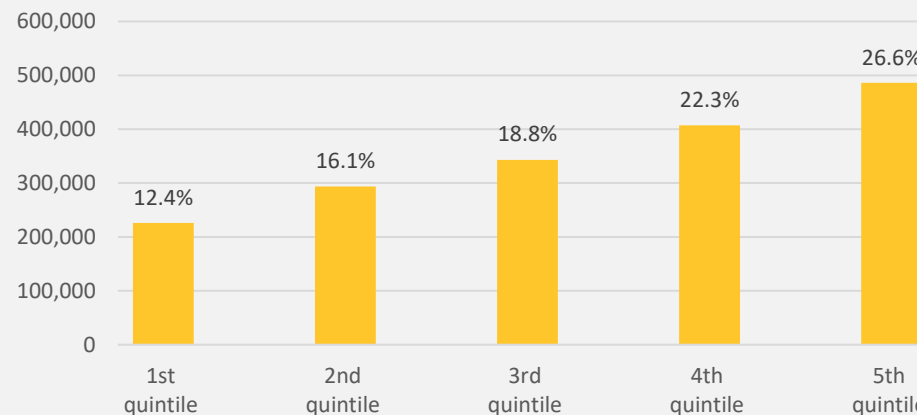


Participants by ethnicity



*81.0% White British participants

Distribution of participants across IMD* quintiles



*Index of Multiple Deprivation - 1st quintile is most deprived; 5th is least deprived



Questionnaire and
Linked Health Records



Baseline health questionnaire

- The current baseline health questionnaire has a total of 288 questions, with 2479 distinct response codes
 - There are two main versions of the questionnaire (currently v2.2).
 - Minor versions have logic or minor data type fixes.
- There are 68 core questions that all respondents see, and 220 dynamic questions.

You/Your Household

age
gender
ethnicity
language
housing
vehicle ownership
marital status

Work and Education

income
job type
work patterns
qualifications

Lifestyle

physical activity
socialising
TV/com use
sleep
smoking
alcohol

Family History

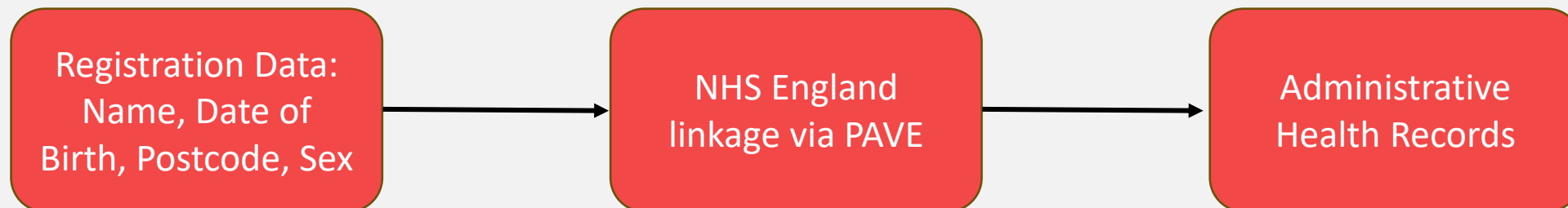
parent status
parent medical history
sibling medical history

Your Health History

BMI
disability status
COVID
UV exposure
oral health
falls
respiratory symptoms
amputation surgery
pain
cancer screening
parenthood
obstetric/gynae history
female contraception
comorbidities
medications
mental health

Linked Health Records

- Linked health records have been available on our Trusted Research Environment since June 2024.
- Only participants who have completed and submitted a baseline health questionnaire are eligible for linkage.
- Currently we are only linking participants to records available through NHS England.



Linked Health Records Datasets

- The datasets contain a curated set of variables, focusing on diagnoses and procedures.
- In the future, we are planning to release more HES and NDRS Cancer datasets as well as new datasets, such as the National Diabetes Audit.
 - We also actively working with the NHS to make primary care data available for research.

Office of National Statistics
Death Registration

Hospital Episode Statistics
Admitted Patient Care
Accident & Emergency
Outpatient

National Disease Registry Service
Cancer Pathways
Cancer Registration

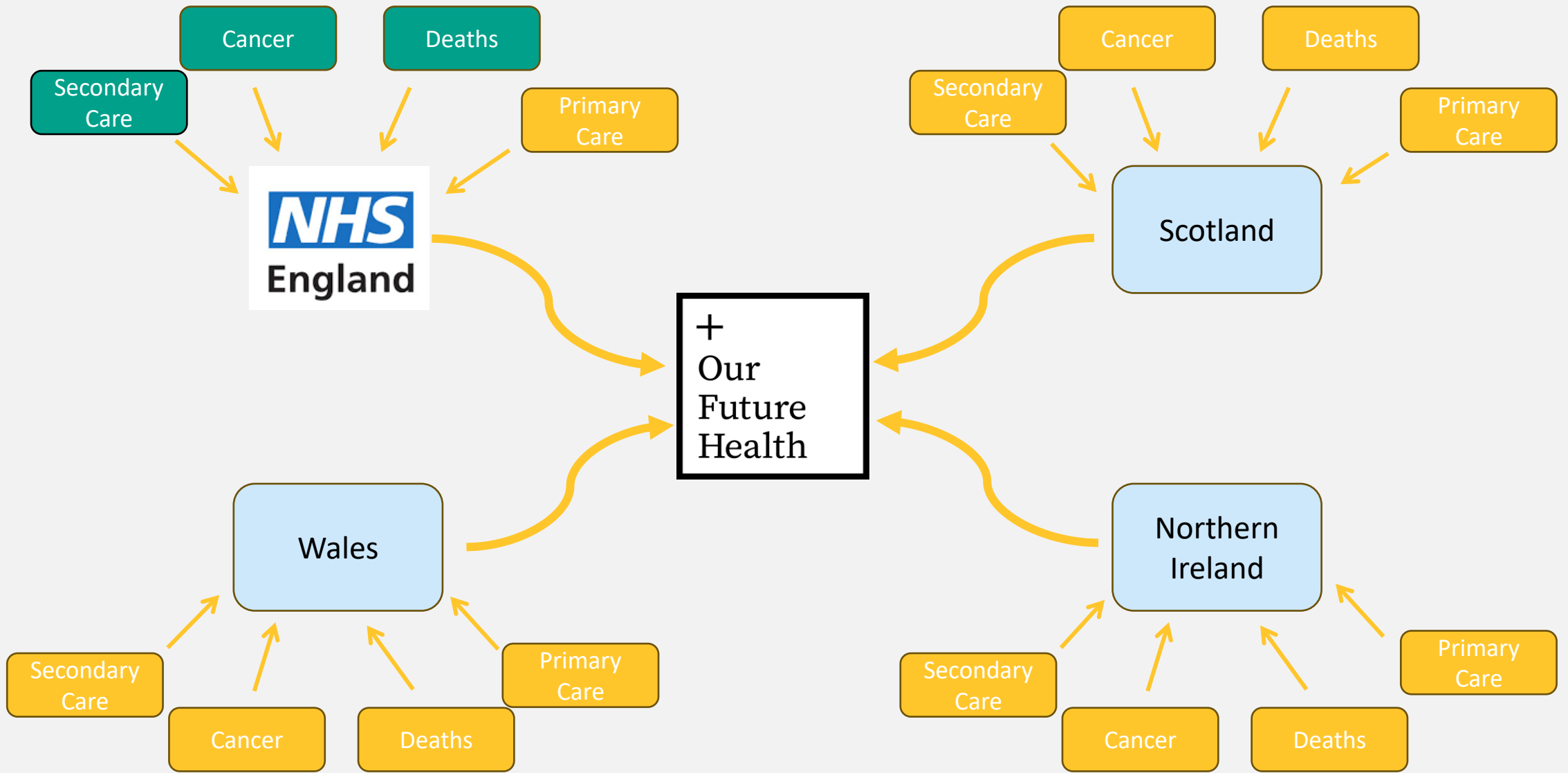
Our Future Health
Linked Participants

Participants with linked health records data (almost 34m records total)

Linked health dataset	Participants (N)	Percent (%)
Any linked record	957,444	80.3
Admitted Patient Care (HES APC)	661,810	55.5
Accident and Emergency (HES A&E)	659,716	55.3
Outpatient (HES OP)	864,360	72.5
Office of National Statistics (ONS) Death Registration	1,068	<0.1
National Disease Registration Service (NDRS) Cancer Pathways	39,462	3.3
NDRS Cancer Registry Patient Tumour	66,899	5.6
NDRS Cancer Registry Treatment	56,056	4.7

Records included: 2007-2023 (all HES); 2021-2024 (ONS deaths, post-recruitment); 2013-2023 (NDRS Cancer Pathways); 1995-2021 (NDRS Patient Tumour); 1977-2023 (NDRS Treatment)

Plans for Future Linkage





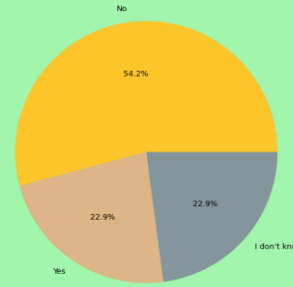
Our Future Health &
OMOP



Early work on OMOP at Our Future Health

User demand

- **23% of respondents to our TRE user survey reported they had used data models** such as OMOP
- (23% didn't know)
- OMOP was **the most frequently used data standard**
- Strong evidence of demand from other UK programmes (e.g. UK Biobank, GEL) and NHS SDE Network



Questionnaire mapping



- Preliminary exploration of OMOP and SNOMED mapping for selected data
- Some simple questions map, but more often we have specific granularity
- e.g. **Self-reported health conditions** may not align to a single point in SNOMED hierarchy (e.g. unspecified “bladder cancer”)

Existing resources for linked data

- NHS England linked data mapped to OMOP
- Could align with e.g. GEL repository, or data providers themselves
- Will need solutions for other devolved nations also (to be explored)



OMOP & opportunities for the Our Future Health resource

- One of our aims is to support common data models
 - OMOP was the most cited data models in from our user research
 - Use for replicating and validating studies
 - Importance for code and data re-use
- Extensive work has already been completed on mapping the NHS datasets
 - Researcher familiarity with mapped versions of the datasets
 - OMOP is being adopted by e.g. the NHS SDE Network
- Opportunity to collaborate with researchers to determine a strategy to begin mapping our data to maximise value.

Challenges: Data and validation

- There isn't a direct mapping for a lot of questionnaire data
 - Our baseline questionnaire has almost 300 questions, 2500 response codes
 - We also need to consider versioning due to large/small changes to the questionnaire
- We expect to need a mixture of standard, non-standard (but common), and bespoke vocabularies/coding systems
- All our data go through quality assurance processes on ingestion and for release
 - Desire to guarantee parity of results from OMOP and non-OMOP data analyses
 - Questionnaire branching logic has complex validation

Challenges: Data releases and Trusted Research Environment

- We release data on a quarterly schedule, including both questionnaire and linked data; scale and pace mean solutions need to be automatable
 - Should OMOP releases be on the same schedule?
- Dealing with dynamic nature of OMOP and data over long-term program
 - How to handle OMOP versioning?
 - Evolution of our data resource, including follow-up questionnaires, new linkages
- Incorporating OMOP data within our Trusted Research Environment (TRE)
 - Additional storage/processing requirements on already large-scale data
 - Exploration of on-TRE tooling to support analysis

Ambitions

- We are creating a world-leading resource for researchers to conduct discovery and translational research to better prevent, detect and treat disease
- We're at the beginning of our journey with OMOP, but building and growing our data amid increasing scale and demand for OMOP data in the UK and globally
- Common data models increase accessibility and research value of data, supporting our aim to help everyone live longer and healthier lives

Thank you

Our Future Health is a company limited by guarantee registered in England and Wales (number 12212468) and a charity registered with the Charity Commission for England and Wales (charity number 1189681) and OSCR, Scottish Charity Regulator (charity number SC050917). Registered office: 2 New Bailey, 6 Stanley Street, Manchester M3 5GS



Genetic Data

- Our Future Health has developed a custom genotype array in partnership with Illumina to genotype approximately 700,000 genetic variants
- Optimised for the UK population and incorporates up to date sets of disease- and phenotype-associated variants including polygenic risk score, GWAS catalog, ACMG, and ClinVar sets
- The array is also designed to predict blood types and assess pharmacogenetic variants to further enhance healthcare insights.
- We have built an integrated genotyping workflow and have contracted with Genomics plc to provide imputation and polygenic/integrated risk scores.



Physical Measurements

- Physical measurements are collected at the clinics where blood samples are collected.
- Measurements include height, weight, waist circumference, heart rate, heart rhythm, blood pressure, cholesterol
- All participants are offered a first and second reading. If both readings are irregular, participants are offered a repeat reading.
- First data release planned for December 2024

