UK Biobank – an update for our participants

Assoc Prof Naomi Allen
UK Biobank Senior Epidemiologist
What is UK Biobank?

- A readily accessible resource for scientists to study the causes of health and disease in middle and old age
- 500,000 UK adults age 40-69 at recruitment
- Information on lifestyle, environment, medical history, physical measures & biological samples

And then…..

- Additional data collection and analyses of samples
- Participants followed-up for many years to find out about their health
503,000 participants
40-69 years old
22 recruitment centres
89% England
7% Scotland
4% Wales
Baseline assessment

Touchscreen questions
- Socio-demographic information
- Lifestyle factors (diet, physical activity, smoking, sleep)
- Environmental exposures
- Reproductive history & screening
- Sexual history
- Family history of common diseases
- General health & medical history

Large subsets
- Noise exposure
- Psychological status
- Cognitive function tests
- Hearing test

Physical measures
- Blood pressure
- Hand grip strength
- Body composition
- Lung function test
- Heel ultrasound

Large subsets
- Vascular reactivity
- Exercise test/ECG
- Eye measures (visual acuity, refractive error, OCT scan)

Biological samples
- Blood, urine, saliva
Comparative images of retinal detachment: Retinal Photograph versus Optical Coherence Tomogram
Using machine learning techniques, the anatomy of the retinal blood vessels can predict your:

- Age (within 3 years)
- Sex (97% of the time)
- Smoking status (71%)
- Blood pressure (within 11 mmHg)
- Heart problems (70%)
Numbers of participants by socio-demographic factors

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Category</th>
<th>Numbers (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>40-49</td>
<td>119,000 (24%)</td>
</tr>
<tr>
<td></td>
<td>50-59</td>
<td>168,000 (34%)</td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>213,000 (42%)</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>228,000 (46%)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>270,000 (54%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>White</td>
<td>473,000 (95%)</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>27,000 (5%)</td>
</tr>
<tr>
<td>Deprivation</td>
<td>More</td>
<td>92,000 (18%)</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>166,000 (33%)</td>
</tr>
<tr>
<td></td>
<td>Less</td>
<td>241,000 (46%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>500,000</td>
</tr>
</tbody>
</table>

Generally better off than UK average but a wide range of backgrounds represented
Industrial scale processes needed: UK Biobank samples during recruitment

- 700 participants per day
- 4,900 sample tubes per day
- 25,000 aliquots produced per day
- 15 million 0.85ml aliquots
• Blood
• Urine
• Saliva

Total > 15 million aliquots
– Measurement of 850,000 parts of the genome
– Available for research July 2017
– Exome sequencing now underway
  • $100M investment from pharma
  • Measures parts of the genes that make proteins (~2% of the genome)
  • Could be important to identify targets for new drugs
  • Available for researchers in 2 years time
Largest study to date in the area identifies genetic variants associated with chronic obstructive pulmonary disease (COPD)

Spirometry measurements also enabled the researchers to explore different aspects of lung function.

Those in the highest risk group based on genetic susceptibility were 4 times more likely to develop COPD compared to those in the lowest risk group.

By quitting smoking in early adulthood, COPD can be reduced in 5 out of 10 people at high genetic risk.
Measurement of 36 blood and urine biomarkers of most relevance to the research community

• Established risk factors for disease
  – e.g. lipids for vascular disease; sex hormones for cancer

• Diagnostic markers
  – e.g. HbA1c and glucose for diabetes; Rheumatoid Factor for arthritis

• To assess body function
  – e.g. biomarkers for renal and liver function
Other planned measures: external investment

- **Telomere length in all 500,000**
  - Funded by MRC, BHF, BBRSC
  - Important in ageing

- **Blood metabolites in all 500,000**
  - Funded by Nightingale, Finland
  - 200+ circulating lipids and amino acids

- **Whole-genome sequencing in all 500,000?**
  - Measuring all 3B ‘letters’ or ‘base-pairs’ of the genome
  - Will massively extend research into personalised medicine
Accelerometry data collected on 100,000 participants
- Worn continuously for 7 days
- Mailed in the post and returned to UK Biobank

- Started repeat collection in 2,500 participants
- To collect 4 times over a year to assess seasonal change in activity
Variation in physical activity across the day by age and sex in 100,000 UK Biobank participants

Published in Doherty et al., Plos One, 2017
Much more than a biobank....

Size +++
Lifestyle measures +++
Biochemical measures +++
Genotyping +++

and

PROSPECTIVE....i.e. follow-up (for decades)

Diseases    Disability    Quality of life    Death
Following the health of half a million people

All participants:
  registered with a GP in the NHS
  consented to linkage to health-related records

NHS provides majority of healthcare in UK

National datasets about healthcare & health outcomes exist

.....so link to these datasets.....
Follow-up through health records

Key national sources:

- Death registrations ✓
- Cancer registrations ✓
- Hospital episode data ✓
- Primary care data (✓)
Expected numbers of people developing disease during follow-up: some examples

<table>
<thead>
<tr>
<th>Condition</th>
<th>2012</th>
<th>2017</th>
<th>2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>10,000</td>
<td>25,000</td>
<td>40,000</td>
</tr>
<tr>
<td>Heart attack</td>
<td>7,000</td>
<td>17,000</td>
<td>28,000</td>
</tr>
<tr>
<td>Stroke</td>
<td>2,000</td>
<td>5,000</td>
<td>9,000</td>
</tr>
<tr>
<td>Chronic obstructive lung disease</td>
<td>3,000</td>
<td>8,000</td>
<td>14,000</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>2,500</td>
<td>6,000</td>
<td>10,000</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>1,500</td>
<td>3,500</td>
<td>7,000</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>1,500</td>
<td>3,500</td>
<td>7,000</td>
</tr>
<tr>
<td>Hip fracture</td>
<td>1,000</td>
<td>2,500</td>
<td>6,000</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>1,000</td>
<td>3,000</td>
<td>9,000</td>
</tr>
</tbody>
</table>
Follow-up through questionnaires

- Email for 2/3 of the cohort (330,000)

- To assess lifestyle and health outcomes we can’t find out about through linking to health records:
  - Details of dietary intake
  - Occupational history
  - Cognitive function
  - Mental health (thoughts and feelings)
  - Gastrointestinal symptoms
How many researchers are using UK Biobank?

>7,000 registered researchers; > 1,000 research projects

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Applications</th>
<th>International %</th>
<th>UK %</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>130</td>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>2014</td>
<td>330</td>
<td>24</td>
<td>76</td>
</tr>
<tr>
<td>2015</td>
<td>440</td>
<td>44</td>
<td>56</td>
</tr>
<tr>
<td>2016</td>
<td>550</td>
<td>58</td>
<td>42</td>
</tr>
<tr>
<td>2017</td>
<td>680</td>
<td>68</td>
<td>32</td>
</tr>
</tbody>
</table>
Where are the researchers based?

- 2230 researchers are based in Canada.
- 3115 researchers are based in the United States.
- 978 researchers are based in the United Kingdom.
- 110 researchers are based in India.
- 20 researchers are based in Brazil.
- 268 researchers are based in Australia.
What research is being done?

- Genetics
- Obesity
- Cancer
- Cardiovascular mortality
- Vision
- Physical activity
- Imaging
- Sleep
- Smoking
- Dementia
- Cognition
- Diet
- Biomarkers
- Metabolic traits
- Hypertension
- Liver disease
- Reproduction
- Osteoporosis
- COPD
- Multiple sclerosis
- Chron’s disease
- Hypertension
- Stroke
- Diabetes
- Ageing
- Ethnicity
- Heritability
- Infectious disease
- Parkinson’s
- Air pollution
- Social isolation
- Multi-morbidity
- Socio-economic factors
- Liver disease
- Oral health
- Asthma
- Haematology
- Allergy
- Primary care
- Primary care
- Metabolic traits
- Mendelian randomisation
- Brain
- Personalised medicine
- Built environment
- Psoriasis
Thank you!