UK Biobank data access and how apply to use the resource

Enabling scientific discoveries that improve human health

Lauren Carson
Head of Research Development
Lauren.Carson@ukbiobank.ac.uk

What is UK Biobank?

UK Biobank is a large-scale **biomedical database** and research **resource**, containing in-depth genetic and health information from **half a million** UK participants. The database is **regularly augmented** with additional data and is **globally accessible** to approved researchers undertaking vital research into the most **common and life-threatening diseases**. It is a major contributor to the **advancement of modern medicine** and treatment and has enabled several scientific discoveries that **improve human health**.

Non-profit charity, established by:

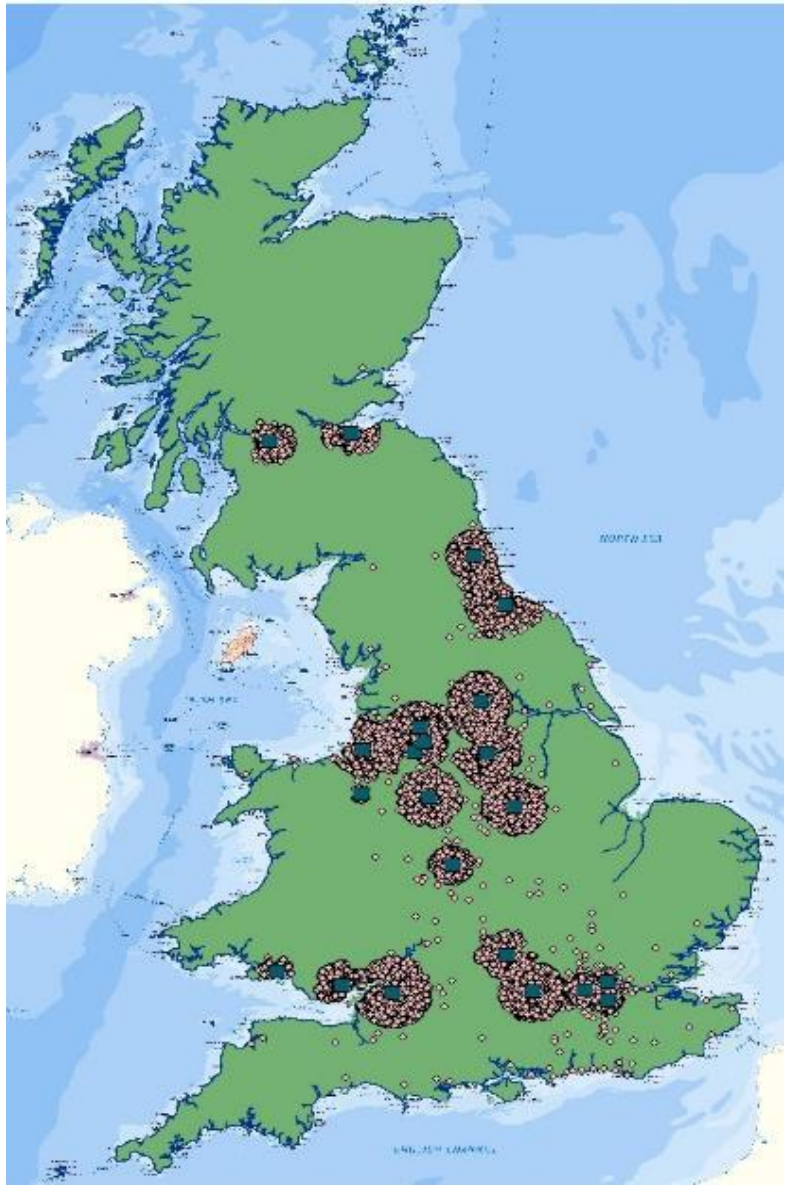
Ongoing core funding and additional funding from:

Link to 'The ground breaking UK Biobank Resource' video: https://www.youtube.com/watch?v=NGdegXRx8U0&t=152s

Link to 'What is UK Biobank?' video: https://www.youtube.com/watch?v=66moI1ZHMYs

Link to 'Celebrating 20 years of UK Biobank': https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/our-impact

Participant recruitment



- 500,000 people aged 40-69 in 2006-10 from England (89%), Scotland (7%) & Wales (4%)
- 22 assessment centres located around the UK in order to enhance the cohort's diversity (e.g. age, gender, ethnicity, socio-economic status, geographic location, urban/rural, etc)
- No selection based on disease (i.e prospective)
- Comprehensive baseline data collection
  - Lifestyle and environmental factors
  - Personal and family medical history
  - Cognitive function, hearing & vision tests
  - Physical measures (e.g. BMI, BP, FEV/FVC)
  - Biological samples (blood, urine, saliva)
- Consent to access all medical and other health-related records, and to re-contact participants for further assessments

Unique combination of BREADTH and DEPTH (and, now, DURATION)

Value increases over time

- More data and linkage is gathered from participants every year
- As participants age, more incident cases of common disease

Condition	Year of diagnosis		
	Observed	Predicted	
	2020	2027	2032
Diabetes	31,000	54,000	70,000
Myocardial infarction	15,000	30,000	46,000
Stroke	12,000	25,000	37,000
COPD	25,000	47,000	65,000
Depression	25,000	39,000	47,000
Breast cancer	9,000	14,000	18,000
Colorectal cancer	5,000	8,000	11,000
Lung cancer	4,000	6,000	8,000
Prostate cancer	10,000	16,000	20,000
Hip fracture	5,000	13,000	22,000
Rheumatoid arthritis	4,000	6,000	8,000
Parkinson's disease	4,000	10,000	14,000
Alzheimer's disease	5,000	17,000	37,000

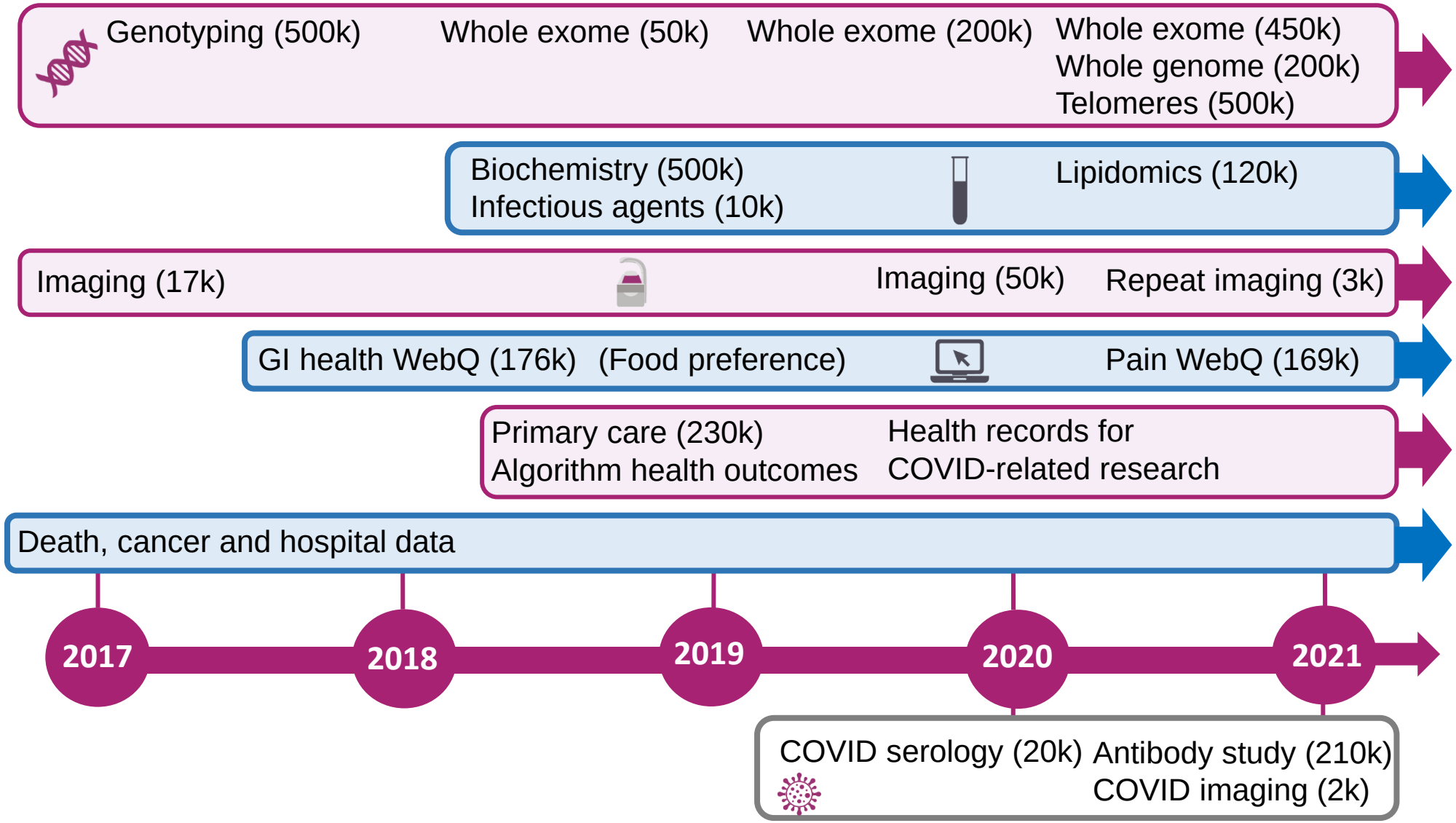
Data type	Period of complete follow-up
Death	2006 – 2021
Cancer	1957 – 2019
Hospital admissions	1997 – 2021
Primary care	1938 – 2017
for COVID research	1938 – 2021
SARS-CoV-2 antigen tests	2020 –2021

Web-based questionnaires	Number of participants	Year of data collection
Cognitive function	121,000	2015
Mental health	158,000	2016
Gastrointestinal health	176,000	2017
Pain	169,000	2018

- Potential to expand linkage further as more records become digitised

Data releases and enhancements over the last 5 years

Data releases



Opportunities to turn samples into data

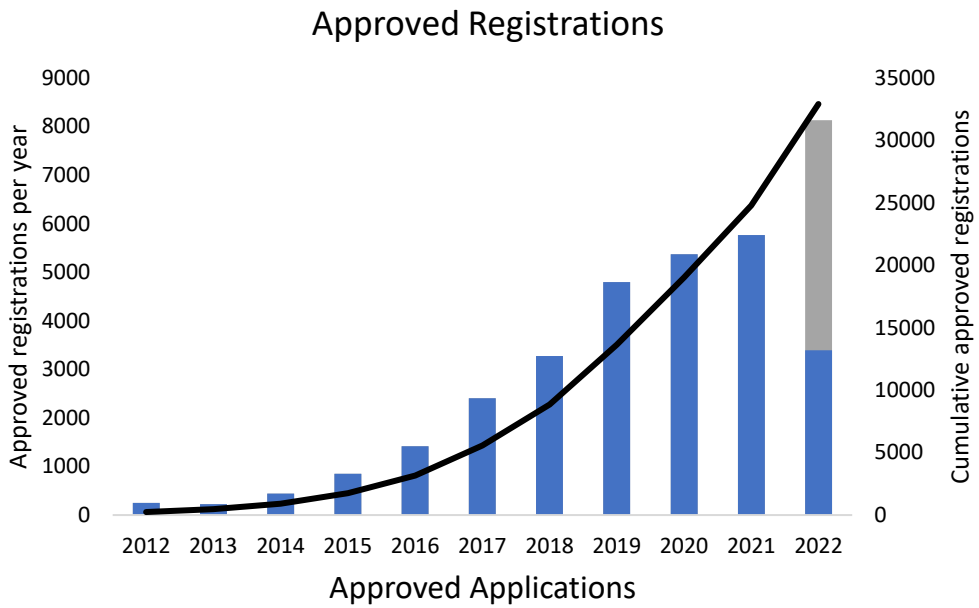
- Assays have started across the –omics and demand expected to continue over coming years as pilots turn to whole cohort studies with increased momentum from industry partners
- Increased interest in longitudinal samples
- Expect increase in use of complementary –omics platforms alongside untargeted approaches for discovery

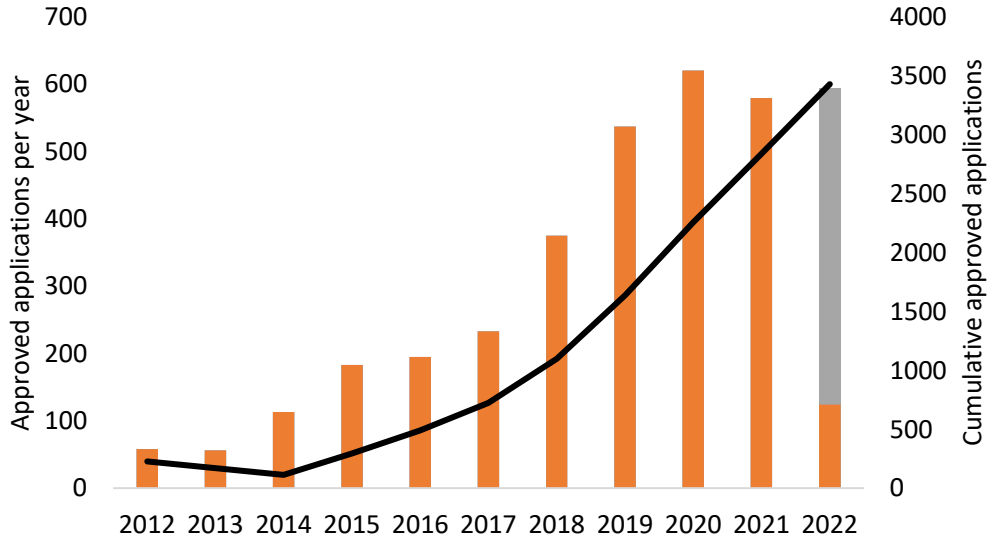
Calico
Genentech
Alnylam PHARMACEUTICALS
novo nordisk
Takeda
Pfizer

Pharma Proteomics Project (PPP)

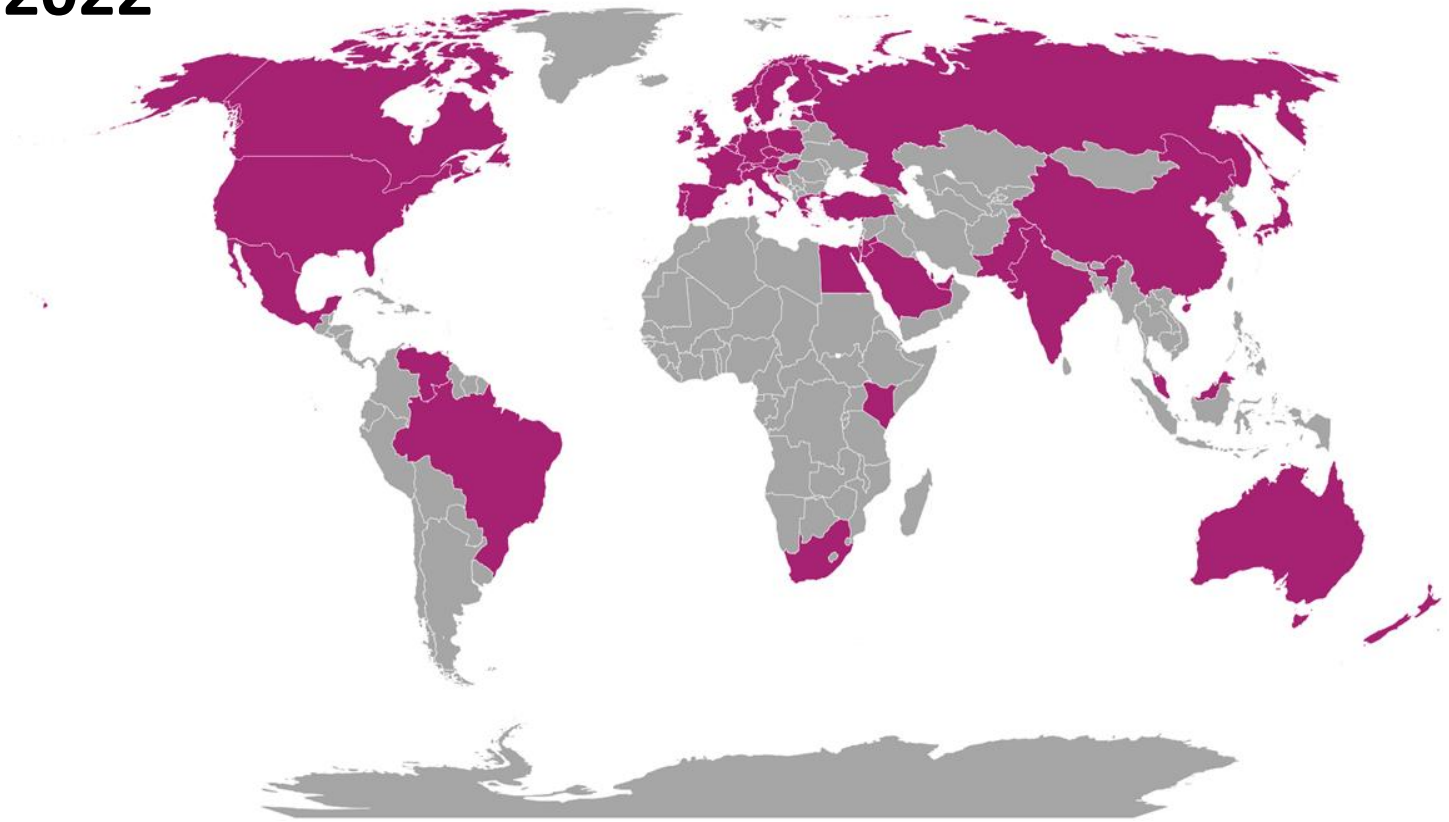
Biogen
Bristol-Myers Squibb
REGENERON
gsk GlaxoSmithKline
AstraZeneca
Janssen Pharmaceutical Companies of Johnson & Johnson
AMGEN

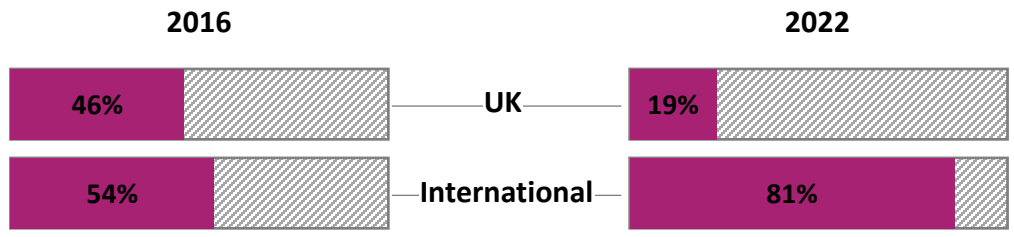
	2022	2023 onwards
Metabolome	Nightingale 500,000 plasma ~250 metabolites	10,000's samples Pilots 500,000 samples Targeted assays 500,000 samples Untargeted assays
Proteome	Olink 62,000 plasma 3k proteins Imperial College London 2,500 plasma Neurobiomarkers	
Transcriptome	wellcome sanger institute Chan Zuckerberg Initiative 5,000 whole blood scRNA sequencing	Repeat visit (60k) participants scRNA sequencing


Approved Registrations
Approved registrations per year
Cumulative approved registrations
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
35000
30000
25000
20000
15000
10000
5000
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022


Approved Applications
Approved applications per year
Cumulative approved applications
700
600
500
400
300
200
100
0
4000
3500
3000
2500
2000
1500
1000
500
0
2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022

2022

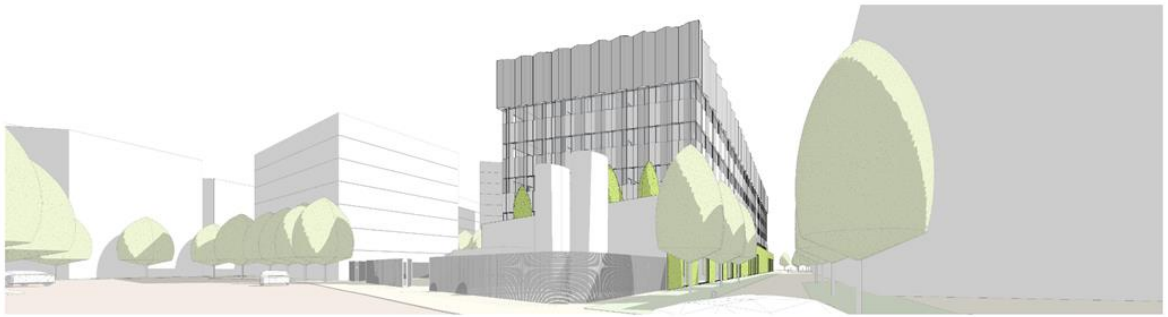



2016
2022
46%
UK
19%
54%
International
81%

- Recognised by UKRI as a part of national research infrastructure with £127.6m award for the next phase of our development
- Will include replacement infrastructure (to increase our responsiveness to sample access ideas)
- New facility in Manchester (to make us more visible and accessible to users)
- Development of new links to environmental and social sciences data
- An ‘industry hub’ to encourage liaison

Perspective one

Perspective two



UK Biobank: Principles of Access

- Available to all bona fide researchers for all types of health-related research that is in public interest.
- No preferential or exclusive access; use of the Resource on same basis for academic or commercial researchers.
- Researchers only have to pay for the costs of using the Resource (and not for any of the costs of setting it up).
- Access to the biological samples that are limited and depletable are carefully controlled and coordinated.
- Researchers are required to publish their findings and return the data so that other researchers can use them (including data from analyses of samples, images, etc.).

biobank uk


1 | REGISTER
biobank
Register


2 | APPLY
Apply for access


3 | MANAGE
RESULTS
NEWS
NEWS
Manage your project

Step 1: Go to AMS - https://bbams.ndph.ox.ac.uk/ams/.

Step 2: Complete the online form including:

- CV or link to a personal profile at your institute.
- PubMed References (if available).
- Email account at your institute.

Step 3: Submit your registration.

Step 4: Access Team will review and approve your registration.

Step 1: Complete the application form and the data you require.

Step 2: Add researchers who will access UK Biobank data to the application as collaborators.

Step 3: Add contact information for a signatory who is authorised to sign the MTA on behalf of each institute.

Step 4: The UK Biobank team will review your application.

Step 5: On approval of your application, an access fee request will be raised, and the MTA will be sent to you and your authorised signatory.

Step 6: Following payment of the access fee and return of signed MTA, the Access team will inform you when your data have been released and are available to access.

biobank uk

Description	Tier 1	Tier 2	Tier 3
Core data • Questionnaires and physical measurements • Linked health data • Health Outcome phenotypes • Web-based questionnaires	✓	✓	✓
Assay data and enhanced measures * • Biochemical and haematological assays • Measured and imputed genotypes • Other platform based assays • Other enhancements		✓	✓
Very large datasets • Imaging data ** • Whole genome sequence data • Other large-scale assay data • Whole exome sequence data			✓ Via platform only
First 3 years - access to data with scheduled updates	£3,000	£6,000 (+£3,000 vs Tier 1)	£9,000 (+£3,000 vs Tier 2)
Additional Institution fee - each additional institution added to an application	£1,000 for first 3 years (£500 p.a. extension)		
Low & Middle Income Countries and Student Researchers *** - access to all datasets via the Research Analysis Platform (full fees apply to downloaded data)	£500 for first 3 years (£175 p.a. extension)		

Democratising access and the Research Analysis Platform

https://ukbiobank.dnanexus.com/

- Cloud-based Platform; secure and audited access enabled by DNAnexus and powered by AWS hosted in London
- Bringing the researcher to the data; making data access & compute more widely accessible

1400+

Number of Users

300+

Number of Applications

The UK Biobank Platform Credits Programme

aws

Step 1: Keep your contact details up to date.

Step 2: Submit extensions to your projects as required.

Step 3: Provide Annual Project Reports or Annual Confirmation Forms as required by the MTA.

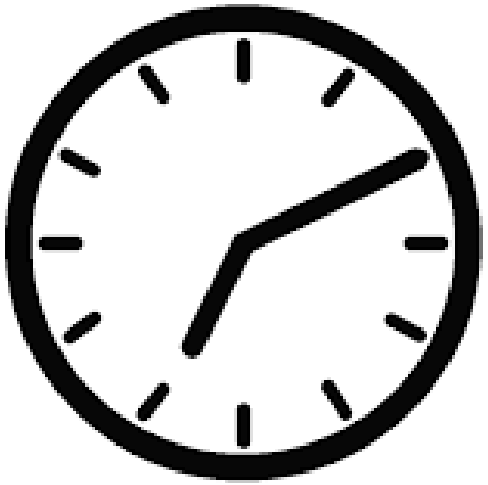
Step 4: Alert us to publications.

Step 5: Respond quickly to audit requests.

Step 6: Return your results.

Tips to minimise the application timeline:

- Ensure your legal team has seen our standard MTA at the point of applying.
- Amend your application in response to our feedback as soon as you can.
- Pay the access fee online via AMS.
- Please liaise with your Finance Departments to resolve any issues with online payment to ensure prompt payment.
- Sign the MTA we send to you and your authorised signatory (via DocuSign) as soon as you can.



You can always contact us via AMS or email to access@ukbiobank.ac.uk

Questions?